SCIENCE CHINA
Information Sciences

December 2020, Vol. 63 220302:1–220302:14
https://doi.org/10.1007/s11432-019-2906-y

• REVIEW •

Special Focus on Challenges and New Insights for Network Security in 5G-and-Beyond

Security in edge-assisted Internet of Things: challenges and solutions

Shuaiqi SHEN[1], Kuan ZHANG[1], Yi ZHOU[2*] & Song CI[3]

[1]*Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, Omaha* NE 68182*, USA;*
[2]*School of Computer and Information Engineering, Henan University, Kaifeng* 475001*, China;*
[3]*Department of Electrical Engineering, Tsinghua University, Beijing* 100084*, China*

Received 2 December 2019/Revised 27 February 2020/Accepted 12 May 2020/Published online 2 November 2020

Abstract The flourish of 5th generation wireless systems (5G) network has brought numerous benefits to the Internet of Things (IoT) with universal connectivity, improved data rate, and decreased latency. The development of IoT extends the computation-intensive applications from centralized servers to the edge of the network, promoting the paradigm of edge computing. Edge computing brings great assistance to IoT architecture by efficiently accomplishing tasks with lower latency, less energy consumption and reduced network bandwidth. Edge-assisted IoT emerges to provide location-aware services and offload computational tasks to edge nodes near the IoT devices. However, a series of security challenges still exist in edge-assisted IoT owing to the inherent vulnerabilities of edge nodes and sensitive nature of the collected data. Among the emerging and existing security schemes in IoT applications, security and energy consumption are two overriding yet conflicting requirements. The tradeoff between security and energy is critical to the sustainability of the IoT ecosystem but still lacks sufficient discussion. In this article, we investigate the security challenges in edge-assisted IoT and how the constraints in energy consumption can affect the design of security schemes. Specifically, we firstly present the architecture of edge-assisted IoT and its unique characteristics. Secondly, we identify the security threats in the edge-assisted IoT applications and the security-energy tradeoff during implementation. Thirdly, in a case study of distributed denial of service (DDoS) and malware injection attack, we propose a preliminary solution to address the conflict between security and energy requirements. Finally, we discuss some open issues and identify future research directions for security and energy efficiency in edge-assisted IoT.

Keywords Internet of Things, edge computing, network security, machine learning, data analytics

Citation Shen S Q, Zhang K, Zhou Y, et al. Security in edge-assisted Internet of Things: challenges and solutions. Sci China Inf Sci, 2020, 63(12): 220302, https://doi.org/10.1007/s11432-019-2906-y

1 Introduction

Towards the era of 5th generation wireless systems (5G), Internet of Things (IoT) is gaining increasing popularity, because smart devices are expected to play a major role in the daily life. The 5G network efficiently transmits the massive volume of data at a high bandwidth, allowing information to be accessed and shared anywhere and anytime as long as a device is deployed. With the evolution of 5G technology, the number of smart devices employed in IoT increases exponentially, and billions of devices with 6 or 7 ones per sensors are expected to be connected to IoT by 2020 [1]. Devices, including wearable

[*] Corresponding author (email: zhouyi@henu.edu.cn)

equipments, versatile sensors, mobile phones and embedded systems, work together in IoT architecture to collect, process and analyze large amounts of data for automated decision making. A wide range of IoT applications, such as smart city, smart home, smart grid, e-healthcare and intelligent transportation, can provide better interactions among things and human for a quality life [2].

Traditional IoT architecture places huge amount of devices throughout the community and connects them to a centralized data center. Because most IoT devices lack enough computational power and storage capacity to process the data, the centralized data center is expected to receive all the generated data including the redundant or insignificant part. The overloaded volume and frequency of the received data can cause a waste in computation and storage resources of the data center, and make it difficult to sift through the unnecessary information to find kernels of insight. With the assistance of edge computing, IoT alleviates the insufficiency of computational power and storage capacity in connected devices by passing complex tasks to edge nodes [3]. As the devices may generate a great volume of data from various services and applications, traditional centralized servers of IoT bear heavy overhead on data transmission, processing and storage. Edge-assisted IoT offloads the burden to the edge of the network, so that a large amount of bandwidth can be saved and latency can be reduced owing to its proximity to local devices. By applying a hierarchy of decentralized edge nodes with better computation and storage capabilities, edge computing can handle mobile and heterogeneous tasks required by IoT devices and applications [4]. As a result, edge-assisted IoT provides location-aware, bandwidth-sufficient, real-time, and low-cost services to support the ever-growing data-intensive and latency-sensitive applications.

However, the edge-assisted IoT brings security threats owing to its distinctive vulnerabilities and features [5]. For instance, data collected from sensor devices may be falsified by attackers to manipulate the results of decision and control in IoT applications. Edge nodes may be compromised or hacked by the attackers in vicinity to launch malicious attacks, such as distributed denial of service (DDoS) aiming to prevent the legitimate use of a service. In addition, IoT services, such as e-healthcare and smart home, collect and manage user's private information, taking the risk of being disclosed by malicious attackers. The drastic rising trend of attacks targeting edge computing infrastructures in recent years [6] makes it urgent to develop effective security mechanisms against those threats. Besides the security requirements, the energy constraints of IoT devices bring more challenges to the development of security schemes in edge-assisted IoT. Considering the massive amount of devices and servers included in IoT, the energy conservation achieved from optimizing security schemes can make a huge difference in the business of IoT applications. In existing studies, these two disciplines are quite isolated and the techniques developed for one rarely address the issues in the other. The tradeoff between energy and security requires more efforts for security issues.

In this paper, we investigate the challenges of enhancing security in edge-assisted IoT and the new requirements for security schemes taking the constraints of energy consumption into consideration. Specifically, the major contributions in this paper are fourfold.

- Firstly, we present the architecture of edge-assisted IoT, its key characteristics, and some promising services and applications. We discuss the challenges of security protection in edge-assisted IoT.
- Secondly, we investigate the tradeoff between energy consumption and security in IoT security architecture and analyze several possible standards to distinguish the requirements of security levels for the optimization of energy consumption.
- Thirdly, we discuss a case study of DDoS and malware injection attack, and then propose a preliminary solution to address security-energy tradeoff through feature selection based on nonadditive measure theory. Simulation results validate that the proposed solution can customize the computational complexity of machine learning based security schemes, reaching a balance between the security demand and energy consumption.
- Finally, we discuss some open issues. We also identify future research directions for the security and energy efficiency of edge-assisted IoT.

The remainder of this article is organized as follows. In Section 2, we illustrate the edge-assisted IoT architecture, applications and characteristics. Security challenges and the security-energy tradeoff are discussed in Section 3. In Section 4, we illustrate how to achieve energy efficiency by adapting the

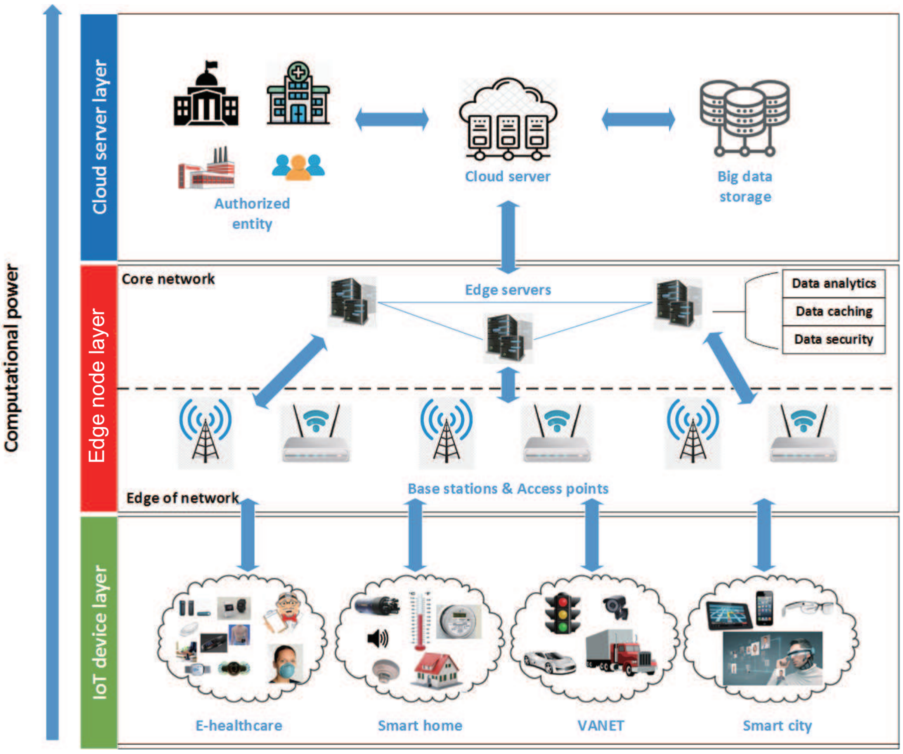

Figure 1 (Color online) Edge-assisted IoT architecture.

complexity of security schemes to various security requirements in a case study of DDoS and malware injection attack. In Section 5, open issues and future research directions are pointed out. Finally, a conclusion is drawn in Section 6.

2 Edge-assisted IoT architecture, applications, and characteristics

Edge-assisted IoT integrates emerging edge nodes with IoT devices, core networks and centralized computing units, connecting the physical world and cyber world. In this section, we present the edge-assisted IoT architecture, applications, and key characteristics.

2.1 Edge-assisted IoT architecture

As illustrated in Figure 1, there are three layers in edge-assisted IoT architecture, i.e., IoT device layer, edge node layer, and cloud server layer. The computational power increases from bottom to top.

2.1.1 *IoT device layer*

Devices deployed in this layer consist of a wide range of industrial sensors, wearable devices, smart meters, smartphones, and video surveillance cameras. They are usually with limited battery, computing, storage and communication capabilities, and are used to measure the information from the objects in physical world and transmit it to the edge nodes. IoT devices can also be controllers in reality, such as servo actuator, navigator or smartphone, leveraging the optimal decisions from the centralized computing unit and control objects to make operations. In edge-assisted IoT architecture, the IoT device layer serves as the bridge linking the physical and cyber world.

2.1.2 *Edge node layer*

The edge node layer is a key component in IoT linking the massive IoT devices to the centralized computing units. Edge nodes are interconnected by heterogeneous networks, incorporating cellular network, wide area network (WAN), wireless local area networks (WLANs), device-to-device (D2D) communications, etc. The seamless switch among different types of networks and devices is supported in edge-assisted IoT. Edge nodes collect the data generated from IoT devices for data processing and send back control flows. If the task complexity exceeds the computation capacity of current edge node level, it offloads the task to edge nodes at higher level until the task reaches cloud servers.

2.1.3 *Cloud server layer*

The cloud server layer hosts the centralized computing unit, such as cloud servers and data centers with powerful computing capability and big data storage. This layer is responsible for the highest level authentication, processing, and integration of different tasks offloaded from edge nodes. It analyzes the data from IoT devices, collaborates with edge nodes to process the data, and makes decision or feedback control to IoT devices.

2.2 Edge-assisted IoT applications

Edge-assisted IoT expands the traditional IoT services into a variety of aspects. A wide range of IoT applications have shown promising perspectives, in which the security and energy consumption are both highly critical issues.

2.2.1 *Intelligent transportation*

Cameras in vehicles or along the road can capture numerous video data for road condition and traffic, which is real-timely processed at the edge of IoT to provide proper driving decision, navigation or intelligent transportation management. Edge nodes take over the majority of the video processing tasks from the centralized computing servers, such that the response time can be reduced, while the network bandwidth is saved from a huge burden of transmitting all video data. In 2020, a live traffic and traffic management information service called “TidalWave” has been deployed in nine cities across the USA powered by edge computing and machine learning techniques. TidalWave collects the conditions of moving vehicles from drivers’ smartphone and offloads data processing to the edge nodes in local area. With the assistance of edge computing, the service delivers high resolution and accurate live data for city intersections and traffic locations, providing results in less than a second after actual traffic changes, so that optimal vehicle routing can be achieved.

2.2.2 *Video streaming and gaming*

Video streaming from the edge can support a series of IoT services, especially location-aware applications. Edge nodes cache or stream the video clips toward the specific location. Then, users are able to directly access these clips from edge nodes, such that the end-to-end latency is shortened compared with the old-fashion downloading from the remote/centralized servers. Companies such as Netflix has launched its ubiquitous online-streaming platform to deliver video contents. Google, Microsoft and other heavyweight hardware companies such as Sony have started streaming games on the cloud.

2.2.3 *Location-aware monitoring*

Edge-assisted IoT facilitates location-aware applications, such as environmental and energy monitoring. The environmental monitoring analyzes a series of measurement data, such as air and water pollution, greenhouse gas emissions and urban noise, mostly at the edge of IoT to assess the environment condition and afford the intelligent control. Energy monitoring, on the other hand, keeps track of the energy generation, transmission, distribution and consumption in a community, which can not only facilitate energy conservation in many aspects, but also prevent the blackout of power grid and deficiency in individual energy usage. Edge-assisted IoT utilizes edge nodes to collect and process data generated by the widely-deployed sensors in the community, moving the computation power closer to the sensor layer.

In this case, edge computing provides immediacy and location-awareness of the delivered results to the monitoring applications.

2.3 Characteristics

Under this edge-assisted IoT architecture, there are several key characteristics as follows.

2.3.1 *Low latency*

The first-tier feature of edge-assisted IoT architecture is the reduced latency, since edge nodes facilitate data processing offloaded from the centralized servers to the edge of network, which is closer to local devices. The bi-directional communication between IoT device and an edge node may take approximately milliseconds, while conversing in the same manner with the cloud may take minutes. Furthermore, the pre-processed data at edge nodes improve the efficiency of data transmission and processing at cloud layer. For the distributed IoT services, edge nodes can respond in real time such that the delay can be dramatically reduced.

2.3.2 *Location awareness*

Edge nodes are deployed in the proximity of IoT devices, users, and the service area, extending the centralized computing and processing to where data are collected. Location awareness acquires more information from IoT devices such as the coordinates in a grid and distances to the local edge node, which enable more IoT applications including intelligent navigation, surveying, and warehouse routing.

2.3.3 *Heterogeneity*

Edge nodes bridge the cloud servers to various types of IoT devices, offering the compatibility in IoT services. In edge-assisted IoT, edge nodes could be different parties: (1) wireless carriers or the Internet service providers, who may incorporate edges with their existing infrastructures including cellular network, WAN, WLANs and D2D communications; (2) cloud service providers, who want to expand their cloud services from centralized computing units to the edge of the IoT, may build their dedicated edge infrastructures; and (3) users or devices, who have a local powerful computing capabilities (e.g., private cloud) and want to make profits, may turn the local private cloud into edge and lease spare resources.

2.3.4 *Offloading*

Edge-assisted IoT enables computation to be offloaded at the network edge such that a portion of computing tasks can be performed locally near the data sources. By offloading the computation-intensive workloads to the edge nodes, the latency is dramatically reduced, and the computational quality of IoT services is enhanced.

3 Security and energy challenges

When edge-assisted IoT is expanded from the traditional cloud-based one, security issues in traditional IoT are still challenging [7]. The inherent characteristics of edge, such as location awareness and computation offloading flexibility further complicates the security situation. Furthermore, the energy constraints require massive IoT devices and edge nodes to perform on a thrifty energy consumption while maintaining sufficient security level. The tradeoff between energy and security should be addressed to enhance the sustainability of IoT architecture and make IoT applications to be cost-effective. In this section, we point out the security threats in edge-assisted IoT and discuss the challenges brought by energy issues to security schemes.

3.1 Security threats in edge-assisted IoT

The local data storage decreases the dependency on Internet connections and local data exchange reduces the possibility of data exposure. However, edge-assisted IoT still faces a variety of security threats because of the limited resources and the drawbacks in secure communication protocols [8]. The edge layer receives data generated from massive IoT devices to provide local services by sending back control flows, and may also transmit the collected data to cloud server for task integration, further analysis and permanent storage. For example, crowdsensing outsources sensing tasks to a crowd, where a group of IoT devices interconnected by the edge of network help the task publishers to report the local area information in real time. The generated data and processing results are aggregated in edge nodes to measure, analyze or estimate the common interest of the crowd. In edge-assisted IoT applications, edge nodes are in charge of most core computing functions, such as authentication, authorization, data analysis, task offloading, and data storage. Therefore, rogue or compromised edge nodes can bring serious security threats to edge-assisted IoT applications, leading to privacy leakage, data corruption, malware injection and service shutdown.

In addition, the heterogeneity of edge-assisted IoT raises in service, data and devices. In terms of security, each service requires a specific configuration of policies and parameters based on the distinct requirements. Some may deliver produced data frequently and some may need rapid authentication responses. To recognize the identities of accessing devices, service-oriented identity authentication is needed by edge nodes. However, it is difficult to find a "one-size-fit-all" authentication approach to support all IoT services. Furthermore, IoT devices usually form a group to communicate with each other and upload data to edge node together. This group is not stable, because some old or malfunctioning devices may leave the group and new devices may join, which leads to huge obstacles on group identity authentication. The challenge in implementing efficient authentication mechanism needs to be resolved; otherwise compromised devices and edge nodes may be involved in IoT services to launch cyber attacks.

3.2 Security-energy tradeoff

To protect edge-assisted IoT against the aforementioned threats, various security schemes have been proposed [9]. Most of the state-of-art schemes [10–12] are based on machine learning techniques to detect malicious behaviors of attackers from normal activities in an autonomous and self-evolving way. Despite the powerful characteristics of machine learning techniques in detecting security threats, their high computational cost is still a serious concern when deployed in edge-assisted IoT. Many IoT devices and edge nodes are required to perform with thrifty energy consumption for longer longevity of the energy source, such as battery. Besides, some wearable devices have low energy requirement because high energy consumption can cause the device to heat up beyond the human tolerance limit, making the device unusable. Preserving energy is also beneficial from the business aspect. As edge-assisted IoT consists of massive devices and edge nodes, even slight energy conservation achieved from the development of security schemes can save massive amount of energy expenses for IoT applications. However, the demands for security and energy efficiency are usually in conflict. The tradeoff between security and energy challenges the security protection of edge-assisted IoT in two aspects, which are side channel attack enabled by energy optimization and the determination of security demands in different scenarios.

3.2.1 *Vulnerability to side channel attack*

In order to preserve energy consumption, specific hardware components in IoT devices are often designed to optimize power in frequently occurring cases. These hardware components keep track of the power consumption as log files. However, the log files may also be monitored by attackers. By profiling different patterns of power consumption, the attacker can launch a side channel attack and apply machine learning algorithm to recognize the operations executed in the device [13]. In this case, the attacker may identify the internal execution information of the IoT device via the vulnerability of power-optimization hardware.

To prevent side channel attack, hardware-level energy optimization may not be the best option in IoT devices and edge nodes.

3.2.2 *Variation in security demands*

As discussed above, the hardware-level optimization may not be the most practical approach to preserve energy consumption owing to the vulnerability of side channel attack. In this paper, we propose a preliminary solution to achieve security-energy tradeoff in software level. The conflict between security and energy can be addressed through the adjustment of computational complexity of the deployed security schemes. For machine learning based schemes, increasing algorithm complexity leads to higher accuracy in authenticating, detecting, and monitoring. On the other hand, the increasing complexity also rises energy consumption of the device. To address security-energy tradeoff, the complexity of security scheme needs to be adjusted to achieve reasonable energy consumption while maintaining a sufficient degree of security. This leads to the challenge on how to determine the security demands according to different scenarios. We present four aspects to evaluate the security demands in edge-assisted IoT.

- Application features. The intrinsic features of IoT applications are critical to the demands for security schemes. For example, e-healthcare stores a lot of private health information and medical data in edge nodes. The user security can be seriously threatened if the edge nodes are compromised, so that the demand for security should be satisfied over the demand for energy conservation [14]. On the other hand, video streaming and gaming involves less sensitive data in most cases, so that security demand can be weakened for preserving energy consumption.
- Characteristics of devices. Owing to the heterogeneity of edge-assisted IoT, the interconnected devices have a variety of characteristics that require different degree of security and energy consumption. For example, a smart phone has more complex operating system than a sensor, such that the smart phone can be more vulnerable to various types of attacks. Smart phones also store sensitive and private data, and thus deserve higher level of security. In addition, IoT devices have various kinds of power sources that lead to different energy constraints. For example, a video surveillance camera has little energy constraint as it is charged through power lines, while sensors and other small devices rely on wireless power sources such as battery and solar energy. Those devices with wireless power sources have to reduce energy consumption by sacrificing security to guarantee their sustainability in IoT applications.
- Vulnerable attacks. Different IoT applications, devices and servers are vulnerable to various cyber attacks. Depending on the vulnerabilities to different attacks, the security demand may also be customized manually. Some users, for instance, may consider man-in-the-middle attack to be more dangerous than DDoS attack. A man-in-the-middle attacker can obtain complete control of all communication between users to capture or falsify data, while DDoS attack only blocks user's access to IoT service but the data are still secured. Therefore for these users, the applications or devices vulnerable to man-in-the-middle attack should demand higher security level than DDoS attack.
- Frequency of invoking. The tradeoff between security and energy is also influenced by the frequency of invoking a security scheme. For example, if a sensor network for environmental monitoring collects physical data once a week, then the application has less concern on energy cost. For an edge node handling massive data transmission, however, the security scheme has to be invoked frequently and requires lower energy cost by reducing the security demand.

4 A case study to address security-energy tradeoff

As discussed in Section 3, different IoT applications and devices have a variety of security demands, which can be distinguished through the aforementioned aspects. The security-energy tradeoff for edge-assisted IoT can be addressed by adapting the complexity of security schemes to a specific security level, so that the energy consumption of the security scheme can be conserved. For machine learning based security schemes, one of the approaches to adjust their computational complexity is to change the number of features [9]. In this section, we propose a preliminary solution to address the security-energy tradeoff

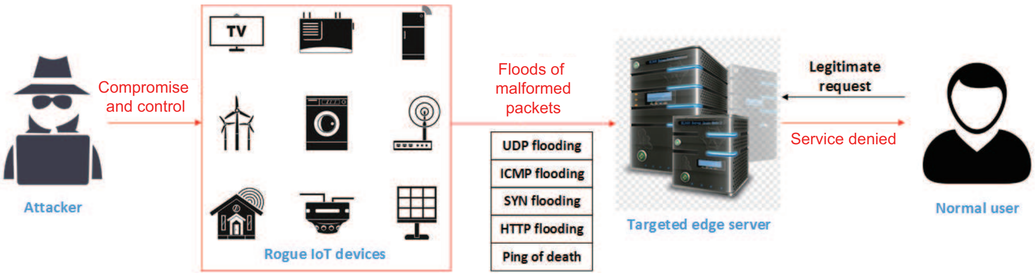

Figure 2 (Color online) System model of typical DDoS attack.

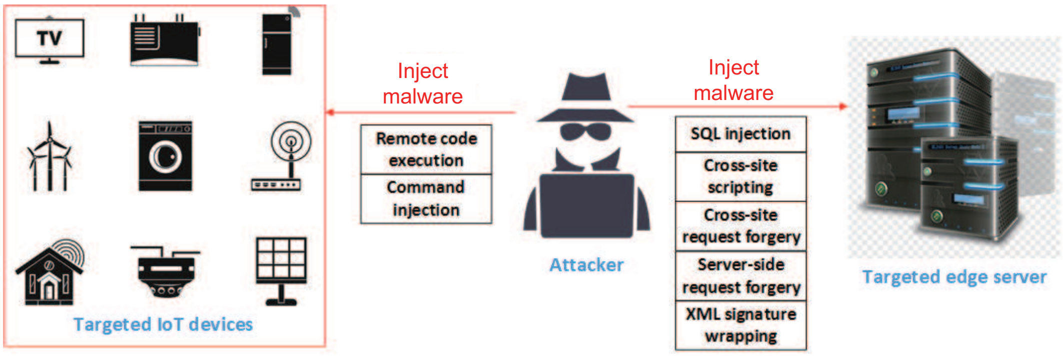

Figure 3 (Color online) System model of typical malware injection attack.

by performing feature selection to adjust the number of features in machine learning based security schemes. A case study of DDoS attack and malware injection attack is investigated to demonstrate how the complexity of security schemes can be adjusted flexibly with feature selection solution.

4.1 Case study of DDoS and malware injection attack

4.1.1 *Attack models*

The major security challenges faced by edge-assisted IoT include DDoS attack and malware injection attack. The system models of two types of attacks are presented in Figures 2 and 3. DDoS attacks may occur when rogue IoT devices are connected with the edge nodes. The attacker compromises a group of IoT devices to take their control, and then commands those devices to launch denial-of-service attack towards the target edge node to terminate its services [15]. The attack can be launched by either flooding large amount of malicious network packets, or finding an unknown vulnerability in the code running on the target node to corrupt its program. On the other hand, a malware injection attack [16] is also a major threat to the data confidentiality and integrity of edge-assisted IoT. Owing to the limitation of computational power, IoT devices and low-level edge nodes can barely be protected by traditional firewall, making them more vulnerable to both client-side and server-side injection attack.

4.1.2 *Related work*

For security in edge-assisted IoT, machine learning based schemes become increasingly popular, because they are capable of detecting cyber attacks from normal activities in an autonomous and self-evolving way. Livadas et al. [17] detected compromised devices launching DDoS attacks by applying several basic machine learning schemes including naive Bayes and Bayesian network classifiers. Zolotukhin et al. [18] proposed a deep learning model using an auto-encoder to detect encrypted DDoS traffics. Neural networks are also popular for identification of flooding-based DDoS attacks in software-defined networks [19]. For detection of zero-day attacks, which is another type of DDoS besides flooding-based, deep learning techniques including recurrent neural networks (RNNs) [20], graph neural networks (GNNs) [21], and deep natural language processing (NLP) [22] are adopted to assess firmware security with high accuracy.

Malware injection attacks intend to install malware into computing systems. The implementation of traditional detection schemes requires much expert knowledge in malware injection, which encourages researchers to explore autonomous schemes. For instance, NLP is applied to identify structured query language (SQL) injection vulnerabilities in edge node program [23]. Multiple other machine learning techniques are also evaluated by Ross et al. [24] to compare the accuracy in detecting SQL injection attacks. A cross-site scripting (XSS) classifier [25] is proposed based on three features of URLs, webpage, and social networking services to detect XSS vulnerabilities, achieving over 97% accuracy. However, the existing schemes hardly fulfill the requirements of edge-assisted IoT for energy efficiency when they involve numerous features to generate detection models. In the following subsection, we propose a preliminary solution based on nonadditive measure theory to perform feature selection before applying security schemes, so that the tradeoff between security and energy efficiency can be addressed.

4.2 Preliminary solution to security-energy tradeoff

In this subsection, we present the details of our proposed feature selection solution. Firstly, we introduce a concept of interaction measure among edge node features and its properties. Then, we generate a multivariate regression model with the interaction measures as independent variables, so that the feature interactions can be quantified by solving the regression model. Finally, feature selection can be performed based on the obtained interaction measures to achieve energy efficiency for security schemes against DDoS and malware injection attacks.

4.2.1 *Interaction measure among features*

The security features for DDoS and malware injection attacks can be categorized into static type and dynamic type. Static features are extracted from application source code without executing it, including permissions, hardware components used and data flow usage. Dynamic features of edge nodes are extracted from application behaviors in a running environment, including network connections, system calls, and resource usage. Because machine learning schemes tend to involve numerous features that may be related to malicious detection to achieve high accuracy, efficiency cannot be guaranteed to afford the demands of real-time edge-assisted IoT applications. Pre-processing of feature selection is necessary to reduce the model dimensionality. However, like cross-layer design in communication system, different features may also have correlations among each other. The sensitivity of an individual feature towards the detection result can be affected by the values of other features, which complicates the situation of feature selection. To systematically quantify the significance of all edge node features towards the detection results, we introduce the concept of interaction measure to evaluate significance based on feature sets instead of individual ones.

In our proposed solution, interaction measure is formulated by Choquet integral [26] to obtain nonadditivity. As an expansion of Lebesgue integral, which treats the impact of each feature individually with associated weight, Choquet integral computes the combined impacts among different sets of features by applying interaction measure μ:

$$z = (C) \int_A f \, \mathrm{d}\mu, \tag{1}$$

where $A = \{a_1, a_2, \ldots, a_N\}$ is the set of features, f is the tuple of observed values on subset of A determined by all of the features, and μ is the interaction measure to be estimated based on subset of A. The larger value of $|\mu|$ implies higher significance of the corresponding feature subset. The quantification of μ is non-additive, which means for two feature subsets A and B, $A \subseteq B \subseteq X$ does not necessarily imply $|\mu(A)| < |\mu(B)|$. This property is to reflect the cases that sometimes adding more features into a certain subset may decrease its significance, because features may have opposite impacts towards objective value. As illustrated in Figure 4, the full set of three features $\{x_1, x_2, x_3\}$ is not necessarily the most significant set for detection model. If we only consider x_1 and x_2, the subset has the largest contribution towards detection results, which indicates that x_3 is less important and can be discarded in feature selection to reduce dimensionality of detection model.

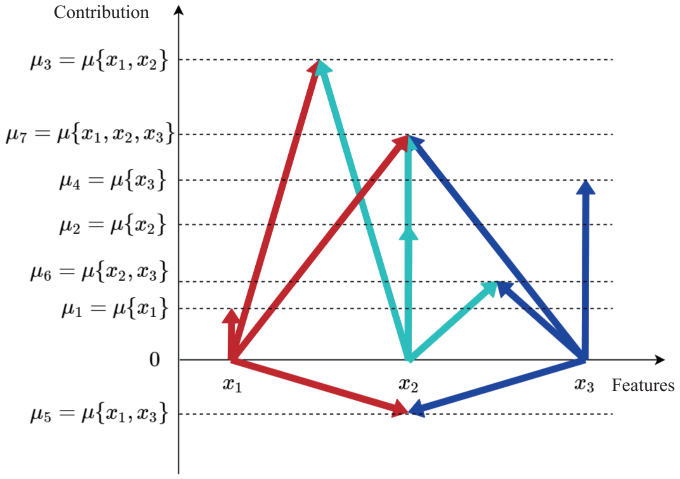
Contribution
$\mu_3 = \mu\{x_1, x_2\}$
$\mu_7 = \mu\{x_1, x_2, x_3\}$
$\mu_4 = \mu\{x_3\}$
$\mu_2 = \mu\{x_2\}$
$\mu_6 = \mu\{x_2, x_3\}$
$\mu_1 = \mu\{x_1\}$
0
x_1
x_2
x_3
Features
$\mu_5 = \mu\{x_1, x_3\}$

Figure 4 (Color online) Illustration of interaction measure nonadditivity among three features.

4.2.2 *Model formulation*

A multivariate regression model can be generated for quantifying interaction measures based on the observations of edge node features and IoT application labels. The preliminary solution provides interaction measures as the outcomes, which indicate the joint impacts of various sets of edge node features towards the detection result. Suppose the training data contains N features and Q samples, it can be organized into an observation matrix F, where f_{ij} denotes the observed value of observation i and feature j, and y_i denotes objective value of observation i. Min-max normalization is firstly applied to the dataset to eliminate the influence of various feature scales on the quantification of feature interactions. Given the formulation of Choquet integral as shown in (1) and normalization of observation data, the interaction measures of all feature subsets can be obtained with a multivariate regression model:

$$y = e + \int_{(C)} g \; \mathrm{d}\mu + N(0, \delta^2), \tag{2}$$

where y is the probability of edge node activity being malicious; μ are the independent variables of the model representing the significance of feature subsets; e is the regression constant that represents the bias of model when no features is used for classification; $\int_{(C)} g \; \mathrm{d}\mu$ is Choquet integral computed as the weighted sum of interaction measures μ and normalized observed values g calculated according to corresponding subset of features; $N(0, \delta^2)$ is the normally distributed random perturbation of regression model; and the variance δ^2 is the measure of regression residue error. To obtain Choquet integral, the tuple g needs to be computed for all feature subsets by extending the matrix of observation data from $Q - by - N$ matrix F into $Q - by - 2^N$ augmented matrix Z. For each row,

$$z_{q0} = 1, \quad z_{qk} = \max\{\min_{n \in \{n|k_n=1\}}\{g_{qn}\} - \max_{n \in \{n|k_n=0\}}\{g_{qn}\}, 0\}, \tag{3}$$

where $q = 1, 2, \ldots, Q$ and $k = 1, 2, \ldots, 2^N - 1$, k_n is the n-th lowest bit of binary representation of k. For example, when $k = 3$, its binary representation is $00 \cdots 011$, so that $k_1 = k_2 = 1$ and $k_3 = \cdots = k_N = 0$. In this case, the values of k can be used to represent different subsets of features by converting k to binary form. The corresponding interaction measure for each subset can also be expressed as $\{e, \mu_1, \mu_2, \ldots, \mu_{2^N-1}\}$, so that all possible feature combinations are considered, and the measure for null set μ_0 can be replaced by regression constant e in regression model, which indicates the bias when no features are applied. With (3), the augmented matrix Z can be generated based on observations matrix F, which extends its column number to the number of feature subsets. The values of g in (2) can be obtained from the entries of matrix Z as the polynomial coefficients, which correspond to the interaction measures of feature subsets, respectively. The regression model of (2) can be expressed in matrix form as

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_Q \end{bmatrix} = \begin{bmatrix} 1 & z_{11} & \cdots & z_{1(2^N-1)} \\ 1 & z_{21} & \cdots & z_{2(2^N-1)} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & z_{Q1} & \cdots & z_{Q(2^N-1)} \end{bmatrix} \times \begin{bmatrix} e \\ \mu_1 \\ \mu_2 \\ \cdots \\ \mu_{2^N-1} \end{bmatrix} + N(0, \delta^2). \tag{4}$$

4.2.3 *Feature selection*

The regression residue σ^2 in (8) can be minimized with least square scheme to determine the independent variables μ:

$$\mu = (Z^{\mathrm{T}} Z)^{-1} Z^{\mathrm{T}} y. \tag{5}$$

By minimizing the regression residue σ^2, we can obtain the solution of $\{e, \mu_1, \mu_2, \ldots, \mu_{2^N-1}\}$, so that the interactions of all combinations of features from null set to full set can be quantified, which reflects the significance of features or their combinations in affecting detection results for DDoS and malware injection attacks. Finally, feature selection can be performed in pre-processing by choosing the feature set with large interaction measure and small number of elements, so that the dimensionality of detection model can be substantially reduced. As the selected features carry the majority of impacts towards the detection results, the machine learning based schemes fulfill the security demands for edge-assisted IoT applications while achieving energy efficiency. The tradeoff between security and energy consumption can be adjusted by the number of selected features.

4.3 Simulation results

In this subsection, we conduct extensive simulations to illustrate how to adjust the complexity of security schemes by using the proposed solution. Two existing security schemes against DDoS attack [18] and malware injection attack [24] are applied in this case study. We present the performance of the two schemes before and after feature selection to validate the effectiveness of our proposed solution to address security-energy tradeoff. Figure 5 presents the results of interaction measure for all possible sets of edge node features, respectively. The values on x-axis represent feature subsets, in which the binary is denoted as 1 if the corresponding feature is selected, otherwise the binary is 0. The values on y-axis represents the quantification of interaction measures. The larger absolute value of interaction measure indicates larger impact of the feature set towards detection result. Either positive or negative, the feature set with the most significant interaction measure can be selected in pre-processing, such that the model dimensionality for security scheme is reduced. We may also select smaller feature sets with less interaction, so that the detection accuracy is partially traded for computational efficiency. Through the feature selection performed by the proposed solution, different tradeoffs between security and energy can be reached depending on the demands of IoT applications.

After feature selection, the state-of-art security schemes [18, 24] are applied against DDoS attack and malware injection attack, respectively. Figure 6 presents performances of two schemes with different numbers of selected features, where the balanced accuracy reflects the security level and computational time reflects the energy consumption. We can observe that for both schemes, the detection accuracy remains favorable as long as more than half of the features are selected, while the computational time decreases exponentially as fewer features are selected. The results verify that the proposed solution is capable of adjusting the accuracy and computational complexity of security schemes to achieve different security-energy tradeoffs.

Finally, the proposed solution is compared with two benchmark feature selection methods, which are principle component analysis (PCA) and factor analysis (FA). Feature selection is performed as the preprocessing step with the proposed solution, PCA, and FA, respectively. Then, the security schemes against DDoS attack and malware injection attack are applied. The balanced accuracy of each case is

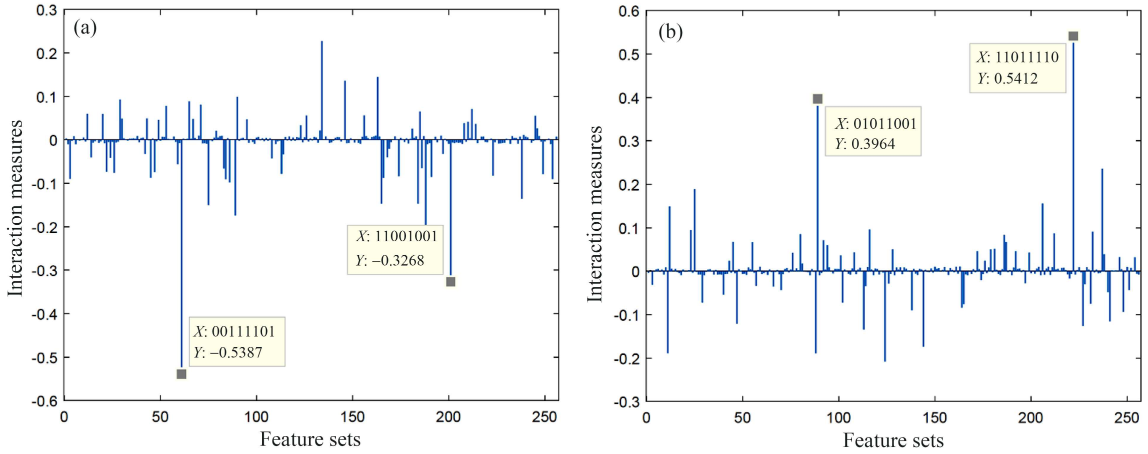

Figure 5 (Color online) Quantification values of interaction measures for security features towards the detection of (a) DDoS attack; (b) malware injection attack.

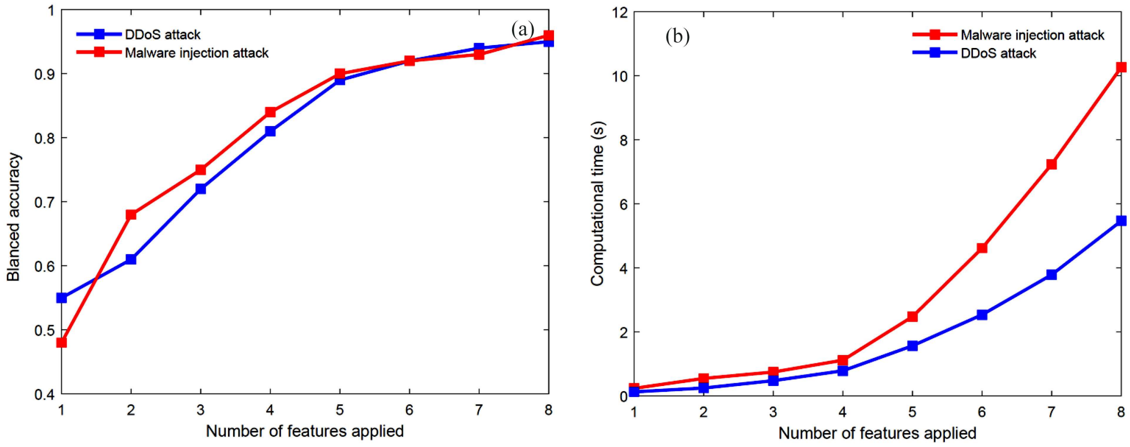

Figure 6 (Color online) Performance evaluation using different numbers of features applied in detection schemes. (a) Balanced accuracy; (b) computational time.

presented in Figure 7. We can observe from the comparison results that the proposed solution outperforms PCA and FA with various numbers of features and is able to help both security schemes achieve the highest accuracy. With fewer selected features, the complexity of security schemes can be substantially reduced to save more energy. The proposed solution is capable of achieving energy efficiency while guaranteeing a high degree of security in this case study.

5 Future research directions

Although some security solutions are proposed for IoT and edge computing based on machine learning techniques, there are still several promising open research directions including but not limited to the followings. Firstly, to achieve energy efficiency in security schemes, the correlation between energy consumption and algorithm complexity needs to be investigated in a quantitative way. This requires more researches on both hardware and software of IoT devices. If the energy consumption can be accurately estimated based on the execution of security schemes, the challenge of security-energy tradeoff in edge-assisted IoT can be formulated and solved as an optimization problem.

Secondly, a more flexible and self-adaptive security scheme is expected in edge-assisted IoT. Because not all user data are equally sensitive, the demands for security protection vary from task to task. How to automatically identify the sensitivity of user data and flexibly alter the security intensity is worthy of

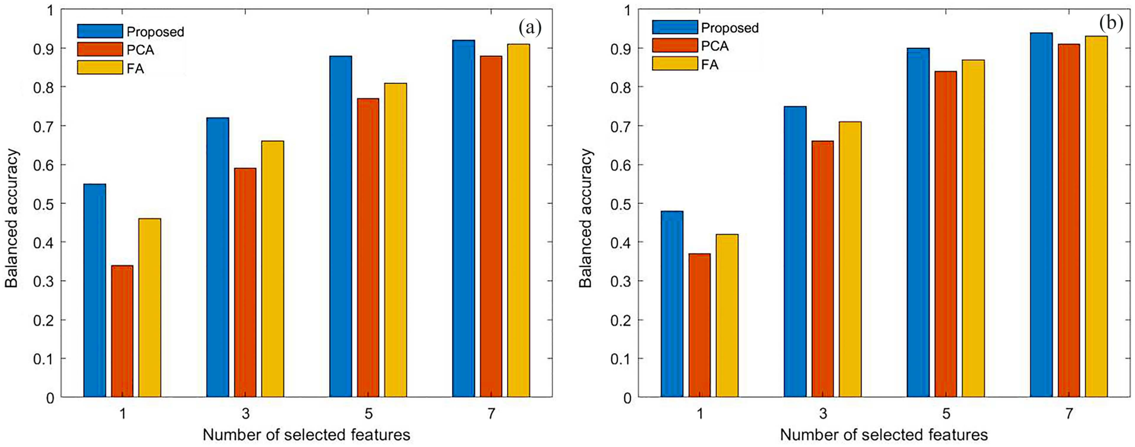

Figure 7 (Color online) Balanced accuracy after feature selection using different numbers of features. (a) DDoS attack; (b) malware injection attack.

exploration for achieving both data confidentiality and efficiency in edge-assisted IoT applications. Owing to limited capacities and resources of edge nodes, a lightweight security scheme without too many complex operations is in urgent demand, especially for the real-time applications with low-latency requirements.

Thirdly, except for being constraints, the energy consumption of IoT devices and edge nodes can also be treated as status indicators to facilitate security schemes against cyber attacks. When an edge node is compromised and used to launch an attack or perform malware injection, the edge node is expected to execute addition operations compared to legitimate behaviors so that abnormal energy consumption may be detected. By tracking the energy status of IoT devices and edge nodes, machine learning techniques can be applied to extract energy features from hardware components, recognize different patterns of energy consumption and infer devices' kernel conditions. Therefore, the development of security schemes integrated with energy monitoring is worth exploring to enhance the capability of malicious detection in edge-assisted IoT from a different angle.

Last but not least, the security schemes in traditional communication networks have been broadly explored and verified to be reliable in countering various attacks. Nevertheless, the simple migration of those traditional security schemes to edge-assisted IoT encounters inevitable obstacles, such as limited computation power, diverse operating systems and software, and different network topologies. The heterogeneity of IoT devices, services, collected data and network protocols usually hinders the migration of security schemes to another application or scenario. The security of edge-assisted IoT still lacks a universal solution, as most schemes focus on addressing one or few particular types of attacks but fail to adapt to the majority of other attacks. Therefore, it is challenging to incorporate various security schemes to protect the whole IoT architecture in an unified way.

6 Conclusion

In this paper, we have investigated the security issues in edge-assisted IoT and discussed the tradeoff between security and energy efficiency in security schemes. Specifically, we introduced the edge-assisted IoT architecture, promising applications and key advantages over traditional IoT. Then, we have presented several major security threats in edge-assisted IoT, and discussed the challenges brought by the conflicting demands of security and energy efficiency. In addition, we have proposed a preliminary solution to address the security-energy tradeoff, which is illustrated in a case study of DDoS and malware injection attack. Simulation results have validated the effectiveness of the proposed solution. Finally, several promising open issues are discussed to shed light on the further research on security and energy efficiency in edge-assisted IoT.

References

1 Ejaz W, Anpalagan A, Imran M A, et al. Internet of Things (IoT) in 5G wireless communications. IEEE Access, 2016, 4: 10310–10314

2 Zhang K, Ni J, Yang K, et al. Security and privacy in smart city applications: challenges and solutions. IEEE Commun Mag, 2017, 55: 122–129

3 Ni J, Zhang K, Lin X, et al. Securing fog computing for internet of things applications: challenges and solutions. IEEE Commun Surv Tut, 2018, 20: 601–628

4 Xiao Y, Jia Y, Liu C, et al. Edge computing security: state of the art and challenges. Proc IEEE, 2019, 107: 1608–1631

5 Shirazi S N, Gouglidis A, Farshad A, et al. The extended cloud: review and analysis of mobile edge computing and fog from a security and resilience perspective. IEEE J Sel Areas Commun, 2017, 35: 2586–2595

6 Antonakakis M, April T, Bailey M, et al. Understanding the Mirai botnet. In: Proceedings of the 26th USENIX Security Symposium (USENIX Security 17), 2017. 1093–1110

7 Roman R, Lopez J, Mambo M. Mobile edge computing, fog et al.: a survey and analysis of security threats and challenges. Future Gener Comput Syst, 2018, 78: 680–698

8 Ni J, Lin X, Shen X S. Toward edge-assisted Internet of Things: from security and efficiency perspectives. IEEE Netw, 2019. 33: 50–57

9 Liu D, Yan Z, Ding W, et al. A survey on secure data analytics in edge computing. IEEE Internet Things J, 2019, 6: 4946–4967

10 Liu Z, Yin X, Hu Y. CPSS LR-DDoS detection and defense in edge computing utilizing DCNN Q-learning. IEEE Access, 2020, 8: 42120–42130

11 Xiao L, Xie C, Chen T, et al. A mobile offloading game against smart attacks. IEEE Access, 2016, 4: 2281–2291

12 Shi C, Liu J, Liu H, et al. Smart user authentication through actuation of daily activities leveraging wifi-enabled iot. In: Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2017. 1–10

13 Hlavacs H, Treutner T, Gelas J P, et al. Energy consumption side-channel attack at virtual machines in a cloud. In: Proceedings of 2011 IEEE 9th International Conference on Dependable, Autonomic and Secure Computing, 2011. 605–612

14 Zhang K, Yang K, Liang X, et al. Security and privacy for mobile healthcare networks: from a quality of protection perspective. IEEE Wirel Commun, 2015, 22: 104–112

15 Kolias C, Kambourakis G, Stavrou A, et al. DDoS in the IoT: Mirai and other botnets. Computer, 2017, 50: 80–84

16 Martin M C, Lam M S. Automatic generation of XSS and SQL injection attacks with goal-directed model checking. In: Proceedings of USENIX Security Symposium, 2008. 31–44

17 Livadas C, Walsh R, Lapsley D, et al. Usilng machine learning technliques to identify botnet traffic. In: Proceedings of the 31st IEEE Conference on Local Computer Networks, 2006. 967–974

18 Zolotukhin M, Hämäläinen T, Kokkonen T, et al. Increasing web service availability by detecting application-layer DDoS attacks in encrypted traffic. In: Proceedings of the 23rd International Conference on Telecommunications (ICT), 2016. 1–6

19 Niyaz Q, Sun W, Javaid A Y. A deep learning based DDoS detection system in software-defined networking (sdn). 2016. ArXiv: 161107400

20 Chua Z L, Shen S, Saxena P, et al. Neural nets can learn function type signatures from binaries. In: Proceedings of the 26th USENIX Security Symposium (USENIX Security 17), 2017. 99–116

21 Song W, Yin H, Liu C, et al. Deepmem: learning graph neural network models for fast and robust memory forensic analysis. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2018. 606–618

22 Zuo F, Li X, Young P, et al. Neural machine translation inspired binary code similarity comparison beyond function pairs. 2018. ArXiv: 180804706

23 Jackson K A, Bennett B T. Locating SQL injection vulnerabilities in java byte code using natural language techniques. In: Proceedings of SoutheastCon 2018, 2018. 1–5

24 Ross K, Moh M, Moh T S, et al. Multi-source data analysis and evaluation of machine learning techniques for SQL injection detection. In: Proceedings of the ACMSE 2018 Conference, 2018. 1–8

25 Rathore S, Sharma P K, Park J H. XSSClassifier: an efficient XSS attack detection approach based on machine learning classifier on SNSs. J Inf Process Syst, 2017, 13: 1014–1028

26 Murofushi T, Sugeno M. An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure. Fuzzy Sets Syst, 1989, 29: 201–227