Information Fusion to Defend Intentional Attack in Internet of Things

Pin-Yu Chen, *Student Member, IEEE*, Shin-Ming Cheng, *Member, IEEE*, and Kwang-Cheng Chen, *Fellow, IEEE*

***Abstract*—Robust network design against attacks is one of the most fundamental issues in Internet of Things (IoT) architecture as IoT operations highly rely on the support of the underlaying communication infrastructures. In this paper, the vulnerability of IoT infrastructure under intentional attacks is investigated by relating the network resilience to the percolation-based connectivity. Intentional attacks impose severe threats on the network operations as it can effectively disrupt a network by paralyzing a small fraction of nodes, and therefore deteriorating IoT operations. A fusion-based defense mechanism is proposed to mitigate the damage caused by such attacks, where each node feedbacks minimum (one-bit) local decision to the fusion center for attack inference. By formulating the attack and defense strategy as a zero-sum game, the outcome of the game equilibrium is used to evaluate the effectiveness of the proposed mechanism. The robustness of the Internet-oriented and the cyber-physical system (CPS)-oriented networks are specifically analyzed to illustrate the foundation of future IoT infrastructure. Both analytical and empirical results show that the proposed mechanism greatly enhances the robustness of IoT, even in the weak local detection capability and fragile network structure regime.**

***Index Terms*—Attack and defense, connectivity, cyber-physical system (CPS), machine-to-machine (M2M) communications, network vulnerability, zero-sum game.**

NOMENCLATURE

N	Number of nodes in a network.
d_i	Degree of node i, $d_1 \geq d_2 \geq \ldots \geq d_N$.
$\mathbf{D}$	Degree of a randomly selected node.
$\mathbf{D_0}$	Degree of a randomly selected node in the original network.
$P(d)$	Degree distribution.
$P_0(d_0)$	Degree distribution of the original network.
$\widetilde{d}_{\max}$	Highest degree of the remaining network.
q	Fraction of removed nodes.
q_c	Critical value of removed fraction for network disruption.
$\widetilde{q}$	Probability of a randomly selected link leading to a deleted node.
T	Set of attack strategy, $T \in \{1, 2, \ldots, N\}$.
S	Set of defense strategy, $S \in \{1, 2, \ldots, N\}$.
$\underline{t}$	Probability distribution of the attack strategy.
$\underline{s}$	Probability distribution of the defense strategy.
C_q	Cost at the network level.
H_i	Hypothesis of node level defense made by node i.
H_C	Hypothesis of network level defense made by the fusion center.
P_D	Detection probability at the node level.
P_F	False alarm probability at the node level.
$\lfloor \eta' \rfloor$	Threshold of the fusion-based defense.
P_{TS}	Payoff matrix of the zero-sum game.
α	Parameter of the Internet-oriented network.
β	Parameter of the CPS-oriented network.

Manuscript received October 30, 2013; revised May 18, 2014; accepted June 29, 2014. Date of publication July 08, 2014; date of current version August 01, 2014. This work was supported by the Ministry of Science and Technology, Taiwan, under Contract MOST 103-2221-E-011-008-MY3 and Contract NSC 102-2221-E-011-046-MY2.

P.-Y. Chen is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: pinyu@umich.edu).

S.-M. Cheng is with the Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei 10607, Taiwan (e-mail: smcheng@mail.ntust.edu.tw).

K.-C. Chen is with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan (e-mail: chenkc@cc.ee.ntu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JIOT.2014.2337018

I. INTRODUCTION

WITH THE advance of communication technology and the penetration of various networking applications and services in our daily lives, employing an ubiquitous Internet of Things (IoT) system [1], [2] has been proposed to empower a full-mechanical automation [such as the machine-to-machine (M2M) communication and the smart grid], which embraces autonomous operations and distributed computations in large-scale networks (i.e., the network diameter increases with the order of the network size) [3]–[8]. Generally speaking, in addition to physical infrastructures, an IoT infrastructure involves a communication network to collect and exchange useful information to fully facilitate the advantages of IoT. Despite a wide range of IoT applications and deployments, the robustness of IoT (namely, cyber-physical security [9]) is far from realized owing to the absence of theoretical characterizations. As the network functionality and robustness are closely related to the network structure, the disruption or disfunction of some devices in an IoT infrastructure may incur disastrous threats to the operation and reliability, which is an ever-increasing concern for the deployments of IoT technology [10]–[13]. For instance, the U.S. Department of Energy (DOE) has identified attack resistance to be one of the seven major properties required for the operation of smart grid [14].

As IoT systems are expected to operate in an autonomous fashion, while being capable of supporting communications

among relatively large number of machines compared with current communication systems, the collective features of a large-scale network of our interest can be characterized with the aid of complex network theory [15]–[21], which provides analytically tractable tools to investigate a network consisting of tremendous number of nodes and intricate interconnections, either in nature or engineered systems. The Internet and the cyber-physical system (CPS) are considered as the typical examples of IoT infrastructures.

From the viewpoint of a complex network, an intentional attack is known to be quite effective in disrupting a network by paralyzing some fraction of nodes with the highest degree [22], [23], which is equivalent to node removals on the corresponding network graph [24]. Similarly, an intentional attack incurs fatal threats to IoT by intruding the nodes (e.g., launching denial of service attacks on the devices) with the most physical connections. However, owing to the network resilience of modern communication networks [19], [22], [25]–[29], a network is expected to recover from temporal malfunctioning as long as most of the nodes are still connected, which coincides with the *percolation* phenomenon in statistic physics [30], [31].

More importantly, IoT technology indicates the possibility of sparing tremendous investments in installing nodal defense modules on each device provided that a robust defense mechanism can be implemented at the network level for attack inference and defense reactions in order to mitigate the damage caused by network disruption, which is particularly essential in networks with enormous network size and stringent energy budget. As the adversary and the defender tend to maximize their own profits by attacking/defending a subset of nodes in the network, and their payoffs are coupled with the resulting network robustness, a zero-sum game [32] is naturally formed between these two parties. The attack/defense strategies at the game equilibrium are in general highly nontrivial, yet they play an essential role in network robustness and defense capability. To the best of our knowledge, this is the first series of efforts [12], [33] that utilize the outcome of the game equilibrium to analyze the interactions between the adversary and the defender, where the game payoff is coupled with the corresponding network robustness.

Considering the increasing computation capability of an adversary, an intentional attack is more difficult to be detected if the adversary is aware of the network topology and intelligently sabotage some central nodes in the network. To tackle intentional attacks in an IoT infrastructure, a fusion-based defense mechanism is proposed to enhance the network robustness. As illustrated in Fig. 1, since installing individual defense modules on each machine may incur excessive implementation costs for IoT deployment owing to its enormous number of devices, each node simply performs local detection via intrusion detection or anomaly detection on suspicious activities [34]–[38] and then feedbacks minimum (one-bit) decision to the fusion center for attack inference and defense reactions at the network level. The fusion center launches immediate defense reactions if an attack at the network level is detected, otherwise it keeps surveillance on the network to mitigate the potential damage caused by the false alarms. Note that distinct from the traditional distributed detection scheme

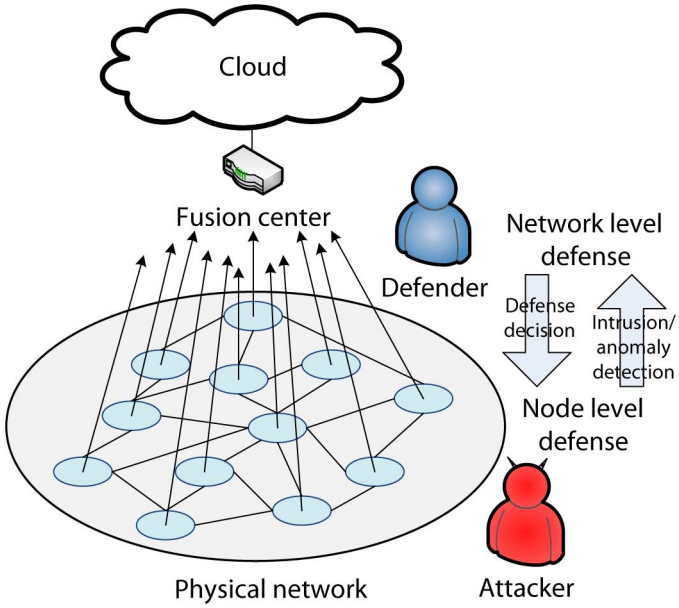

Fig. 1. System model of an IoT infrastructure and the proposed fusion-based defense mechanism. Each node feedbacks minimum (one-bit) decision to the fusion center for attack inference and further defense reactions at the network level. Based on the collected information, the fusion center launches defense reactions such as nodal quarantine or reset upon the presence of the attack, or it keeps surveillance on the network. In reality, the fusion center can be a gateway, a data aggregator, or an intelligent machine. The adversary's attack strategy is to attack a subset of nodes to disrupt the network while reducing the risks of being detected. On the other hand, the defender's strategy is to infer potential attacks from the feedback information. Note that this system model can also represent a hierarchical multilevel network defense mechanism where each node is a subfusion center in a subsystem.

where each node feedbacks a local decision for inferring of a common event [39], the adversary can intelligently sabotage some but not all the nodes to disrupt the network such that the attack is not a common event to all nodes, which hinders the precision of attack inference and thereby results in severe threats on the network robustness. It is worth noting that this system model can also represent a hierarchical multilevel network defense mechanism where each node is a subfusion center in a subsystem.

This paper specifically analyzes the network robustness of the Internet-oriented network and the CPS-oriented network (smart grid as the representative case [40]), as these two networks possess distinct topological features, and they are the foundation of future IoT infrastructures owing to their maturity and well-developed communication protocols. It is worth mentioning that the proposed framework can be used to evaluate the network robustness of any large-scale IoT infrastructure when the corresponding network parameters are specified. In addition, both analytical results and empirical data extracted from the real-world large-scale networks support that the fusion-based defense mechanism is quite efficacious against intentional attacks, even with weak local detection capability and inherently fragile network structure. This paper, therefore, offers novel avenues to the theoretical analysis and network robustness enhancement for IoT.

This paper is organized as follows. Section II elucidates the related works. The percolation-based connectivity and the

impacts of node removal on an arbitrary network are introduced in Section III. The attack strategy, defense mechanism, and network resilience are specified in Section IV. The vulnerabilities of the Internet-oriented network and the CPS-oriented network are discussed in Section V. In Section VI, we formulate the fusion-based defense mechanism via optimal data fusion approaches and solve the detection threshold for attack inference at the network level. Due to the competing nature between the adversary and the defender, the interactions among the two parties are formulated as a zero-sum game to evaluate the effectiveness of the proposed fusion-based defense mechanism in Section VII. In Section VIII, the performance of the proposed mechanism is presented by deploying the mechanism on synthetic network models and empirical network data. Finally, Section IX concludes this paper.

II. Related Works

Network vulnerability to attack is a fundamental security issue in IoT [10], [41]. An attack on IoT infrastructure can be categorized into two types according to its purpose: 1) *manipulation attack* and 2) *disruption attack*. Manipulation attacks take advantage of the communication vulnerabilities to manipulate the measurements or decisions in IoT, such as the emulation attacks in cognitive radio networks [42], Byzantine attacks in spectrum sensing [43], sybil attacks in open-access distributed systems [44], and data injection attacks in smart grid [45]. Kosut *et al.* investigated the impacts of the number of meters under manipulation on the attack observability and specified the smallest set of meters sufficient for the adversary to control the smart grid [46]. Kim and Poor [47] demonstrated that it is possible to defend against malicious data injection if a small subset of measurements can be made immune to the data injection attacks. Although manipulation attack may impose severe threats on IoT, the damage is often limited to revenue loss or performance degradation.

In contrast to fraudulently manipulating the system state estimators without deteriorating the system's functionality, disruption attacks aim to paralyze IoT operations by launching denial of service attacks [48] to jam the entire system [49]. An intentional attack takes a more progressive approach to disrupt a network by paralyzing a set of nodes with the highest degree and incurring the disintegration of the entire system. Albert *et al.* [22] first showed that a power-law distributed network is quite tolerant to random node failures (removals), while it is very fragile to selective removals. In [25], Cohen *et al.* studied the resilience of the Internet to random breakdowns, and they proposed an analytical model to evaluate the critical value for the network breakdown under an intentional attack in [50]. Solé *et al.* [51] analyzed the topology and robustness of the European power grids under an intentional attack. Moreover, Xiao *et al.* [23] verified that intentional attack is the most fatal attack to disrupt the network when the network topology is known to the adversary. Our previous results focus on network resilience in scale-free complex networks [33].

It is worth mentioning that the aforementioned research on intentional attacks mainly focuses on the inherent network resilience to an intentional attack and no further defense mechanisms are considered in the literature. However, from an engineering point of view, with the support of IoT and the advance of the anomaly/intrusion detection techniques, the damage caused by an intentional attack can be greatly mitigated via our proposed fusion-based defense mechanism with minimum (one-bit) feedback from each node. Moreover, by formulating the interactions between the adversary and the defender as a two-player zero-sum game with its game payoff coupled by the network resilience, the game equilibrium and the resulting optimal attack/defense strategies are far from being realized.

The main contributions of this paper are summarized as follows.

1) By relating the network resilience to the percolation-based connectivity [30], the critical value of an arbitrary network to remain connection in percolation sense after removing the targeted nodes is introduced. This critical value is associated with the fragility of IoT infrastructure under an intentional attack.
2) A fusion-based defense mechanism is proposed to mitigate the damage caused by an intentional attack. The proposed defense mechanism requires only minimum (one-bit) local decision from each node to reduce the additional communication overheads and computation complexity.
3) Given the percolation-based connectivity of an IoT infrastructure, the competing nature of the adversary and the defender is formulated as a zero-sum game, where the game payoff is coupled with the corresponding network robustness, and the outcome of the game equilibrium is used to evaluate the performance of the proposed defense mechanism, which quantizes the network robustness in a precisely defined manner.
4) Considering the Internet-oriented network and the CPS-oriented network, we analyze the critical value for the percolation-based connectivity under intentional attacks as these networks possess distinct topological features and they are the foundation of future IoT infrastructure.
5) By implementing the fusion-based defense mechanism on the Internet router-level topology and the European power grid, both analytical results and empirical network data show that the proposed mechanism greatly enhances the network robustness to prevent IoT infrastructure from disruption.

III. Percolation-Based Connectivity and Node Removal

A. Percolation-Based Connectivity

In large-scale networks with enormous number of nodes, considering full connectivity to collect information from all nodes and edges may not be practically feasible. For example, we could exploit the second smallest eigenvalue of the *Laplacian matrix* [52] of a network to determine the existence of full connectivity at the cost of generating tremendous computation overheads. Therefore, in large-scale networks, we would like to only consider if the largest component in the network is able to maintain the main operation of the entire

system. From this perspective, the characterization of network connectivity corresponds to the investigation for the properties of the largest connected component (giant component) [17], [18], [24]. It is suggested that percolation theory be useful for the analysis of connectivity in large-scale networks owing to its tractability and engineering interpretations. For instance, percolation theory is used to study the sensor coverage problem in IoT [53] and interference analysis in wireless networks [54]. Nonetheless, in case of small-scale networks, the critical value to sustain percolation-based connectivity can be obtained by exhaustively running experimental trials, and thereby, it can be applied to the fusion-based defense mechanism.

According to the seminal work in [55], given a degree distribution $P(d)$ of an arbitrary network, a giant component containing the majority of the nodes exists if the degree distribution satisfies the condition $\sum_d d(d-2)P(d) > 0$, which is equivalent to

$$\frac{\mathbb{E}[\mathbf{D}^2]}{\mathbb{E}[\mathbf{D}]} = \frac{\mathbf{var}(\mathbf{D}) + (\mathbb{E}[\mathbf{D}])^2}{\mathbb{E}[\mathbf{D}]} > 2 \tag{1}$$

where $\mathbf{D} \in [d_{\min}, d_{\max}]$ is the random variable representing the degree of a randomly selected node with probability distribution function $P(d)$, and $d_{\min}$ ($d_{\max}$) is the smallest (largest) degree of the network. The first observation for (1) is that if we fix the average number of links (degrees), a network with larger degree variance is prone to possess a giant connected component since a node with extremely high degree is more likely to occur in such a network. The second observation for (1) is that in order to guarantee the existence of a giant component, we are interested in the case that the highest degree increases with the order of the network size and it is sufficiently large such that the degree variance diverges eventually, which we refer as the large-scale network limit.

B. Random Node Removal

Following [25], [50], given the original degree distribution $P_0(d_0)$, the new degree distribution of the network after randomly removing q fraction of nodes (the links emanating from the nodes are removed as well) is

$$P(d) = \sum_{d_0=d}^{d_{\max}} P_0(d_0) \binom{d_0}{d} (1-q)^d q^{d_0-d}. \tag{2}$$

Applying (2) to (1), the criterion for the percolation-based connectivity after random node removal becomes

$$\theta \triangleq \frac{\mathbb{E}[\mathbf{D}^2]}{\mathbb{E}[\mathbf{D}]} = \frac{(1-q)^2\mathbb{E}[\mathbf{D_0}^2] + q(1-q)\mathbb{E}[\mathbf{D_0}]}{(1-q)\mathbb{E}[\mathbf{D_0}]} = 2. \tag{3}$$

Reorganizing (3), we obtain the critical threshold q_c for percolation as

$$q_c = 1 - \frac{1}{\theta_0 - 1} \tag{4}$$

where $\theta_0 \triangleq \frac{\mathbb{E}[\mathbf{D_0}^2]}{\mathbb{E}[\mathbf{D_0}]}$ is calculated from the original degree distribution. The critical value q_c is an important indicator of the network robustness, since it means that a network with original degree distribution $P_0(d_0)$ will be disintegrated into many small components once we randomly remove more than q_c fraction of nodes from the network.

C. Targeted Node Removal

In addition to random node removal, if the q fraction of nodes with the highest degree is removed from a network consisting of N nodes, the cutoff degree (the highest degree in the remaining network) is reduced to some value $\widetilde{d}_{\max} < d_{\max}$, and it can be evaluated from

$$\sum_{d=\widetilde{d}_{\max}+1}^{d_{\max}} P(d) = \sum_{d=\widetilde{d}_{\max}+1}^{\infty} P(d) - \frac{1}{N} = q \tag{5}$$

with the relation $\sum_{d_{\max}}^{\infty} P(d) = \frac{1}{N}$, and the fact that the q fraction of nodes with the highest degree are removed from the network such that $\sum_{d=\widetilde{d}_{\max}+1}^{d_{\max}} P(d) = q$.

Since targeted node removal disconnects the links emanating from the removed nodes to the remaining nodes, removing the q fraction of nodes with the highest degree results in the change of degree distribution. For an arbitrary node, the deletion probability $\widetilde{q}$ of a randomly selected link leading to a deleted node equals the ratio of the number of links belonging to the deleted nodes to the number of links, i.e., $\widetilde{q} = \sum_{d=\widetilde{d}_{\max}+1}^{d_{\max}} \frac{dP(d)}{\mathbb{E}[\mathbf{D_0}]}$ [50], [56] due to the fact that a node with higher degree has more chance to be deleted. Consequently, the critical value for percolation-based connectivity under targeted node removal can be obtained from (4) by replacing q_c and $d_{\max}$ with $\widetilde{q}_c$ and $\widetilde{d}_{\max}$ as removing q fraction of nodes with the highest degree is equivalent to removing $\widetilde{q}$ fraction of nodes randomly.

IV. System Model

Suppose there are N nodes in the network sorted in descending degree order, i.e., $d_1 \geq d_2 \geq \ldots \geq d_N$. Based on the anomaly/intrusion detection techniques of each node, every node feedbacks a local decision to the fusion center for attack inference and further defense reactions. Upon validation of the attack at the network level, the fusion center takes immediate reaction (e.g., nodal quarantine) on the suspicious nodes reporting the presence of attack. Without loss of generality, we assume that there is only one fusion center in the network. Nonetheless, this work can be extended to a multistage hierarchical network structure composed of several autonomous fusion centers when distributed computation mechanisms are involved, which will be considered as our future works. In IoT, the fusion center can be a gateway, a data aggregator or an intelligent machine. For clear reading, the notations throughout this paper are summarized in the Nomenclature.

A. Intentional Attack

We consider the worst-case scenario that the adversary knows the complete topological information (degree of every node) of the network and it is capable of sabotaging all the nodes simultaneously. The attack strategy of an adversary is to sabotage $T \in \{1, 2, \ldots, N\}$ nodes in descending degree order, where $T = N$ refers to an undifferentiated attack or an uniform attack, while $T < N$ contributes to an intentional attack on T nodes with the highest degree. The attack on the ith node is in vain if the fusion center detects the presence of the attack and takes immediate reaction such as nodal

quarantine or reset on the ith node. Intuitively, a uniform attack is an inadequate strategy for the adversary at the risk of exposed activity, whereas an intentional attack is more effective since it makes the most use of the network topology to achieve its aim, and in the meanwhile, it is more difficult to be detected at the network level by paralyzing a small fraction of nodes with the highest degree.

B. Node Level Defense: Local Detection

Since the *a priori* probability of launching an attack and the impacts of nodal defense (e.g., a node turns off its operations once an attack is detected on its side) on the overall network robustness are unknown at the node level, each node employs Neyman–Pearson criterion for hypothesis testing with detection probability P_{D_i} and false alarm probability P_{F_i}. $H_i = 1$ denotes the hypothesis that ith node is under attack, otherwise $H_i = 0$. For simplicity, we assume every node possesses identical anomaly/intrusion detection capability, i.e., $P_{D_i} = P_D$ and $P_{F_i} = P_F$. Based on the local detection, every node feedbacks one-bit information u_i to the fusion center for attack inference and defense reactions, where $u_i = 1$ if the ith node declares that it is under attack, otherwise $u_i = -1$.

C. Network Level Defense: Surveillance and Quarantine

Regarding the defense strategy at the network level, the fusion center infers the presence of an attack by keeping surveillance on the local decisions from $S \in \{1, 2, \ldots, N\}$ nodes in descending degree order. Since the impacts of nodal quarantines on the network robustness are known at the network level while the *a priori* probability of attack is still unknown, a binary hypothesis testing based on minimax criterion is employed at the fusion center to minimize the potential cost, where $H_C = 1$ if attack occurs in the network, otherwise $H_C = 0$. The fusion center quarantines the nodes which feedback the local decision $u_i = 1$ when $H_C = 1$, or it declares a null attack and keeps surveillance on the network. Note that if the defense at the network level is not considered, the optimal strategy for the adversary is to launch an undifferentiated attack to sabotage as many nodes as possible to disrupt the network.

D. Network Resilience

By relating the network resilience to the percolation-based connectivity, one is able to characterize the network robustness of an arbitrary network for quantitative analysis. Considering the resilience of an IoT infrastructure and the critical value to sustain percolation-based connectivity under intentional attacks, we define the cost at the network level as

$$C_q = \begin{cases} 1, & \text{if } q > q_c \\ -1, & \text{if } q \le q_c. \end{cases} \tag{6}$$

The resilience of IoT infrastructure is determined by analyzing the critical value q_c to sustain percolation-based connectivity. The network transitions from the connected phase to the disconnected phase if more than q_c fraction of nodes are paralyzed. Note that the cost of erroneous node quarantine due to false attack alarm is identical to the role of node removal since immediate reaction is taken by the fusion center once the attack is detected at the network level.

V. Intentional Attack on Canonical Networks

Due to the facts that the Internet and the CPS are probably the largest man-made engineering networks in the world, in this section, we specifically investigate the damage caused by intentional attacks on the synthetic network models of the Internet router-level topology and the CPS. These two types of networks are known to possess quite distinct topological features, and they are very likely to be the foundation of future IoT infrastructure owing to their mature deployments and well-developed communication protocols.

A. Internet-Oriented Network

For an Internet-oriented network, the degree distribution follows a power-law distribution [57]

$$P(d) = c_1 \cdot d^{-\alpha}, \qquad d = d_N, d_{N-1}, \ldots, d_1 \tag{7}$$

with exponent α and normalization coefficient c_1. A power-law distributed network is also renowned as a *scale-free* network [17], [18] when $2 \le \alpha \le 3$ since its second and higher-order moments of the degree distribution are usually divergent.

By relaxing the degree d to be real-valued, we have the continuous approximation in the large-scale network limit $(d_1 \to \infty)$ as $c_1 = \frac{1-\alpha}{d_1^{1-\alpha} - d_N^{1-\alpha}} \overset{d_1 \to \infty}{=} (\alpha - 1) d_N^{\alpha - 1}$ for $\alpha > 1$ because of the power-law distribution. By relaxing the degree to be real-valued, substituting (7) into (5) and integrating (5), we obtain the relation between the cutoff degree after intentional attacks and the fraction of removed nodes as

$$\widetilde{d}_{\max} = d_N \left(q + \frac{1}{N} \right)^{\frac{1}{1-\alpha}}. \tag{8}$$

When the size of the network is huge, the critical value for percolation-based connectivity can thus be evaluated as $q_c \overset{N \to \infty}{=} \left(\frac{\widetilde{d}_{\max}}{d_N} \right)^{1-\alpha}$. Moreover, removing q fraction of nodes with the highest degree is equivalent to the probability $\widetilde{q}$ of reaching a removed node by following a randomly selected link as discussed in Section III-C, where

$$\widetilde{q} = \sum_{d=\widetilde{d}_{\max}}^{d_1} \frac{dP(d)}{\mathbb{E}[\mathbf{D_0}]} \overset{d_1 \to \infty}{=} \left(\frac{\widetilde{d}_{\max}}{d_N} \right)^{2-\alpha} = q^{\frac{2-\alpha}{1-\alpha}} \tag{9}$$

for $\alpha > 2$. Note that $\widetilde{q} \to 1$ as $\alpha \to 2$ suggests that a power-law-distributed network with exponent $\alpha < 2$ is very vulnerable to intentional attacks since removing only a few nodes is able to disrupt the entire network. Applying the result to (3) with $\theta_0 = \left(\frac{2-\alpha}{3-\alpha} \right) \frac{d_1^{3-\alpha} - d_N^{3-\alpha}}{d_1^{2-\alpha} - d_N^{2-\alpha}}$, the cutoff degree after intentional attacks is obtained by solving the following equation

$$\left(\frac{\widetilde{d}_{\max}}{d_N} \right)^{2-\alpha} - d_N \left(\frac{2-\alpha}{3-\alpha} \right) \left[\left(\frac{\widetilde{d}_{\max}}{d_N} \right)^{3-\alpha} - 1 \right] - 2 = 0. \tag{10}$$

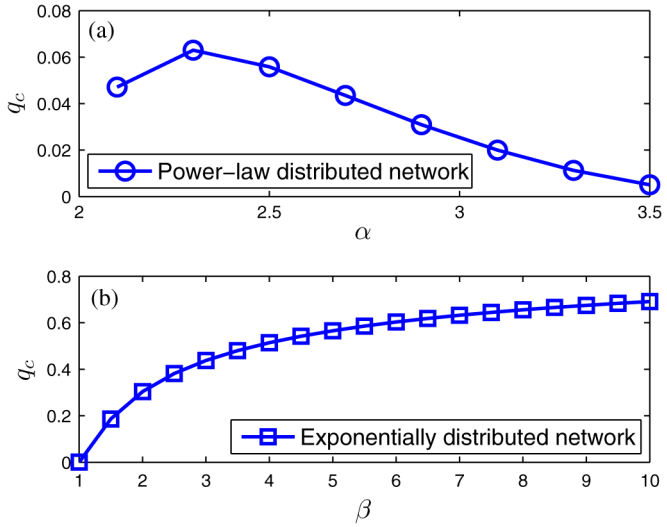

Fig. 2. Robustness of the two network prototypes under intentional attacks. (a) Internet-oriented network. The degree distribution follows a power-law distribution with the power-law exponent α, i.e., $P(d) \sim d^{-\alpha}$. (b) CPS-oriented network. The degree distribution follows an exponential distribution with the exponential exponent β, i.e., $P(d) \sim \frac{1}{\beta} e^{-\frac{d}{\beta}}$.

The critical value for percolation-based connectivity under intentional attacks is thus $q_c = \left(\frac{\widetilde{d}_{\max}}{d_N}\right)^{1-\alpha}$ for $\alpha > 2$.

As demonstrated in Fig. 2 (a), a power-law-distributed network is very vulnerable to intentional attacks such that paralyzing a small fraction of nodes with the highest degree contributes to network disruption. As α increases, the size of the giant component becomes smaller due to weak connectivity [22], [50], and it leads to the decline in q_c. The decline in q_c as $\alpha \to 2$ is explained by the critically high degree of just a few nodes. Note that it has been validated in [50] that the large-scale network limit assumption has negligible impacts on the accuracy of the critical value provided that the network size is large enough. Roughly speaking, the rate of convergence depends on the underlaying network topology and most of the datasets have extremely large number of nodes such that running experimental trials on these datasets to obtain the exact critical values are highly infeasible and intractable. In addition, in case of time-varying network topology, we can update the degree distribution of a network and recalculate q_c.

B. CPS-Oriented Network

For CPS, particularly the smart grid, the degree distribution follows an exponential distribution [51]

$$P(d) = c_2 \cdot e^{-\frac{d}{\beta}}, \qquad d = d_N, d_{N-1}, \ldots, d_1 \tag{11}$$

with an exponent β and normalization coefficient c_2. Similarly, by continuous approximation at the large-scale network limit, we have $c_2 \overset{d_1 \to \infty}{=} \frac{1}{\beta} e^{\frac{d_N}{\beta}}$, and from (5), the relation between the cutoff degree after intentional attacks and the fraction of removed nodes is

$$\widetilde{d}_{\max} = -\beta \ln\left(q + \frac{1}{N}\right) + d_N. \tag{12}$$

The probability of reaching a removed node following a randomly selected link is

$$\widetilde{q} = \frac{\exp\left(\frac{d_N}{\beta}\right)}{d_N + \beta} \cdot \left(\widetilde{d}_{\max} + \beta\right) \exp\left(-\frac{\widetilde{d}_{\max}}{\beta}\right). \tag{13}$$

If d_N is negligible, (13) can be simplified as $\widetilde{q} \overset{d_N \to 0}{=} \left[1 - \ln\left(q + \frac{1}{N}\right)\right]\left(q + \frac{1}{N}\right)$. Applying the result to (3) with $\theta_0 = 2\beta$, the critical value for percolation-based connectivity can be obtained by solving the equation

$$\left[1 - \ln\left(q_c + \frac{1}{N}\right)\right]\left(q_c + \frac{1}{N}\right) = 1 - \frac{1}{2\beta - 1}. \tag{14}$$

As demonstrated in Fig. 2(b), the critical value q_c increases with the exponent β, since larger β indicates that the network is more strongly connected, and the adversary has to sabotage more nodes to successfully disintegrate the network.

VI. Fusion-Based Defense Analysis

In addition to the inherent vulnerability of an IoT infrastructure, the attack inference at the fusion center plays an essential role in enhancing the network robustness and mitigating the damage caused by intentional attacks. In the proposed defense mechanism, the fusion center infers the presence of the attack based on the collected one-bit feedback from each node. Leveraging optimal fusion rule [58], [59], let $\mathbf{u} = \sum_{i=1}^{S} a_i u_i$ denote the observation obtained at the fusion center when keeping S nodes with the highest degree under surveillance, where

$$a_i = \begin{cases} \log\left(\dfrac{P_D}{P_F}\right), & \text{if } u_i = 1 \\ \log\left(\dfrac{1 - P_F}{1 - P_D}\right), & \text{if } u_i = -1 \end{cases} \tag{15}$$

is the optimal weighted coefficient for data fusion. The likelihood ratio test (LRT) at the fusion center is

$$\mathbf{u} \underset{H_C=0}{\overset{H_C=1}{\gtrless}} \eta \tag{16}$$

for some threshold η. Since the fusion center infers the presence of the attack based on the local decisions of S nodes with the highest degree, adopting the *k-out-of-n* decision rule as consistent with [58], [59], (16) can be written as

$$\mathbf{k}\left(\log\left[\frac{P_D(1-P_F)}{P_F(1-P_D)}\right]\right) \underset{H_C=0}{\overset{H_C=1}{\gtrless}} \eta + S\log\left(\frac{1-P_F}{1-P_D}\right) \tag{17}$$

where $\mathbf{k}$ out of S nodes report attacks. Without loss of generality, we assume $P_D > P_F$ so that (17) becomes

$$\mathbf{k} \underset{H_C=0}{\overset{H_C=1}{\gtrless}} \left(\log\left[\frac{P_D(1-P_F)}{P_F(1-P_D)}\right]\right)^{-1} \left\{\eta + S\log\left(\frac{1-P_F}{1-P_D}\right)\right\} \triangleq \eta' \tag{18}$$

and $\mathbf{k}$, therefore, has a binomial distribution $\mathbf{BIN}(S, P_D)$ when $H_C = 1$ and $\mathbf{BIN}(S, P_F)$ when $H_C = 0$.

Let C_{xy} denotes the cost when the event is $H_C = x$ while the decision is $H_C = y$ at the fusion center. Regarding the defense cost at the network level, we set $C_{00} = C_{11} = 0$ and $C_{01} = C_F = C_{10} = C_M = C_q$. The cost of false alarm C_F is identical to the cost of miss detection C_M, since erroneous node quarantine may also lead to network disruption. Adopting minimax criterion and randomized decision rule with probability ϵ at the network level, η' can be solved by setting $P_M^C = P_F^C$ [60], where P_M^C (P_F^C) is the probability of miss detection (false alarm) at the fusion center. We have

$$\begin{aligned}&\sum_{k=0}^{\lfloor\eta'\rfloor-1} P(\mathbf{k}=k \mid H_C=1) + (1-\epsilon)P(\mathbf{k}=\lfloor\eta'\rfloor \mid H_C=1)\\ &= \epsilon P(\mathbf{k}=\lfloor\eta'\rfloor \mid H_C=0) + \sum_{k=\lfloor\eta'\rfloor+1}^{S} P(\mathbf{k}=k \mid H_C=0)\end{aligned} \tag{19}$$

where $\lfloor\eta'\rfloor$ is the greatest integer that is smaller or equal to η', since $\mathbf{k}$ is a discrete random variable.

Let $F(k;n,p)$ denote the cumulative distribution function (CDF) of $\mathbf{k} \sim \mathbf{BIN}(n,p)$, we have

$$\begin{aligned}F(k;n,p) &= P(\mathbf{k}\le k)\\ &= I_{1-p}(n-k,k+1)\\ &= (n-k)\binom{n}{k}\int_0^{1-p} t^{n-k-1}(1-t)^k dt\end{aligned} \tag{20}$$

where $I_z(a,b) = \frac{B(z;a,b)}{B(a,b)}$ is the regularized incomplete beta function, $B(z;a,b) = \int_0^z t^{a-1}(1-t)^{b-1}dt$ is the incomplete beta function, and $B(a,b) = B(1,a,b) = \int_0^z t^{a-1}(1-t)^{b-1}dt$ is the complete beta function. With (20), (19) can be rewritten as

$$\begin{aligned}&F(\lfloor\eta'\rfloor-1;S,P_D) + (1-\epsilon)P(\mathbf{k}=\lfloor\eta'\rfloor \mid H_C=1)\\ &= \epsilon P(\mathbf{k}=\lfloor\eta'\rfloor \mid H_C=0) + 1 - F(\lfloor\eta'\rfloor;S,P_F).\end{aligned} \tag{21}$$

Consequently, given S and (20), the threshold $\lfloor\eta'\rfloor$ can be obtained by solving the following equation

$$\begin{aligned}&(S-\lfloor\eta'\rfloor+1)\binom{S}{\lfloor\eta'\rfloor-1}\int_0^{1-P_D} t^{S-\lfloor\eta'\rfloor}(1-t)^{\lfloor\eta'\rfloor-1}dt\\ &+(1-\epsilon)\binom{S}{\lfloor\eta'\rfloor}P_D^{\lfloor\eta'\rfloor}(1-P_D)^{S-\lfloor\eta'\rfloor}\\ &= \epsilon\binom{S}{\lfloor\eta'\rfloor}P_F^{\lfloor\eta'\rfloor}(1-P_F)^{S-\lfloor\eta'\rfloor} + 1 - (S-\lfloor\eta'\rfloor)\binom{S}{\lfloor\eta'\rfloor}\\ &\times\int_0^{1-P_F} t^{S-\lfloor\eta'\rfloor-1}(1-t)^{\lfloor\eta'\rfloor}dt.\end{aligned} \tag{22}$$

The relation between the threshold $\lfloor\eta'\rfloor$ and the number of nodes with the highest degree under surveillance (S) is shown in Fig. 3. The threshold has a linear scalability with respect to the number of nodes with the highest degree under surveillance, and higher false alarm probability contributes to larger $\lfloor\eta'\rfloor$ in order to minimize the potential cost introduced by erroneous node quarantine. A direct observation from Fig. 3 is that higher false alarm probability tends to benefit the adversary since the adversary is prone to disrupt the network without being detected at the fusion center if the threshold is too high. It is, therefore, of crucial importance for the defender to determine the optimal value of S against intentional attacks so that the damage of network disruption can be minimized.

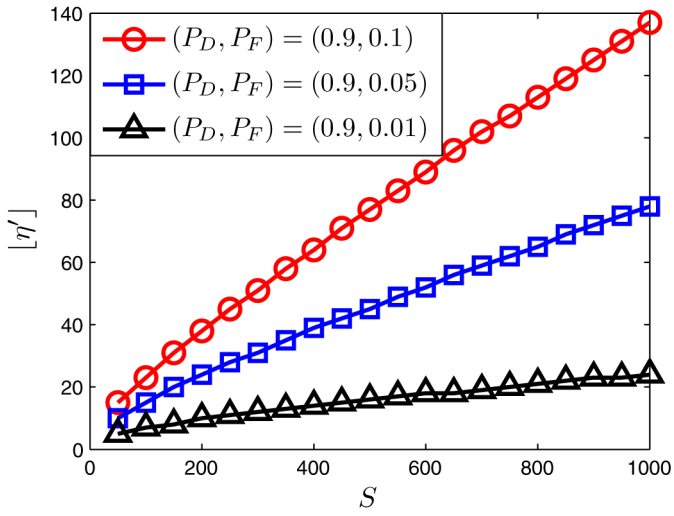

Fig. 3. Threshold $\lfloor\eta'\rfloor$ with respect to S for network level defense with different (P_D, P_F) configurations. $N = 1000$ and $\epsilon = 0.5$. The threshold has a linear growth with the number of nodes with the highest degree under surveillance. The threshold increases with the false alarm probability for attack inference at the network level in order to minimize the overall cost.

VII. Game-Theoretic Analysis

With the network resilience described in Section IV-D, an intentional attack on an IoT infrastructure is regarded as effective if the adversary sabotages at least $\lfloor Nq_c \rfloor$ nodes with the highest degree to disrupt the network without being detected by the fusion center (i.e., $H_C = 0$). As derived in Section VI, the fusion center declares a null attack ($H_C = 0$) if less than $\lfloor\eta'\rfloor$ nodes with the highest degree report that they are under attack simultaneously. More interestingly, since the threshold $\lfloor\eta'\rfloor$ increases linearly with the increase in S as shown in Fig. 3, if the fusion center knows the adversary's strategy T (the number of nodes with the highest degree under attack), the fusion center manages to adjust its defense strategy (the number of nodes with the highest degree under surveillance) to smaller S in order to detect the attack with high precision. On the other hand, if the adversary knows the defender's strategy S and the detection capability P_D at the node level, the adversary tends to sabotage as many nodes as possible provided that the fusion center regards the abnormal feedbacks as a null attack and takes no reaction on the suspicious nodes. There is clearly a tradeoff between the attack strategy and the defense strategy among these two parties, and the effectiveness of the defense mechanism can be evaluated by analyzing the network robustness at the stable state (game equilibrium), where both the adversary and the defender simultaneously choose its optimal strategy against each other to maximize their own payoffs, and the payoff of each player cannot be improved via unilateral change in its own strategy at the game equilibrium.

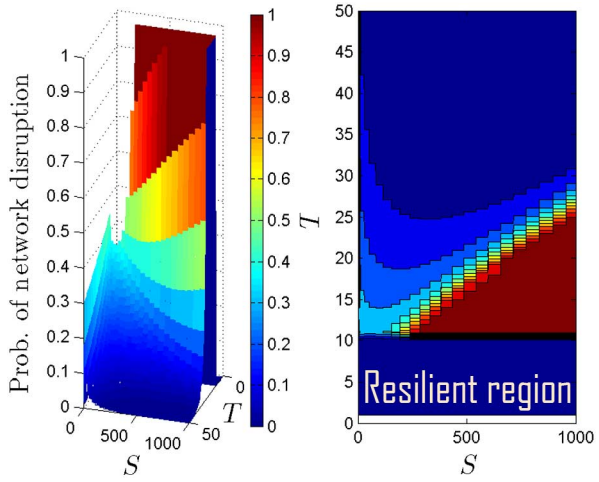

Fig. 4. Probability of an adversary to disrupt IoT infrastructure with respect to the attack strategy T and the defense strategy S. $N = 1000$, $q_c = 0.01$, $P_D = 0.9$, $P_F = 0.01$, and $\epsilon = 0.5$. An intentional attack in the resilient region turns out to be in vain, since it fails to sabotage enough nodes to paralyze the entire network owing to the network resilience. In this case, the adversary has to attack at least $\lfloor N \cdot q_c \rfloor = 10$ nodes to disrupt the network with nonzero probability.

For the purpose of demonstration, Fig. 4 displays the probability for an adversary to successfully disrupt an IoT infrastructure given the attack strategy T and the defense strategy S. An intentional attack in the resilient region (i.e., $T \leq \lfloor Nq_c \rfloor$) turns out to be in vain since it fails to sabotage enough nodes to paralyze the entire network owing to the network resilience. In other words, an adversary has to sabotage at least $\lfloor Nq_c \rfloor$ nodes to disrupt an IoT infrastructure with nonzero probability. In case that $\lfloor Nq_c \rfloor \leq T \leq \lfloor \eta' \rfloor - 1$, the adversary is able to disrupt the entire network without being detected by the fusion center. It is observed that the increase in T leads to the decrease in the disruption probability since the attack can be detected with higher probability as T increases. On the other hand, decreasing S may enhance the precision of detecting intentional attacks. However, as the attack and defense strategies are not known by its opponent, the outcome of the game equilibrium therefore offers insightful analysis on the robustness of IoT.

Using game theory, the interactions between the adversary and the defender can be formulated as a two-player, zero-sum, matrix game, where the adversary's strategy is to attack T nodes with the highest degree and the defender's strategy is to keep S nodes with the highest degree under surveillance. The payoff of one player is set to be the opponent's cost as defined in (6). The payoff of the matrix game is specified by a payoff matrix $\mathcal{P} \in \mathbb{R}^{N \times N}$, where the entry $\mathcal{P}_{TS}$ denotes the payoff of the defender when the adversary's strategy is T and the defender's strategy is S. With the network cost defined in (6) and the detection capability of every node (P_D), we denote $\mathbf{T_D} \sim \mathbf{BIN}(T, P_D)$ as a binomial random variable representing the number of nodes detecting the attack when T nodes are under attack. An attack is in vain if the adversary sabotages less than $\lfloor Nq_c \rfloor$ nodes because the network is still connected in percolation sense, while the attack is effective if $\lfloor Nq_c \rfloor \leq T \leq \lfloor \eta' \rfloor - 1$, where the fusion center fails to detect the attack prior to the network disruption. Regarding the randomized decision rule, when $T \geq \lfloor Nq_c \rfloor$ and $\mathbf{T_D} = \lfloor \eta' \rfloor$, with (6), the payoff is

$$\begin{aligned}\widetilde{C} &= 1 \cdot \epsilon P(\mathbf{T_D} = \lfloor \eta' \rfloor) - 1 \cdot (1-\epsilon) P(\mathbf{T_D} = \lfloor \eta' \rfloor) \\ &= (2\epsilon - 1) \binom{T}{\lfloor \eta' \rfloor} P_D^{\lfloor \eta' \rfloor} (1 - P_D)^{T - \lfloor \eta' \rfloor}. \end{aligned} \tag{23}$$

Moreover, given $T \geq \max\{\lfloor Nq_c \rfloor, \lfloor \eta' \rfloor\}$, the payoff becomes

$$\begin{aligned}\widehat{C} &= 1 \cdot P(\mathbf{T_D} \geq \lfloor \eta' \rfloor + 1) - 1 \cdot P(\mathbf{T_D} \leq \lfloor \eta' \rfloor - 1) + \widetilde{C} \\ &= 1 - F(\lfloor \eta' \rfloor; T, P_D) - F(\lfloor \eta' \rfloor - 1; T, P_D) + \widetilde{C} \\ &= 1 - I_{1-p}(T - \lfloor \eta' \rfloor, \lfloor \eta' \rfloor + 1) \\ &\quad - I_{1-p}(T - \lfloor \eta' \rfloor + 1, \lfloor \eta' \rfloor) + \widetilde{C}. \end{aligned} \tag{24}$$

From (22), the detection threshold η' is a function of the number of nodes under surveillance (S), i.e., $\eta'(S)$. Consequently, we write the payoff matrix $\mathcal{P}$ as

$$\mathcal{P}_{TS} = \begin{cases} 1, & \text{if } T < \lfloor Nq_c \rfloor \\ -1, & \text{if } \lfloor Nq_c \rfloor \leq T \leq \lfloor \eta'(S) \rfloor - 1 \\ \widehat{C}, & \text{if } T \geq \max\{\lfloor Nq_c \rfloor, \lfloor \eta'(S) \rfloor\}. \end{cases} \tag{25}$$

Due to the fact that there exists at least one (mixed strategy) Nash equilibrium in a finite matrix game such that no players can be better off by a unilateral change in their strategies, and the Nash equilibria are equivalent in the sense that the payoffs are identical [32], denoting $\underline{t} = (t_1, \ldots, t_N)$ and $\underline{s} = (s_1, \ldots, s_N)$ as the probability distribution on the strategy of the adversary and the defender, respectively, the adversary manages to choose an optimal $\underline{t}$ to minimize the defender's payoff, which is the solution of the optimization problem in nonnegative orthant

$$\begin{aligned} &\text{minimize} \quad \max_{s=1,\ldots,N} (\mathcal{P}^{\mathrm{Tr}} \underline{t})_s \\ &\text{subject to} \quad \underline{t} \succeq 0, \quad \underline{1}^{\mathrm{Tr}} \underline{t} = 1 \end{aligned} \tag{26}$$

where $(\cdot)^{\mathrm{Tr}}$ denotes matrix transpose and $\succeq$ denotes componentwise inequality. As proved in [61], the optimization problem in (26) is equivalent to the linear programming problem

$$\begin{aligned} &\text{minimize} \quad v \\ &\text{subject to} \quad \underline{t} \succeq 0, \quad \underline{1}^{\mathrm{Tr}} \underline{t} = 1 \\ &\qquad\qquad \mathcal{P}^{\mathrm{Tr}} \underline{t} \preceq v \underline{1} \end{aligned} \tag{27}$$

which is particularly suitable in analyzing the network robustness of IoT having tremendous nodes owing to its computation efficiency. Therefore, the solution of (27) v^* is the optimal expected payoff of the defender for the attack and defense zero-sum game, and the optimal expected payoff of the adversary is $-v^*$. More importantly, v^* quantifies the capability of the proposed defense mechanism subject to intentional attacks, and therefore, it can be used as a performance benchmark for network robustness. As shown in Fig. 5, the optimal expected payoff v^* of the defender increases with the critical value q_c, which is quite reasonable since the adversary needs to pay more efforts to disrupt the network if IoT infrastructure holds stronger resilience. Moreover, the adversary benefits

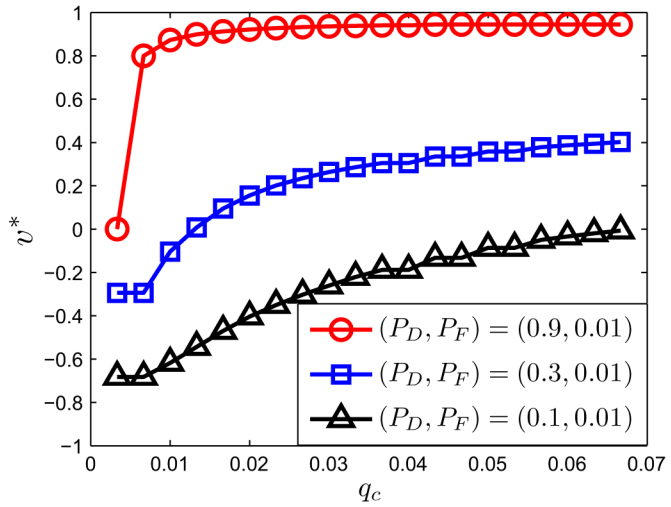

Fig. 5. Optimal expected payoff of the zero-sum game with respect to q_c with different (P_D, P_F) configurations. $N = 300$ and $\epsilon = 0.5$. Low critical value and weak detection capability indeed benefit the adversary in the sense that the adversary has more chance to win this game.

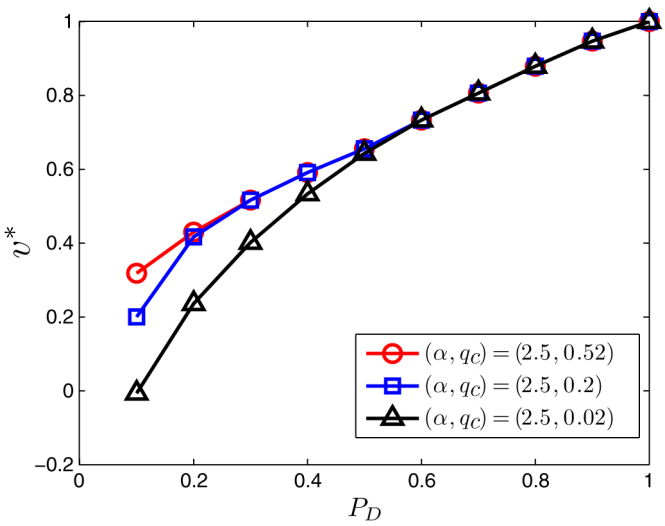

Fig. 6. Optimal expected payoff of the zero-sum game with respect to P_D under different q_c of the Internet-oriented network. $N = 400$, $P_F = 0.01$, $\epsilon = 0.5$, and $d_N = 5$, 2, and 1.

from weak detection capability (low P_D) if the local detection fails to distinguish the occurrence of attack.

In practice, the proposed fusion-based defense mechanism can be implemented in real-time via multiplicative weights update methods [62], [63] by updating the defense strategy according to the system loss of the previous stage. In other words, the Nash equilibrium in (26) can be achieved within ϵ precision of v^* in $O(\log N/\epsilon^2)$ stages, which is particularly preferable to large-scale IoT since the number of stages required to reach the game equilibrium within ϵ precision scales with $\log N$. Note that in addition to the percolation-based connectivity, other connectivity measures can also be used to define q_c for game-theoretic analysis. For instance, we can define q_c as the least fraction of node removals such that the remaining largest network size is no more than 10% of its original network size.

VIII. Performance Evaluation

For performance evaluation, the network robustness is quantitatively analyzed by solving the linear programming problem (27) with respect to the specified network parameters such as q_c, P_D, and P_F. The solution provides nontrivial optimal attack/defense strategies for the adversary and the defender, and the game equilibrium suggests that no player's payoff can be increased by unilaterally changing its own strategy, and therefore, the outcome of the game equilibrium turns out to be a stable network robustness measure, and it serves as a reliable benchmark for performance evaluation. Note that in reality, the critical value for network disruption can be either obtained by performing experimental trials on the collected network data or using complex network theory to identify the collective network features such as the skewness of the power-law degree distribution of the Internet-oriented network. In this section, both synthetic network models and parameters collected from real-world dataset will be investigated for performance comparisons.

Incorporating the concepts of network resilience in Section IV-D and the fusion-based defense in Section VI, in this section, we investigate the outcome of the attack and defense game at the game equilibrium as discussed in Section VII, where both the adversary and the defender aim to maximize their own payoffs. By formulating the intentional attack and fusion-based defense as a zero-sum game, the attack incurs severe threats on IoT if the optimal expected payoff $v^* < 0$, otherwise the defense is regarded as efficacious, since $v^* > 0$ implies that the defender has higher chance to win the game. The synthetic complex network models and the empirical network data extracted from Internet router-level map and European power grid are used to evaluate the robustness of IoT.

Regarding an Internet-oriented network, since most of the real-world power-law distributed networks have the degree exponent $2 \leq \alpha \leq 3$ [17], [18], in Fig. 6, we demonstrate the effectiveness of the fusion-based defense mechanism when $\alpha = 2.5$ with respect to the local detection capability, where the critical value q_c is obtained from (8) to (10) given the minimum degree d_N. The payoff of the defender increases with the local detection capability owing to better precision of the attack inference at the network level, and it asymptotically approaches to 1 when $P_D = 1$, suggesting that the attack is in vain when the detection probability of local node is high. In addition, lower d_N contributes to smaller q_c and v^*, since the network is more vulnerable to intentional attacks. Nonetheless, the fusion-based defense still takes advantage ($v^* > 0$) of such fragile network structure even with weak detection capability and small critical value to sustain the percolation-based connectivity.

Similar results can be found in a CPS-oriented network as shown in Fig. 7. The critical value q_c is obtained from (14) given the exponential exponent β. The payoff of the defender also increases with the local detection capability and $v^* \to 1$ as $P_D \to 1$. Larger exponential exponent β provides better defense performance since $\mathbb{E}[\mathbf{D_0}] = \beta$ suggests that a node has more links to other nodes and hence the robustness of

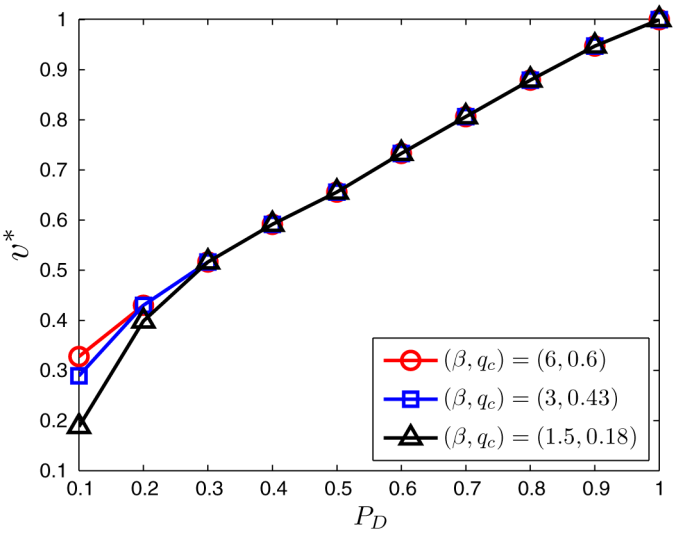

Fig. 7. Optimal expected payoff of the zero-sum game with respect to P_D under different q_c of the CPS-oriented network. $N = 400$, $P_F = 0.01$, and $\epsilon = 0.5$.

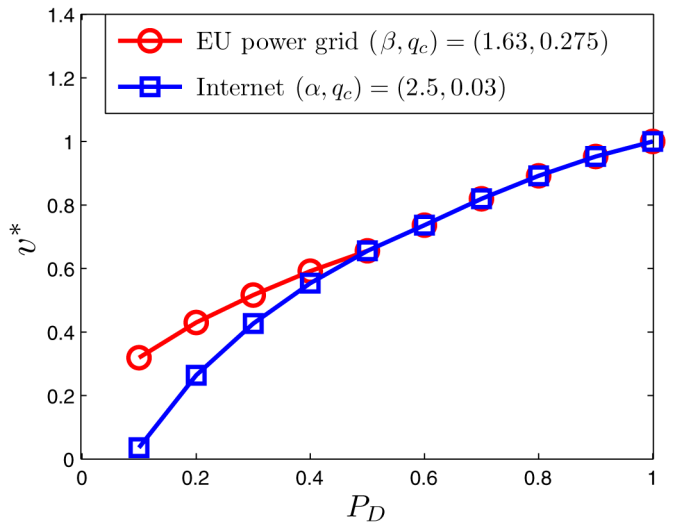

Fig. 8. Optimal expected payoff of the zero-sum game with respect to P_D and the empirical data collected in [22] and [51]. $P_F = 0.01$ and $\epsilon = 0.5$. The topological map of the Internet contains 6209 nodes and 12 200 links with $\mathbb{E}[\mathbf{D_0}] = 3.4$. The EU power grid contains 2783 nodes and 3762 links with $\mathbb{E}[\mathbf{D_0}] = 2.7$.

the network is enhanced. The results of these two networks also show that the proposed fusion-based defense mechanism indeed provides reliable and efficient protection against intentional attacks and is able to enhance the network robustness by acquiring minimum feedbacks from local nodes.

Following the empirical network data collected in [22] and [51], we analyze the performance of the fusion-based defense mechanism on the Internet router-level topology and the European power grid as the foundation of future IoT infrastructure. As shown in Fig. 8, the Internet is observed to be more vulnerable to intentional attacks due to the existence of hubs (nodes with much higher degree) [19] and relatively small critical value for percolation-based connectivity [22], while the defender is able to prevent the network from disruption

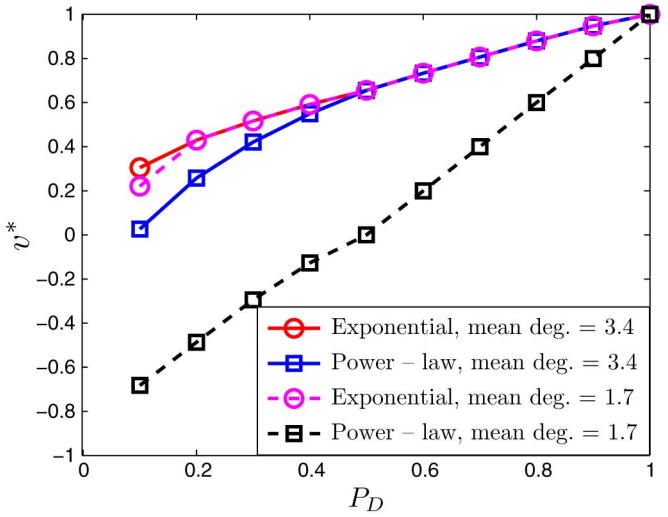

Fig. 9. Optimal expected payoff of the zero-sum game with respect to P_D and the same mean degree. $N = 400$, $P_F = 0.01$, and $\epsilon = 0.5$. The mean degree of the Internet-oriented network and the CPS-oriented network is set to be the same for fair comparison. The power-law distributed network is shown to be more vulnerable to intentional attacks, especially when the mean degree is small.

even under weak detection capability. When $P_D \to 0.1$, the fusion-based defense mechanism tends to lose its advantage for the Internet, suggesting that the adversary may disrupt the network if the local detection capability is inherently awkward.

To make a fair comparison between these two network configurations, we set the same mean degree with the assumption that $d_N = 1$ in the large-scale network limit, i.e., $\mathbb{E}[\mathbf{D_0}] = \frac{\alpha-1}{\alpha-2} = \beta$ from their degree distributions. The critical values for network disruption can be evaluated from (10) and (14), and therefore the corresponding game payoffs in (25) are specified. The mean degree of a network is closely related to the network construction cost since the degree of a node represents the number of the physical connections (e.g., power lines). As shown in Fig. 9, when $\mathbb{E}[\mathbf{D_0}] = 3.4$, both the Internet-oriented network and the CPS-oriented network have their payoff $v^* > 0$. When P_D is small, the Internet-oriented network is more vulnerable to intentional attacks than the CPS-oriented network owing to the existence of nodes with extremely high degree, which is consistent with the observations of (1). More interestingly, when the mean degree is halved, no significant impacts have been imposed on the CPS-oriented network, whereas the Internet-oriented network suffers severe performance degradation since the fragility of the network allows the adversary to sabotage less nodes with the highest degree to disintegrate the network without being detected. It is worth mentioning that the optimal expected payoff is not restrictive to any specific types of networks, and it only depends on the network structure (i.e., the critical value q_c) and the associated detection capability P_D and P_F, which provide useful guidance and ubiquitous performance benchmarks toward designing a robust network topology and defense mechanisms against fatal attacks.

IX. Conclusion

To tackle the damage caused by intentional attacks on an IoT infrastructure entrenched with complex network structure, a fusion-based defense mechanism is proposed for the attack inference at the network level by means of optimal data fusion approaches. The proposed mechanism requires only minimum (one-bit) local decision of each node to reduce the additional communication overheads and nodal defense module installments toward efficient defense framework in IoT operations. The vulnerability of IoT under intentional attacks is investigated by relating the network resilience to the percolation-based connectivity. A zero-sum game is introduced between the adversary and the defender, and the outcome of the game equilibrium is used to evaluate the network robustness of the proposed defense mechanism.

We specifically analyze the critical values of the Internet-oriented network and CPS-oriented network to sustain connectivity in percolation sense as these two networks possess distinct topological features and they are very likely to be the foundation of future IoT infrastructure. The Internet-oriented network is shown to be more vulnerable to intentional attacks than the CPS-oriented network owing to the existence of nodes with relatively high degree. The results on the synthetic network models and empirical data show that the proposed defense mechanism can effectively enhances the network robustness and counter the damage caused by intentional attacks, even with weak local detection capability and the inherently fragile nature of the network structure such as the power-law distributed networks. This paper, therefore, provides a general theoretic framework for network robustness analysis and enhancement in large-scale networks, in particular to IoT, CPS, and M2M communications.

References

[1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, 2010.

[2] M. Zorzi, A. Gluhak, S. Lange, and A. Bassi, "From today's INTRAnet of things to a future INTERnet of Things: A wireless- and mobility-related view," *IEEE Wireless Commun. Mag.*, vol. 17, no. 6, pp. 44–51, Dec. 2010.

[3] Z. Fadlullah *et al.*, "Toward intelligent machine-to-machine communications in smart grid," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 60–65, Apr. 2011.

[4] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. Johnson, "M2M: From mobile to embedded Internet," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 36–43, Apr. 2011.

[5] Y. Zhang *et al.*, "Home M2M networks: Architectures, standards, and QoS improvement," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 44–52, Apr. 2011.

[6] S.-Y. Lien, K.-C. Chen, and Y. Lin, "Toward ubiquitous massive accesses in 3GPP machine-to-machine communications," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 66–74, Apr. 2011.

[7] K.-C. Chen and S.-Y. Lien, "Machine-to-machine communications: Technologies and challenges," *Ad Hoc Netw.*, vol. 18, pp. 3–23, Jul. 2014.

[8] I. Stojmenovic, "Machine-to-machine communications with in-network data aggregation, processing, and actuation for large-scale cyber-physical systems," *IEEE Internet Things J.*, vol. 1, no. 2, pp. 122–128, Apr. 2014.

[9] S. K. Das, K. Kant, N. Zhang, *Handbook on Securing Cyber-Physical Critical Infrastructure*. Waltham, MA, USA: Morgan Kaufmann, 2012, pp. 1–817.

[10] R. H. Weber, "Internet of Things—New security and privacy challenges," *Comput. Law Security Rev.*, vol. 26, no. 1, pp. 23–30, 2010.

[11] O. Yagan, D. Qian, J. Zhang, and D. Cochran, "Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures, and robustness," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, pp. 1708–1720, 2012.

[12] P.-Y. Chen, S.-M. Cheng, and K.-C. Chen, "Smart attacks in smart grid communication networks," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 24–29, Aug. 2012.

[13] S. Keoh, S. Kumar, and H. Tschofenig, "Securing the Internet of Things: A standardization perspective," *IEEE Internet Things J.*, vol. 1, no. 3, pp. 265–275, 2014.

[14] U.S. Department of Energy (DOE), *A System View of the Modern Grid*. National Energy Technology Laboratory (NETL), Pittsburgh, PA, USA: U.S. DOE, 2007.

[15] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998.

[16] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.

[17] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Modern Phys.*, vol. 74, no. 1, pp. 47–97, Jan. 2002.

[18] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, Mar. 2003.

[19] D. Alderson, L. Li, W. Willinger, and J. Doyle, "Understanding Internet topology: Principles, models, and validation," *IEEE/ACM Trans. Netw.*, vol. 13, no. 6, pp. 1205–1218, Dec. 2005.

[20] A.-L. Barabási, "The architecture of complexity," *IEEE Control Syst. Mag.*, ivol. 27, no. 4, pp. 33–42, Aug. 2007.

[21] L. Cui, S. Kumara, and R. Albert, "Complex networks: An engineering view," *IEEE Circuits Syst. Mag.*, vol. 10, no. 3, pp. 10–25, 2010.

[22] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, Jul. 2000.

[23] S. Xiao, G. Xiao, and T. H. Cheng, "Tolerance of intentional attacks in complex communication networks," *IEEE Commun. Mag.*, vol. 45, no. 1, pp. 146–152, Feb. 2008.

[24] P.-Y. Chen and A. Hero, "Node removal vulnerability of the largest component of a network," in *Proc. IEEE GlobalSIP*, 2013.

[25] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Resilience of the Internet to random breakdowns," *Phys. Rev. Lett.*, vol. 85, no. 21, pp. 4626–4628, Nov. 2000.

[26] P. Cholda, A. Mykkeltveit, B. Helvik, O. Wittner, and A. Jajszczyk, "A survey of resilience differentiation frameworks in communication networks," *IEEE Commun. Surveys Tuts.*, vol. 9, no. 4, pp. 32–55, 2007.

[27] M. Menth, M. Duelli, R. Martin, and J. Milbrandt, "Resilience analysis of packet-switched communication networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 6, pp. 1950–1963, Dec. 2009.

[28] J. P. Sterbenz *et al.*, "Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines," *Comput. Netw.*, vol. 54, no. 8, pp. 1245–1265, 2010.

[29] P. Smith *et al.*, "Network resilience: A systematic approach," *IEEE Commun. Mag.*, vol. 49, no. 7, pp. 88–97, Jul. 2011.

[30] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Network robustness and fragility: Percolation on random graphs," *Phys. Rev. Lett.*, vol. 85, no. 25, pp. 5468–5471, Dec. 2000.

[31] M. Franceschetti and R. Meester, *Random Networks for Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[32] M. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA, USA: MIT Press, 1999.

[33] P.-Y. Chen and K.-C. Chen, "Intentional attack and fusion-based defense strategy in complex networks," in *Proc. IEEE Globecom*, Dec. 2011, pp. 1–5.

[34] B. Mukherjee, L. T. Heberlein, and K. N. Levitt, "Network intrusion detection," *IEEE Netw.*, vol. 8, no. 3, pp. 26–41, May–Jun. 1994.

[35] T. Bass, "Intrusion detection systems and multisensor data fusion," *Commun. ACM*, vol. 43, no. 4, pp. 99–105, Apr. 2000.

[36] A. Patcha and J.-M. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends," *Comput. Netw.*, vol. 51, no. 12, pp. 3448–3470, Aug. 2007.

[37] G. Androulidakis, V. Chatzigiannakis, and S. Papavassiliou, "Network anomaly detection and classification via opportunistic sampling," *IEEE Netw.*, vol. 23, no. 1, pp. 6–12, Jan.–Feb. 2009.

[38] G. Thatte, U. Mitra, and J. Heidemann, "Parametric methods for anomaly detection in aggregate traffic," *IEEE/ACM Trans. Netw.*, vol. 19, no. 2, pp. 512–525, Apr. 2011.

[39] P. K. Varshney, *Distributed Detection and Data Fusion*. Berlin, Germany: Springer-Verlag, 1996.

[40] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," *Proc. IEEE*, vol. 100, no. 1, pp. 210–224, Jan. 2012.

[41] H. Ning, H. Liu, and L. T. Yang, "Cyberentity security in the Internet of Things," *IEEE Comput.*, vol. 46, no. 4, pp. 46–53, 2013.
[42] R. Chen, J.-M. Park, and J. Reed, "Defense against primary user emulation attacks in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 25–37, Jan. 2008.
[43] A. Rawat, P. Anand, H. Chen, and P. Varshney, "Collaborative spectrum sensing in the presence of Byzantine attacks in cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 774–786, Feb. 2011.
[44] H. Yu, P. Gibbons, M. Kaminsky, and F. Xiao, "Sybillimit: A near-optimal social network defense against sybil attacks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 885–898, Jun. 2010.
[45] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. ACM Conf. Comput. Commun. Security*, Nov. 2009, pp. 21–32.
[46] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.
[47] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 326–333, Jun. 2011.
[48] K. Pelechrinis, M. Iliofotou, and S. Krishnamurthy, "Denial of service attacks in wireless networks: The case of jammers," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 2, pp. 245 –257, Quarter 2011.
[49] W. Xu, K. Ma, W. Trappe, and Y. Zhang, "Jamming sensor networks: Attack and defense strategies," *IEEE Netw.*, vol. 20, no. 3, pp. 41–47, May–Jun. 2006.
[50] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, "Breakdown of the Internet under intentional attack," *Phys. Rev. Lett.*, vol. 86, no. 16, pp. 3682–3685, Apr. 2001.
[51] R. V. Solé, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde, "Robustness of the European power grids under intentional attack," *Phys. Rev. E*, vol. 77, p. 026102, Feb. 2008.
[52] B. Mohar, "The Laplacian spectrum of graphs," in *Graph Theory, Combinatorics, and Applications*, Y. Alavi, G. Chartrand, O. Ollermann, and A. Schwenk, Eds. Hoboken, NJ, USA: Wiley, 1991, pp. 871–898.
[53] L. Liu, X. Zhang, and H. Ma, "Percolation theory-based exposure-path prevention for wireless sensor networks coverage in Internet of Things," *IEEE Sensors J.*, vol. 13, no. 10, pp. 3625–3636, 2013.
[54] W.-C. Ao, S.-M. Cheng, and K.-C. Chen, "Connectivity of multiple cooperative cognitive radio ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 263–270, 2012.
[55] M. Molloy and B. Reed, "A critical point for random graphs with a given degree sequence," *Random Struct. Algorithms*, vol. 6, pp. 161–179, Mar. 1995.
[56] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top*, vol. 64, p. 026118, Jul. 2001.
[57] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *ACM SIGCOMM*, Oct. 1999, pp. 251–262.
[58] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 1, pp. 98–101, Jan. 1986.
[59] S. C. A. Thomopoulos, R. Viswanathan, and D. C. Bougoulias, "Optimal decision fusion in multiple sensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, no. 5, pp. 644–653, Sep. 1987.
[60] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Berlin, Gremany: Springer-Verlag, 1994.
[61] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[62] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997.
[63] S. Arora, E. Hazan, and S. Kale, "The multiplicative weights update method: A meta-algorithm and applications," *Theory Comput*, vol. 8, no. 6, pp. 121–164, 2012.

Pin-Yu Chen (S'10) received the B.S. degree in electrical engineering and computer science (undergraduate honors program) from National Chiao Tung University, Hsinchu City, Taiwan, in 2009, the M.S. degree in communication engineering from National Taiwan University, Taipei, Taiwan, in 2011, and is currently working towards the Ph.D. degree in electrical engineering and computer science at the University of Michigan, Ann Arbor, MI, USA. His research interests include network science, interdisciplinary network analysis, and their applications to communication systems.

Dr. Chen is a member of the Tau Beta Pi Honor Society. He was the recipient of a Chia-Lun Lo Fellowship. He was also the recipient of the IEEE GLOBECOM 2010 GOLD Best Paper Award.

Shin-Ming Cheng (S'05–M'07) received the B.S. and Ph.D. degrees in computer science and information engineering from National Taiwan University, Taipei, Taiwan, in 2000 and 2007, respectively.

From 2007 to 2012, he was a Postdoctoral Research Fellow with the Graduate Institute of Communication Engineering, National Taiwan University. Since 2012, he has been an Assistant Professor with the Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan. His research interests include mobile networks, wireless communication, information security, and complex networks.

Dr. Cheng was the recipient of the IEEE PIMRC 2013 Best Paper Award.

Kwang-Cheng Chen (M'89–SM'94–F'07) received the B.S. degree from the National Taiwan University, Taipei, Taiwan, in 1983, and the M.S. and Ph.D. degrees from the University of Maryland, College Park, MD, USA, in 1987 and 1989, respectively, all in electrical engineering.

From 1987 to 1998, he was in mobile communications and networks with Systems Service Enterprises Inc. (SSE), Communications Satellite Corporation (COMSAT), the IBM Thomas J. Watson Research Center, and National Tsing Hua University. Since 1998, he has been the Distinguished Professor and Associate Dean for academic affairs with the College of Electrical Engineering and Computer Science, National Taiwan University. From 2013 to 2014, he is a Sungkyunkwan University (SKKU) Fellow Professor. He has authored or coauthored over 250 technical papers and more than 20 granted US patents. He co-edited (with R. DeMarca) *Mobile WiMAX* (Wiley, 2008), authored *Principles of Communications* (River, 2009), coauthored (with R. Prasad) *Cognitive Radio Networks* (Wiley, 2009), and coauthored a few award-winning papers published in the IEEE ComSoc journals and conferences. His research interests include wireless communications, cognitive science, and network science.

Dr. Chen has been actively involving in the organization of various IEEE conferences as General/Technical Program Committee (TPC) Chair/Co-Chair and has served in editorships with a few IEEE journals and many international journals and in various positions with IEEE and various societies. He also actively participates in and has contributed essential technology to various IEEE 802, Bluetooth, and 3GPP wireless standards. He was the recipient of a number of awards including the 2011 IEEE COMSOC WTC Recognition Award.