On Secure Wireless Communications for IoT Under Eavesdropper Collusion

Yuanyu Zhang, Yulong Shen, Hua Wang, Jianming Yong, *Member, IEEE*, and Xiaohong Jiang, *Senior Member, IEEE*

***Abstract*—Wireless communication is one of the key technologies that actualize the Internet of Things (IoT) concept into the real world. Understanding the security performance of wireless communications lays the foundation for the security management of IoT. Eavesdropper collusion represents a significant threat to wireless communication security, while physical-layer security serves as a promising approach to providing a strong form of security guarantee. This paper studies the important secrecy outage performance of wireless communications under eavesdropper collusion, where the physical layer security is adopted to counteract such attack. Based on the classical Probability Theory, we first conduct analysis on the secrecy outage of the simple noncolluding case in which eavesdroppers do not collude and operate independently. For the secrecy outage analysis of the more hazardous M-colluding scenario, where any M eavesdroppers can combine their observations to decode the message, the techniques of Laplace transform, keyhole contour integral, and Cauchy Integral Theorem are jointly adopted to work around the highly cumbersome multifold convolution problem involved in such analysis, such that the related signal-to-interference ratio modeling for all colluding eavesdroppers can be conducted and thus the corresponding secrecy outage probability can be analytically determined. Finally, simulation and numerical results are provided to illustrate our theoretical achievements. An interesting observation suggests that the SOP increases first superlinearly and then sublinearly with M.**

***Note to Practitioners*—This paper was motivated by the problem of securing the important data collection process for the Internet of Things (IoT) applications based on the wireless communication technologies (e.g., wireless sensor networks). The traditional cryptography security method might not be well suited for the highly distributed data collection process in IoT, where edge points (e.g., sensing nodes) are extremely constrained in terms of computing capabilities. This paper explored the application of the promising physical layer (PHY) security approach in the data collection process of IoT, which simply exploits the inherent randomness of wireless channels and noise to provide a strong form of security guarantee. To model the security performance of the data collection process, we mathematically characterized the outage probability that the data is successfully intercepted by the malicious adversaries (i.e., eavesdroppers). Such outage metric serves as the guideline for system designers to initiate the tradeoff between data collection efficiency and data security guarantee, and thus to identify the suitable PHY security schemes for different IoT application scenarios. Simulation and numerical results were provided in this study, while the experiments based on the test bed or real-world implementations have not been performed, which remain as our future research work.**

***Index Terms*—Eavesdropper collusion, Internet of Things (IoT), physical layer security, secrecy outage performance, wireless communication.**

I. Introduction

THE Internet of Things (IoT), where a variety of *things* (e.g., people, sensors, mobile phones) can interact with one another from any place in the world through an Internet-like infrastructure with unique addressing schemes, serves a crucial architecture for a wide range of promising applications such as e-health, intelligent transportation systems, and environmental monitoring [1], [2]. As one of the key enabling technologies of IoT, wireless communication (e.g., wireless sensor network) plays an important role in collecting the data perceived by edge points (e.g., sensing nodes) from the surroundings so as to link the real world with the digital world. Due to the open nature of the wireless medium, data collection through wireless communication may suffer from eavesdropping attacks by unauthorized *things* (eavesdroppers), posing a significant threat to the IoT security [1]. Thus, extensive research efforts need to be devoted to the security performance study of wireless communications in the data collection process, which serves as a foundation for the security management of IoT.

Conventionally, wireless communication protocols usually use secret keys of various kinds to ensure encryption capability such that, even if an eavesdropper captures packets, it cannot decrypt them without the key. However, as the computing capability of eavesdroppers continues to advance (e.g., by adopting the quantum computing technology [3]), these encryption protocols may face increasingly high risk of being broken in the future. Also, the secret key management and distribution are costly and complex to be implemented in decentralized wireless IoT data collection, where nodes are highly constrained in terms of the computing capability [1]. This is why there is an increasing interest recently in employing the physical-layer security approach to provide a strong form of security guarantee for wireless communications (see [4] and

Manuscript received June 01, 2015; revised September 10, 2015; accepted October 28, 2015. Date of publication December 09, 2015; date of current version June 30, 2016. This paper was recommended for publication by Associate Editor J. Jobin and Editor M. Zhou upon evaluation of the reviewers' comments. This work was supported in part by the Japan JSPS under Grant 15H02692 and the China NSFC under Grant 61571352, Grant 61373173, and Grant U153620014. *(Corresponding author: Yulong Shen.)*

Y. Zhang and X. Jiang are with the School of Systems Information Science, Future University Hakodate, Hokkaido 041-8655, Japan, and also with the School of Computer Science and Technology, Xidian University, Xidian 710071, China (e-mail: yy90zhang@gmail.com; jiang@fun.ac.jp).

Y. Shen is with the School of Computer Science and Technology, Xidian University, Xidian 710071, China (e-mail: ylshen@mail.xidian.edu.cn).

H. Wang is with the Centre of Applied Informatics, College of Engineering and Science, Victoria University, Australia (e-mail: hua.wang@vu.edu.au).

J. Yong is with the School of Management and Enterprise, Faculty of Business, Education, Law and Arts, University of Southern Queensland, Australia (e-mail: Jianming.Yong@usq.edu.au).

Digital Object Identifier 10.1109/TASE.2015.2497663

references therein). The basic principle of the physical-layer security approach is to exploit the inherent randomness of noise and wireless channels to ensure the confidentiality of messages against any eavesdropper regardless of its computing capability and network knowledge (e.g., coding scheme used in legitimate node). Compared with the encryption protocols, the physical-layer security approach can offer some significant advantages, such as no need for secret key management/distribution and an everlasting security guarantee. Actually, applying the physical-layer security approach to provide security guarantee for secret key generation and distribution has also attracted considerable attention [5]–[7]. Thus, physical-layer security technology also serves as a promising complement to the encryption protocols to enhance the security performance of wireless communications. Therefore, physical-layer security could be a new solution to secure the wireless data collection in the IoT.

By now, extensive research efforts have been devoted to the security performance study of wireless communications under physical layer security [8]–[10]. To simplify their analysis, these studies usually consider the noncolluding scenario, where each eavesdropper works independently and decodes the message solely based on its own observation. In a real network environment, however, eavesdroppers may work under a colluding manner such that they can exchange and combine their observations to conduct more hazardous attacks. Based on such observation, some recent works [11]–[24] considered the eavesdropper collusion behavior and explored its impact on the wireless communication security performance in terms of the scaling laws of secrecy capacity and eavesdropper-tolerance capability, secure connection probability, etc. (Please refer to Section VI for related works). These works indicate that the collusion can help eavesdroppers to significantly improve their capability of decoding the secret message, so eavesdropper collusion represents a more hazardous threat to wireless communication security.

It is notable that secrecy outage performance, i.e., the possibility that a secret message can be successfully decoded by eavesdropper(s), serves as an important metric for wireless communication security evaluation and secure transmission scheme design [25]. Despite extensive studies on the security performance of wireless communications, the analysis of the secrecy outage performance under eavesdropper collusion and physical layer security remains a technique challenge. This is mainly due to that such secrecy outage analysis usually involves highly cumbersome multi-fold convolutions related to the modeling of the probability density function (pdf)/cumulative distribution function (cdf) of the aggregate signal-to-interference ratio (SIR) of all colluding eavesdroppers. This paper focuses on the secrecy outage performance study of wireless communications under the eavesdropper collusion, where the physical layer security is adopted to counteract such attack. The main contributions are summarized as follows.

- Based on the classical Probability Theory, we first derive the secrecy outage probability (SOP) for the simple noncolluding case, where each eavesdropper works independently and decodes the message solely based on its own observation.

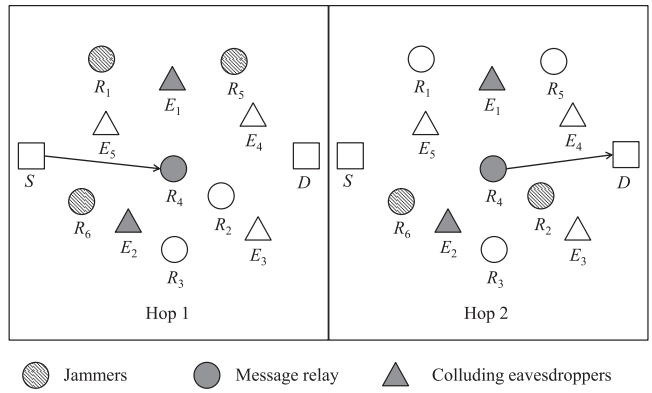

Fig. 1. Network model: a sensing node S is communicating with a sink node D in two hops with the help of relays $R_1, R_2, \ldots, R_n, n = 6$. R_4 is selected as the message relay base on the opportunistic relaying scheme. In Hop 1, R_1, R_5 and R_6 are jammers that generate artificial noise, while R_2 and R_6 are jammers in Hop 2. $E_1, E_2, \ldots, E_m, m = 5$ are eavesdroppers that try to intercept the message, and E_1 and E_2 are colluding eavesdroppers.

- For the secrecy outage analysis of the more hazardous M-colluding scenario, where any M eavesdroppers can combine their observations to decode the message, the techniques of Laplace transform, keyhole contour integral and Cauchy Integral Theorem are jointly adopted to work around the highly cumbersome muti-fold convolution problem involved in such analysis, such that the related SIR modeling for all colluding eavesdroppers can be conducted and thus the corresponding SOP can be analytically determined.
- Finally, we provide simulation and numerical results to validate our theoretical analysis and also to illustrate our theoretical findings.

The remainder of this paper is organized as follows. Section II introduces the system model and problem formulation. The SOP analysis of non-colluding case is presented in Section III. In Section IV, we analyze the SOP of M-colluding case. Simulation/numerical results and the corresponding discussions are provided in Section V. Finally, we introduce the related works in Section VI and conclude this paper in Section VII.

II. System Model and Problem Formulation

A. Network Model

The two-hop wireless network, where a message is first transmitted from its source to intermediate relay(s) and then forwarded by the relay(s) to its destination, serves as an important network model widely adopted in the literature [26], [27]. Actually, understanding the performance of such basic networks lays the foundation for the performance study of general wireless networks. In this paper, we consider a two-hop wireless network (depicted in Fig. 1) for collecting data in the IoT, consisting of a sensing node S, a sink node D, n legitimate relays $R_1, R_2, \ldots, R_n$, and m passive eavesdroppers $E_1, E_2, \ldots, E_m$ of unknown channel information. Each node employs a single antenna and operates in a half-duplex mode. The direct link between S and D is assumed to be unavailable due to deep fading or limited transmit power. The n relays assist in forwarding the message from the sensing node to the sink node

while preventing the eavesdroppers from intercepting the message. We assume that time is slotted and all channels, suffering from Rayleigh fading, remain constant during one time slot and vary randomly and independently from slot to slot. The channel coefficient $h_{i,j}$ of link $i \to j$ is modeled as a complex zero-mean Gaussian random variable with unit variance and thus $|h_{i,j}|^2$ is exponentially distributed with unit mean. All channel coefficients $(S-R, R-D, R-R, S-E, R-E)$ are assumed to be independent and identically distributed (i.i.d.). The network is assumed to be interference-limited, and thus the noise at each receiver is negligible. In this paper, we will explore the SOP performance of such general two-hop wireless networks under eavesdropper collusion. It is notable that the schemes and results in this paper apply to any wireless communication technology suitable for ad hoc scenarios (e.g., Bluetooth, ZigBee and IEEE 802.15.4) [28].

Remark 1: Such a homogeneous network model has been attractive in the security performance study of wireless communications [8]–[10], [12]. Actually, the mathematical tractability of such model allows us to gain some important insights into the structure of SOP analysis, so the results under this model can still help us to have a basic understanding on the SOP performance of wireless communications, and the related theoretical analysis can shed light on the security performance study for more general heterogeneous models.

B. Opportunistic Relaying and Cooperative Jamming

The $S-D$ transmission is conducted in two hops with the help of the opportunistic relaying scheme [29], where a relay R_b with the largest $\min\{|h_{S,R_j}|^2, |h_{R_j,D}|^2\}$ announces itself as the message relay in a distributed manner before the transmission. We assume that only one $S-D$ transmission, including the relay selection, can be conducted in one time slot. To ensure the secure transmission, we consider the cooperative jamming [30], a typical physical layer security method where helper jammers generate artificial noise to counteract the eavesdroppers.

In the first hop, S transmits its message to R_b, while relays with indices in $\mathcal{J}_1 = \{j | j \neq b, |h_{R_j,R_b}|^2 < \tau\}$ serve as helper jammers in the cooperative jamming scheme. Here, τ is the noise-generating threshold to mitigate the interference at intended receivers. The received signal y_{R_b} at R_b can be formulated as

$$y_{R_b} = \sqrt{P} h_{S,R_b} x + \sum_{j \in \mathcal{J}_1} \sqrt{P} h_{R_j,R_b} x_j \tag{1}$$

and y_{E_i} at $E_i, i = 1, 2, \ldots, m$ can be formulated as

$$y_{E_i} = \sqrt{P} h_{S,E_i} x + \sum_{j \in \mathcal{J}_1} \sqrt{P} h_{R_j,E_i} x_j \tag{2}$$

where P is the common transmit power, x denotes the signal from the sensing node S, and x_j is the artificial noise signal from jammer R_j. Hence, the received SIR at R_b and that at E_i can be given by

$$\gamma_{S,R_b} = \frac{|h_{S,R_b}|^2}{\sum_{j \in \mathcal{J}_1} |h_{R_j,R_b}|^2}, \ \gamma_{S,E_i} = \frac{|h_{S,E_i}|^2}{\sum_{j \in \mathcal{J}_1} |h_{R_j,E_i}|^2} \tag{3}$$

where γ_{S,R_b} denotes the SIR from S to R_b and γ_{S,E_i} that from S to E_i.

If R_b is successful in decoding y_{R_b}, it re-encodes the message and then sends it to D in the second hop. Meanwhile, relays with indices in $\mathcal{J}_2 = \{j | j \neq b, |h_{R_j,D}|^2 < \tau\}$ serve as helper jammers to generate artificial noise. If R_b fails to decode y_{R_b}, the transmission will be suspended. We assume that R_b will send back an ACK message to inform S whether a decoding failure happens or not, based on which S will then decide whether to suspend the transmission or not. If a transmission suspension happens during one time slot, it will end at the end of that time slot, and retransmission of the suspended message will be conducted in the next time slot. The received signals and SIR's at D and E_i can be determined in the same way as in the first hop. A legitimate receiver (eavesdropper) is said successful in decoding the received signal if its SIR is above a threshold $\beta(\beta_e)$.

C. Eavesdropper Scenarios

Regarding the eavesdropper behavior, we focus on the following two scenarios.

1) **Noncolluding case**: each eavesdropper works independently and decodes the message from the sensing node solely based on its available observations, i.e., the first-hop observation if the transmission is suspended in the second hop or the cumulative observation in both hops, otherwise.
2) **M-colluding case**: any M eavesdroppers (say, $E_1, E_2, \ldots, E_M, 1 \leq M \leq m, M = 2$ as illustrated in Fig. 1) can combine their available observations to decode the message from the sensing node.

Here, M is referred to as the *collusion intensity* to quantify the level of eavesdropper collusion in this paper. As assumed in [11]–[24], these colluding eavesdroppers can be treated as a super eavesdropper with M antennas, whose SIR is given by the aggregate SIR of all antennas.

D. Problem Formulation

In this paper, we adopt the secrecy outage probability (SOP) to characterize the secrecy outage performance, which is defined as the probability that the received SIR of at least one of the eavesdroppers is above some threshold for a $S-D$ transmission. Therefore, by defining A as the event that the transmission is suspended in the second hop, the SOP P_{so}^{nc} for noncolluding case can be formulated as

$$P_{so}^{nc} = P\left(\bigcup_{i=1}^{m} \{\gamma_{S,E_i} \geq \beta_e\}, A\right) + P\left(\bigcup_{i=1}^{m} \{\gamma_{S,E_i} + \gamma_{R_b,E_i} \geq \beta_e\}, \bar{A}\right) \tag{4}$$

where $\bar{A}$ is the complement of event A. Similarly, the SOP P_{so}^{c} for M-colluding case can be formulated as

$$P_{so}^{c} = P\left(\left\{\gamma_{sp}^{A} \geq \beta_e \text{ or } \bigcup_{i=M+1}^{m} \{\gamma_{S,E_i} \geq \beta_e\}\right\}, A\right) + P\left(\left\{\gamma_{sp}^{\bar{A}} \geq \beta_e \text{ or } \bigcup_{i=M+1}^{m} \{\gamma_{S,E_i} + \gamma_{R_b,E_i} \geq \beta_e\}\right\}, \bar{A}\right) \tag{5}$$

where $\gamma_{sp}^{A} = \sum_{i=1}^{M} \gamma_{S,E_i}$ denotes the aggregate SIR of M colluding eavesdroppers under event A and $\gamma_{sp}^{\bar{A}} = \sum_{i=1}^{M} \gamma_{S,E_i} + \gamma_{R_b,E_i}$ denotes that under event $\bar{A}$.

III. Secrecy Outage Performance Under the Noncolluding Case

Here, we derive the SOP of the noncolluding case, for which we will first establish the following lemma regarding the probability that the transmission is suspended in the second hop, conditioned on the number of jammers in the first hop.

Lemma 1: Define the number of jammers in Hop h ($h = 1, 2$) by J_h and the event $J_h = l$ by J_h^l. For a two-hop $S-D$ transmission under the opportunistic relaying and cooperative jamming schemes as described in the Section II-B, under the condition J_1^l the probability $p_{A|J_1^l}$ that the transmission is suspended in the second hop can be determined as

$$p_{A|J_1^l} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k \frac{1}{2k-1} \left[k \left(\frac{1-e^{-(1+\beta)\tau}}{(1-e^{-\tau})(1+\beta)} \right)^l + (k-1) \left(\frac{1-e^{-(2k\beta+1)\tau}}{(1-e^{-\tau})(2k\beta+1)} \right)^l \right]. \quad (6)$$

Proof: Please refer to Appendix A. ∎

Based on Lemma 1, the SOP of the noncolluding case can be obtained by applying the law of total probability, which is given by the following theorem.

Theorem 1: Consider a two-hop wireless data collection process in the IoT as shown in Fig. 1. For the $S-D$ transmission under the opportunistic relaying, cooperative jamming and non-colluding eavesdropper case as described in Section II, the corresponding SOP P_{so}^{nc} can be formulated as

$$P_{so}^{nc} = 1 - \sum_{l=0}^{n-1} \sum_{t=0}^{n-1} \binom{n-1}{l} \binom{n-1}{t} (1-e^{-\tau})^{l+t} \times e^{-(2n-2-l-t)\tau} \left[(1-p_{A|J_1^l}) p_2^m + p_{A|J_1^l} p_1^m \right] \quad (7)$$

where

$$p_1 = 1 - \left(\frac{1}{1+\beta_e} \right)^l,$$

$$p_2 = \int_0^{\beta_e} \left[1 - \frac{1}{(1+\beta_e - x)^t} \right] \frac{l}{(1+x)^{l+1}} \mathrm{d}x$$

and $p_{A|J_1^l}$ is given as in (6).

Proof: We start the proof with the first term in (4). By the law of total probability,

$$\begin{aligned} &P\left(\bigcup_{i=1}^{m} \{\gamma_{S,E_i} \geq \beta_e\}, A \right) \\ &= \mathbb{E}_{J_1} \left[P\left(\bigcup_{i=1}^{m} \{\gamma_{S,E_i} \geq \beta_e\}, A \,\middle|\, J_1^l \right) \right] \\ &= \sum_{l=0}^{n-1} P\left(\bigcup_{i=1}^{m} \{\gamma_{S,E_i} \geq \beta_e\}, A \,\middle|\, J_1^l \right) P(J_1^l) \\ &= \sum_{l=0}^{n-1} P\left(\bigcup_{i=1}^{m} \{\gamma_{S,E_i} \geq \beta_e\} \,\middle|\, J_1^l \right) P(J_1^l) p_{A|J_1^l} \\ &= \sum_{l=0}^{n-1} \left[1 - P\left(\bigcap_{i=1}^{m} \{\gamma_{S,E_i} < \beta_e\} \,\middle|\, J_1^l \right) \right] P(J_1^l) p_{A|J_1^l} \\ &\overset{(a)}{=} \sum_{l=0}^{n-1} \left[1 - P\left(\gamma_{S,E_i} < \beta_e | J_1^l \right)^m \right] P(J_1^l) p_{A|J_1^l}, \end{aligned} \quad (8)$$

where (a) follows from the fact that $\gamma_{S,E_i}, i = 1, 2, \ldots, m$ are i.i.d. given the event J_1^l. Next, we consider the cdf of γ_{S,E_i} under the condition J_1^l, denoted by $F_{\gamma_{S,E_i}}(x|J_1^l)$. Based on the γ_{S,E_i} in (3), we have

$$\begin{aligned} &F_{\gamma_{S,E_i}}\left(x|J_1^l\right) \\ &= P\left(\gamma_{S,E_i} < x \,\middle|\, J_1^l\right) \\ &= P\left(|h_{S,E_i}|^2 < x \sum_{j \in \mathcal{J}_1} |h_{R_j,E_i}|^2 \,\middle|\, J_1^l \right) \\ &= 1 - \mathbb{E}_{\{|h_{R_j,E_i}|^2, j \in \mathcal{J}_1\}} \left[e^{-x \sum_{j \in \mathcal{J}_1} |h_{R_j,E_i}|^2} \,\middle|\, J_1^l \right] \\ &= 1 - \prod_{j \in \mathcal{J}_1} \mathbb{E}_{|h_{R_j,E_i}|^2} \left[e^{-x|h_{R_j,E_i}|^2} \right] \\ &= 1 - \left(\frac{1}{1+x} \right)^l. \end{aligned} \quad (9)$$

From (9), it is easy to see that

$$P\left(\gamma_{S,E_i} < \beta_e | J_1^l\right) = 1 - \left(\frac{1}{1+\beta_e} \right)^l = p_1. \quad (10)$$

As J_1 is a binomial random variable, it follows that

$$P(J_1^l) = \binom{n-1}{l} (1-e^{-\tau})^l e^{-(n-1-l)\tau}. \quad (11)$$

Hence, substituting (10) and (11) into (8) yields

$$\begin{aligned} &P\left(\bigcup_{i=1}^{m} \{\gamma_{S,E_i} \geq \beta_e\}, A \right) \\ &= \sum_{l=0}^{n-1} \binom{n-1}{l} (1-e^{-\tau})^l e^{-(n-1-l)\tau} (1-p_1^m) p_{A|J_1^l}. \end{aligned} \quad (12)$$

We now turn to consider the second term in (4). Likewise, taking the expectation of (4) in terms of J_1 and J_2 yields

$$\begin{aligned} &P\left(\bigcup_{i=1}^{m} \{\gamma_{S,E_i} + \gamma_{R_b,E_i} \geq \beta_e\}, \bar{A} \right) \\ &= \sum_{l=0}^{n-1} \sum_{t=0}^{n-1} \left[1 - P\left(\gamma_{S,E_i} + \gamma_{R_b,E_i} < \beta_e | J_1^l, J_2^t \right)^m \right] \\ &\quad \times P(\bar{A}|J_1^l) P(J_1^l) P(J_2^t) \end{aligned} \quad (13)$$

where it is straightforward to see $P(\bar{A}|J_1^l) = 1 - p_{A|J_1^l}$, and

$$P\left(J_2^t\right) = \binom{n-1}{t} (1-e^{-\tau})^t e^{-(n-1-t)\tau}. \quad (14)$$

Similar to (9), the cdf of γ_{R_b,E_i} under the condition J_2^t can be given by $F_{\gamma_{R_b,E_i}}(x|J_2^t) = 1 - ((1)/(1+x))^t$. From (9), the pdf of γ_{S,E_i} under the condition J_1^l is $f_{\gamma_{S,E_i}}(x|J_1^l) = (l)/((1+x)^{l+1})$. Hence

$$\begin{aligned} &P\left(\gamma_{S,E_i} + \gamma_{R_b,E_i} < \beta_e | J_1^l, J_2^t\right) \\ &= \int_0^{\beta_e} \left[1 - \frac{1}{(1+\beta_e - x)^t} \right] \frac{l}{(1+x)^{l+1}} \mathrm{d}x = p_2. \end{aligned} \quad (15)$$

Substituting (11), (15), $P(\bar{A}|J_1^l)$ and $P(J_2^t)$ into (13) yields

$$P\left(\bigcup_{i=1}^{m}\{\gamma_{S,E_i}+\gamma_{R_b,E_i}\geq\beta_e\},\bar{A}\right) = \sum_{l=0}^{n-1}\sum_{t=0}^{n-1}\binom{n-1}{l}\binom{n-1}{t} \times(1-e^{-\tau})^{l+t}e^{-(2n-2-l-t)\tau}\left(1-p_2^m\right)\left(1-p_{A|J_1^l}\right). \quad (16)$$

Finally, the theorem follows from summing (12) and (16). ■

IV. Secrecy Outage Performance Under the M-Colluding Case

Here, the secrecy outage performance of the M-colluding case is investigated, for which we will first derive the cdf of the aggregate SIR's γ_{sp}^{A} and $\gamma_{sp}^{\bar{A}}$ of any M colluding eavesdroppers, based on which we then determine the corresponding SOP.

A. Aggregate SIR Analysis

From Section II-D, we can see that the aggregate SIR's γ_{sp}^{A} and $\gamma_{sp}^{\bar{A}}$ are the sums of multiple i.i.d. random variables. The derivation of their cdfs usually involves a multi-fold convolution, which is highly cumbersome in general. To work around this problem, we first take the Laplace transforms of their pdfs and then compute the related inverse Laplace transform by applying the keyhole contour integral and Cauchy Integral Theorem. Finally, the cdfs can be obtained from the pdfs. The related lemma and proof are summarized as follows.

Lemma 2: Define the cdf of γ_{sp}^{A} under event A by $F_M(x)$ and that of $\gamma_{sp}^{\bar{A}}$ under event $\bar{A}$ by $F_{2M}(x)$. For a two-hop $S-D$ transmission under the opportunistic relaying, cooperative jamming and M-colluding eavesdropper case as described in Section II, under the conditions J_1^l and J_2^t, $F_M(x)$ can be given by

$$F_M(x)=l^M\sum_{k=0}^{\lfloor\frac{M-1}{2}\rfloor}\binom{M}{2k+1}(-\pi^2)^k\int_0^\infty\frac{1}{u}(1-e^{-xu})e^{-Mu} \times EI_l(u)^{M-2k-1}\left(\frac{u^l}{l!}\right)^{2k+1}\mathrm{d}u \quad (17)$$

and $F_{2M}(x)$ can be given by

$$F_{2M}(x)=(lt)^M\sum_{k=0}^{\lfloor\frac{M-1}{2}\rfloor}\binom{M}{2k+1}(-\pi^2)^k \times\int_0^\infty\frac{1}{u}(1-e^{-xu})e^{-2Mu} \times\left[EI_l(u)EI_t(u)-\pi^2\frac{u^{l+t}}{l!t!}\right]^{M-2k-1} \times\left[EI_l(u)\frac{u^t}{t!}+EI_t(u)\frac{u^l}{l!}\right]^{2k+1}\mathrm{d}u \quad (18)$$

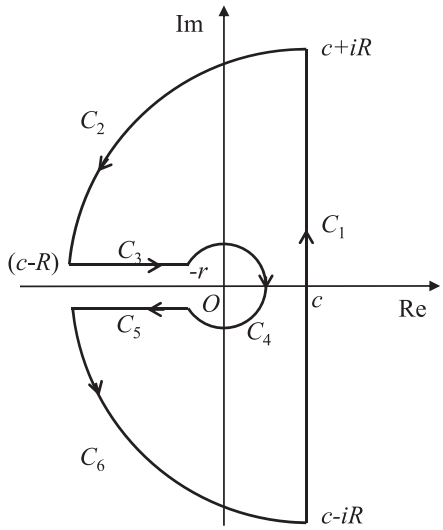

Fig. 2. Diagram of the keyhole contour, where C_1 is a vertical line from $c-iR$ to $c+iR$, C_2 and C_6 forms a large (almost) semi-circle centered at $s=c$ with radius R, C_3 is a line from $c-R$ to $-r$, C_4 is a small (almost) circle centered at the origin with radius r, C_5 is a line from $-r$ to $c-R$.

where

$$EI_l(u)=\frac{u^l}{l!}\left(\sum_{k=1}^{l}\frac{1}{k}-\gamma-\ln u\right)-\sum_{k=0,k\neq l}^{\infty}\frac{u^k}{(k-l)k!}$$

and $\gamma=0.5772156649\ldots$ is the Euler's constant.

Proof: Define $f_M(x)$ the pdf of γ_{sp}^{A} and $f_M^*(s)$ its Laplace transform. Based on the pdf $f_{\gamma_{S,E_i}}(x)=(l)/((1+x)^{l+1})$ of γ_{S,E_i} under the condition J_1^l, $f_M^*(s)$ can be determined by the convolution property of Laplace transform as

$$f_M^*(s)=(le^sE_{l+1}(s))^M \quad (19)$$

where $E_{l+1}(s)=\int_1^\infty(e^{-sv})/(v^{l+1})\mathrm{d}v$ is the generalized exponential integral.

Next, $f_M(x)$ can be obtained by taking the inverse Laplace transform of $f_M^*(s)$, that is,

$$f_M(x)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}e^{sx}f_M^*(s)\mathrm{d}s \quad (20)$$

where $c>0$ is an arbitrary constant greater than the real part of all singularities of $f_M^*(s)$. Since $E_{l+1}(s)$ is analytical in the complex plane except its branch cut along the negative real axis and branch point at the origin, the above integral can be evaluated as a part of the integral along a keyhole contour Ω [31], as illustrated in Fig. 2. As $f_M^*(s)$ is analytical in Ω, by the Cauchy Integral Theorem, we have $\int_\Omega e^{sx}f_M^*(s)\mathrm{d}s=0$. Hence

$$\begin{aligned}f_M(x)=&\frac{1}{2\pi i}\lim_{R\to\infty}\int_{C_1}e^{sx}f_M^*(s)\mathrm{d}s\\ =&-\frac{1}{2\pi i}\lim_{R\to\infty,r\to0}\int_{C_2}+\int_{C_3}+\int_{C_4}+\int_{C_5}\\ &+\int_{C_6}e^{sx}f_M^*(s)\mathrm{d}s\\ =&-\frac{1}{2\pi i}\lim_{R\to\infty,r\to0}\int_{C_3}+\int_{C_5}e^{sx}f_M^*(s)\mathrm{d}s.\end{aligned} \quad (21)$$

To see this, we turn to prove that the integrals along C_2, C_4 and C_6 vanish in the limit. First, letting $s = c + Re^{i\theta}, \theta \in [\pi/2, \pi]$ for any point s on C_2 yields

$$\begin{aligned}
&\lim_{R\to\infty}\left|\int_{C_2} e^{sx} f_M^*(s)\mathrm{d}s\right| \\
&= \lim_{R\to\infty}\left|\int_{\frac{\pi}{2}}^{\pi} e^{x(c+Re^{i\theta})} f_M^*(c+Re^{i\theta}) iRe^{i\theta}\mathrm{d}\theta\right| \\
&\le \lim_{R\to\infty}\int_{\frac{\pi}{2}}^{\pi}\left|e^{x(c+Re^{i\theta})}\right|\left|f_M^*(c+Re^{i\theta})\right|\left|iRe^{i\theta}\right|\mathrm{d}\theta \\
&\le \lim_{R\to\infty}\max_{\theta\in[\pi/2,\pi]}\left|f_M^*(c+Re^{i\theta})\right| Re^{xc}\int_{\frac{\pi}{2}}^{\pi} e^{xR\cos\theta}\mathrm{d}\theta \\
&= \lim_{R\to\infty}\max_{\theta\in[\pi/2,\pi]}\left|f_M^*(c+Re^{i\theta})\right| Re^{xc}\int_0^{\frac{\pi}{2}} e^{-xR\sin\alpha}\mathrm{d}\alpha. \quad (22)
\end{aligned}$$

Since $\sin\alpha \ge (2\alpha)/(\pi)$ for any $\alpha \in [0, (\pi)/(2)]$, then

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} e^{-xR\sin\alpha}\mathrm{d}\alpha &\le \int_0^{\frac{\pi}{2}} e^{-2xR\alpha/\pi}\mathrm{d}\alpha \\
&= \frac{\pi}{2xR}(1-e^{-xR}) \le \frac{\pi}{2xR} \quad (23)
\end{aligned}$$

and thus

$$\begin{aligned}
&\lim_{R\to\infty}\left|\int_{C_2} e^{sx} f_M^*(s)\mathrm{d}s\right| \\
&\qquad \le \lim_{R\to\infty}\frac{\pi e^{xc}}{2x}\max_{\theta\in[\pi/2,\pi]}\left|f_M^*(c+Re^{i\theta})\right|. \quad (24)
\end{aligned}$$

From [32, eq. (5.1.51)], we know

$$e^s E_{l+1}(s) \sim \frac{1}{s} - \frac{l+1}{s^2} + \frac{(l+1)(l+2)}{s^3} + \cdots \quad (25)$$

hence

$$\left|e^{c+Re^{i\theta}} E_{l+1}(c+Re^{i\theta})\right| = O(1/R) \quad (26)$$

and then

$$\max_{\theta\in[\pi/2,\pi]}\left|f_M^*(c+Re^{i\theta})\right| = O(1/R^M) \quad (27)$$

as $R \to \infty$. Therefore

$$\lim_{R\to\infty}\left|\int_{C_2} e^{sx} f_M^*(s)\mathrm{d}s\right| = 0. \quad (28)$$

Likewise, it can be easily seen that the integral along C_6 vanishes as R tends to infinity and that along C_4 vanishes as r tends to zero.

Now, we proceed to evaluate the integrals along C_3 and C_5. By letting $s = ue^{i\pi}$ for the integral along C_3 and $s = ue^{-i\pi}$ for that along C_5, we have

$$\begin{aligned}
f_M(x) &= -\frac{1}{2\pi i}\lim_{R\to\infty, r\to 0}\int_{C_3} + \int_{C_5} e^{sx} f_M^*(s)\mathrm{d}s \\
&= -\frac{1}{2\pi i}\lim_{R\to\infty, r\to 0}\int_{c-R}^{-r} + \int_{-r}^{c-R} e^{sx} f_M^*(s)\mathrm{d}s \\
&= \frac{1}{2\pi i}\int_0^{\infty} e^{-xu}\left(f_M^*(ue^{-i\pi}) - f_M^*(ue^{i\pi})\right)\mathrm{d}u \\
&= \frac{l^M}{2\pi i}\int_0^{\infty} e^{-(x+M)u} \\
&\quad \times \left[E_{l+1}(ue^{-i\pi})^M - E_{l+1}(ue^{i\pi})^M\right]\mathrm{d}u. \quad (29)
\end{aligned}$$

From [32, eq. (5.1.12)], we have

$$E_{l+1}(ue^{\pm i\pi}) = EI_l(u) \mp i\pi\frac{u^l}{l!} \quad (30)$$

where

$$EI_l(u) = \frac{u^l}{l!}\left(\sum_{k=1}^{l}\frac{1}{k} - \gamma - \ln u\right) - \sum_{k=0,k\ne l}^{\infty}\frac{u^k}{(k-l)k!}$$

and $\gamma = 0.5772156649$ is the Euler's constant. Hence,

$$\begin{aligned}
&E_{l+1}(ue^{-i\pi})^M - E_{l+1}(ue^{i\pi})^M \\
&= \left(EI_l(u) + i\pi\frac{u^l}{l!}\right)^M - \left(EI_l(u) - i\pi\frac{u^l}{l!}\right)^M \\
&= 2\pi i\sum_{k=0}^{\lfloor\frac{M-1}{2}\rfloor}\binom{M}{2k+1}(-\pi^2)^k EI_l(u)^{M-2k-1}\left(\frac{u^l}{l!}\right)^{2k+1}. \quad (31)
\end{aligned}$$

Substituting (31) into (29) yields

$$f_M(x) = \int_0^{\infty} e^{-(x+M)u}\varphi(M,l,u)\mathrm{d}u \quad (32)$$

where

$$\begin{aligned}
&\varphi(M,l,u) \\
&= l^M\sum_{k=0}^{\lfloor\frac{M-1}{2}\rfloor}\binom{M}{2k+1}(-\pi^2)^k EI_l(u)^{M-2k-1}\left(\frac{u^l}{l!}\right)^{2k+1}. \quad (33)
\end{aligned}$$

The cdf $F_M(x)$ of γ_{sp}^A can be determined via the integration of $f_M(x)$ as

$$F_M(x) = \int_0^{\infty}\frac{1}{u}(1-e^{-xu})e^{-Mu}\varphi(M,l,u)\mathrm{d}u. \quad (34)$$

Finally, taking the integral for each summand involving u in $\varphi(M,l,u)$ first and then summing the integrals yields (17).

We now turn to consider the cdf $F_{2M}(x)$ of $\gamma_{sp}^{\bar{A}}$. Similarly, we define $f_{2M}(x)$ the pdf of $\gamma_{sp}^{\bar{A}}$ and $f_{2M}^*(s)$ its Laplace transform. Again, by the convolution property of Laplace transform, we have

$$f_{2M}^*(s) = \left(lte^{2s}E_{l+1}(s)E_{t+1}(s)\right)^M \quad (35)$$

under the conditions J_1^l and J_2^t. Taking the inverse Laplace transform of $f_{2M}^*(s)$ along again the keyhole contour in Fig. 2 yields

$$\begin{aligned}
&f_{2M}(x) \\
&= \frac{(lt)^M}{2\pi i}\int_0^{\infty} e^{-(x+2M)u}[(E_{l+1}(ue^{-i\pi})E_{t+1}(ue^{-i\pi}))^M \\
&\quad - (E_{l+1}(ue^{i\pi})E_{t+1}(ue^{i\pi}))^M]\mathrm{d}u \\
&= \int_0^{\infty} e^{-(x+2M)u}\phi(M,l,t,u)\mathrm{d}u \quad (36)
\end{aligned}$$

where

$$\phi(M,l,t,u) = (lt)^M \sum_{k=0}^{\lfloor \frac{M-1}{2} \rfloor} \binom{M}{2k+1} (-\pi^2)^k \times \left[EI_l(u)\frac{u^t}{t!} + EI_t(u)\frac{u^l}{l!} \right]^{2k+1} \times \left[EI_l(u)EI_t(u) - \pi^2 \frac{u^{l+t}}{l!t!} \right]^{M-2k-1} \quad (37)$$

The cdf $F_{2M}(x)$ of $\gamma_{sp}^{\bar{A}}$ is determined as

$$F_{2M}(x) = \int_0^\infty \frac{1}{u}(1-e^{-xu})e^{-2Mu}\phi(M,l,t,u)\mathrm{d}u. \quad (38)$$

Likewise, taking the integral for each summand in $\phi(M,l,t,u)$ first and then summing the integrals yields (18). ∎

Remark 2: Note that, from [32, (5.1.45)], $EI_l(u)$ can also be formulated as

$$EI_l(u) = (-u)^l \mathrm{Re}(\Gamma(-l,-u)) \quad (39)$$

for the simplicity of numerical calculation, where $\mathrm{Re}(\cdot)$ denotes the real part of a complex number and $\Gamma(-l,-u)$ is the upper incomplete gamma function with parameters $-l$ and $-u$.

B. SOP Modeling

Based on Lemma 2, the SOP of the M-colluding case is given by the following theorem.

Theorem 2: Consider a two-hop wireless data collection process in the IoT as shown in Fig. 1. For the $S-D$ transmission under the opportunistic relaying, cooperative jamming and M-colluding eavesdropper case as described in Section II, the corresponding SOP P_{so}^c can be formulated as

$$P_{so}^c = 1 - \sum_{l=0}^{n-1}\sum_{t=0}^{n-1} \binom{n-1}{l}\binom{n-1}{t} \times (1-e^{-\tau})^{l+t} e^{-(2n-2-l-t)\tau} \times \left[(1-p_{A|J_1^l}) p_2^{m-M} F_{2M}(\beta_e) + p_{A|J_1^l} p_1^{m-M} F_M(\beta_e) \right] \quad (40)$$

where

$$p_1 = 1 - \left(\frac{1}{1+\beta_e}\right)^l,$$
$$p_2 = \int_0^{\beta_e} \left[1 - \frac{1}{(1+\beta_e-x)^t}\right] \frac{l}{(1+x)^{l+1}}\mathrm{d}x.$$

$p_{A|J_1^l}$ is given as in (6), and $F_M(\beta_e)$ and $F_{2M}(\beta_e)$ can be directly obtained from (17) and (18).

Proof: Similar to the proof of Theorem 1, the first term in (5) can be determined by taking its expectation in terms of J_1 as

$$P\left(\left\{\gamma_{sp}^A \geq \beta_e \text{ or } \bigcup_{i=M+1}^m \{\gamma_{S,E_i} \geq \beta_e\}\right\}, A\right) = \mathbb{E}_{J_1}\left[P\left(\left\{\gamma_{sp}^A \geq \beta_e \text{ or } \bigcup_{i=M+1}^m \{\gamma_{S,E_i} \geq \beta_e\}\right\}, A | J_1^l\right)\right]$$
$$= \sum_{l=0}^{n-1}\left[1 - P\left(\gamma_{sp}^A < \beta_e \,|J_1^l\right) P\left(\gamma_{S,E_i} < \beta_e \,|J_1^l\right)^{m-M}\right] \times P(J_1^l) p_{A|J_1^l}$$
$$= \sum_{l=0}^{n-1}\binom{n-1}{l}(1-e^{-\tau})^l e^{-(n-1-l)\tau} p_{A|J_1^l} \times \left[1 - p_1^{m-M} F_M(\beta_e)\right]. \quad (41)$$

The second term in (5) can be determined by taking its expectation in terms of J_1 and J_2 as

$$P\left(\left\{\gamma_{sp}^{\bar{A}} \geq \beta_e \text{ or } \bigcup_{i=M+1}^m \{\gamma_{S,E_i} + \gamma_{R_b,E_i} \geq \beta_e\}\right\}, \bar{A}\right)$$
$$= \sum_{l=0}^{n-1}\sum_{t=0}^{n-1} P(\bar{A}|J_1^l) P(J_1^l) P(J_2^t) \Big[1 - P\left(\gamma_{sp}^{\bar{A}} < \beta_e \,|J_1^l, J_2^t\right) \times P\left(\gamma_{S,E_i} + \gamma_{R_b,E_i} < \beta_e | J_1^l, J_2^t\right)^{m-M}\Big]$$
$$= \sum_{l=0}^{n-1}\sum_{t=0}^{n-1}\binom{n-1}{l}\binom{n-1}{t}(1-e^{-\tau})^{l+t} e^{-(2n-2-l-t)\tau} \times (1-p_{A|J_1^l})\left[1 - p_2^{m-M} F_{2M}(\beta_e)\right]. \quad (42)$$

Finally, summing (41) and (42), the theorem then follows. ∎

Remark 3: To design a secure IoT system, the designers first need to specify a maximum allowable SOP for the system. Based on the results of (7) and (40), the designers can then determine the feasible (n,τ) curve and thus the inherent tradeoff between n and τ to ensure the target SOP, as illustrated in Fig. 7.

C. Possible Extensions of the Results

As a step towards the complete SOP performance study of the wireless IoT data collection, this paper focuses on a simple homogeneous network scenario due to its mathematical tractability. Hence, one essential extension of this work is to study the SOP under the more general heterogeneous setting (e.g., heterogeneous relays, heterogeneous eavesdroppers and non-i.i.d. channels). To this end, although the theoretical framework for SOP analysis in this paper can still be applied, we will face some new challenges. For example, under the heterogeneous setting of relays and non-i.i.d relay channels, the relays' SIR are on longer i.i.d, which makes the transmission suspension probability analysis extremely challenging. Thus, the mathematical tools adopted in Lemma 1 cannot be directly applied and some new techniques (e.g., the Bapat-Beg Theorem [33]) are required to derive the joint distribution of the order statistics of non-i.i.d relays' SIR. Similarly, under the heterogeneous setting of eavesdroppers and non-i.i.d eavesdropper channels, the eavesdroppers' SIR are no longer i.i.d, which makes the cdf analysis of the aggregate SIR of eavesdroppers also a challenging issue. Thus, the mathematical tools adopted in Lemma 2 cannot be directly applied and some new techniques (e.g., the Berry-Esseen Theorem [34]) are required to determine the cdf of the sum of non-i.i.d eavesdroppers' SIR. For the extension to multiple source–destination pairs, the most challenging issue is to address the contentions of relay selection and jammer selection from different pairs, which requires much

more complicated relay and jammer selection schemes than those in this paper.

Another possible extension of this work is to apply our theoretical framework to study the SOP scaling laws for a larger and heterogeneous system. The challenging issue there is to determine the asymptotic distributions of the SIR of relays and eavesdroppers as the number of relays tends to infinity. To address this issue, some new mathematical tools widely used in the scaling law studies [35], [36], like the Extreme Value Theory and commonly-used inequalities (e.g., Jensen's inequality and Chebyshev's inequality) should be adopted.

The third possible extension of this work is the SOP analysis under the scenarios where eavesdroppers may employ more sophisticated attacking techniques (e.g., selective gain combining, equal gain combining and known-plaintext attack [37]). Under such new attacking models, although the theoretical framework for SOP analysis in this paper can still be applied, the analysis of aggregate SIR of colluding eavesdroppers there becomes much more complicated, since it cannot be simply modeled by a sum of i.i.d. random variables as in this paper. Thus, the mathematical tools adopted in Lemma 2 cannot be directly applied and some new techniques (e.g., Order Statistics and Q Function [38]) are required to determine the cdf of the aggregate SIR of colluding eavesdroppers under the new attacking models. Therefore, a new and dedicated work is required for the study of SOP under those new attacking models.

V. Numerical Results and Discussions

Here, we first validate our theoretical analysis for SOP modeling through extensive simulations and then explore how the number of relays n, the noise-generating threshold τ, the SIR thresholds β_e and β for eavesdroppers and legitimate receivers, and the collusion intensity M, affect the secrecy outage performance, as well as the security performance, of the wireless data collection process of the IoT.

A. Model Validation

A simulator was developed in C++ to simulate the $S-D$ transmission under the system model as described in Section II, which is now available at [39]. The total number of $S-D$ transmissions is fixed as 100000 and the SOP is measured as the ratio of the number of transmissions suffering from secrecy outage to the total number of transmissions. To verify our theoretical analysis, we conduct extensive simulations for both the non-colluding case and the colluding case under various settings of n and M. The number of eavesdroppers is set as $m = 10$, the noise-generating threshold is set as $\tau = 0.5$, the transmit power is set as $P = 100$, and the decoding thresholds for eavesdroppers and legitimate receivers are set as $\beta_e = 0.5$ and $\beta = 1.0$, respectively. Similar to [15], this paper does not consider the path-loss issue, so the network can be regarded to have a kind of equidistant topology. Simulations with other settings can be easily conducted by our simulator as well. The simulation results and the related theoretical ones are summarized in Fig. 3.

It can be observed from Fig. 3 that the simulation results match fairly well with the theoretical ones for both the non-colluding case and the colluding case with different collusion intensity M, which implies that our theoretical analysis is effective in modeling the secrecy outage performance of the concerned system. A careful observation in Fig. 3 reveals that the curve $M = 1$ of the colluding case coincides with that of the noncolluding case, which is intuitive and further proves the effectiveness of our theoretical analysis.

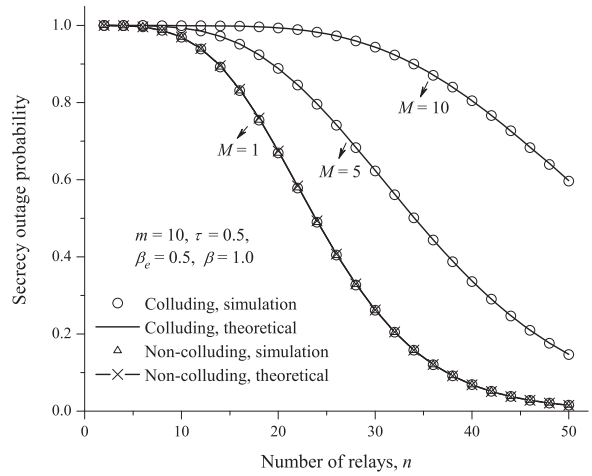

Fig. 3. SOP versus number of relays n for different collusion intensity M, with $m = 10$, $\tau = 0.5, \beta_e = 0.5$ and $\beta = 1.0$.

B. Performance Evaluation

Regarding the impact of the number of relays n on the secrecy outage performance, it can be observed from Fig. 3 that the SOP decreases as n increases for both the noncolluding case and the colluding case with different collusion intensity M. This is mainly due to the reason that, in the cooperative jamming scheme, more interference will be generated at the eavesdroppers for a larger number of relays, and thus the probability that eavesdroppers successfully decode the source message would decrease. This suggests that distributing more relays is an effective approach to decreasing the possibility of secrecy outage, and thus improving the security of the wireless data collection process of the IoT. A careful observation from Fig. 3 indicates that to degrade the SOP to 50%, at least 20 relay nodes are required for $M = 1$ and at least 30 nodes are required for $M = 5$. This is because that the artificial noises generated from the jammers not only degrade the eavesdropper channels but also degrade those of the legitimate transmitter-receiver pairs at the same time. Although the cooperative jamming usually involves a high cost, it is simple and easily deployable, so it still serves as an attractive physical layer security approach to provide security guarantee, in particular for the decentralized wireless networks [8], [9].

To understand the impact of eavesdropper collusion M on the secrecy outage performance, we summarize in Fig. 4 how the SOP varies with M for three different β_e (i.e., $\beta_e = 0.3, \beta_e = 0.5$ and $\beta_e = 1.0$), when $n = 30, m = 10, \tau = 0.5$ and $\beta = 1.0$. We can see from Fig. 4 that as the collusion intensity M increases, so does the SOP, implying that the eavesdropper collusion will significantly increase the possibility of secrecy outage, i.e., deteriorate the security performance of the wireless data collection process of the IoT. For example, the SOP for $\beta_e = 1.0$ when all the eavesdroppers collude (i.e., M

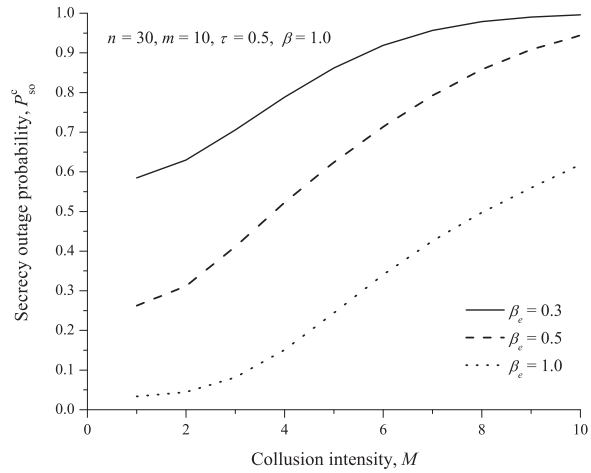

Fig. 4. Secrecy outage probability versus collusion intensity M for different β_e, with $n = 30, m = 10, \tau = 0.5$ and $\beta = 1.0$.

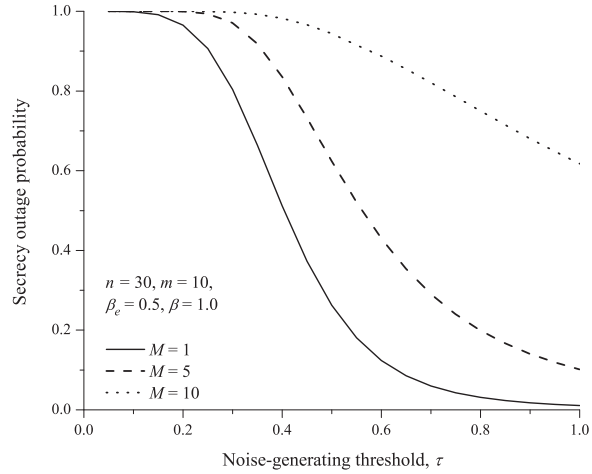

Fig. 5. SOP versus noise-generating threshold τ for different M, with $n = 30, m = 10, \beta_e = 0.5$, and $\beta = 1.0$.

$= m = 10$) is 0.61825, which is much greater than the one 0.03346 when no eavesdroppers collude (i.e., $M = 1$). Another observation from Fig. 4 reveals that the SOP increases as the SIR threshold β_e for eavesdroppers decreases, which is very intuitive since a smaller β_e results in a greater decoding ability for eavesdroppers. As shown in (5), the aggregate SIR is a linear combination of M colluding eavesdroppers' SIR. Thus, an intuitive expectation might be that the SOP is also linear with the collusion intensity M. However, the results in Fig. 4 indicate that this is not the case and the SOP actually increases first superlinearly and then sublinearly with M.

To see how the noise-generating threshold τ affect the secrecy outage performance, we summarize in Fig. 5 how the SOP varies with τ for different collusion intensity M, when $n = 30, m = 10, \beta_e = 0.5$, and $\beta = 1.0$. The results in Fig. 5 indicate that the SOP decreases as the noise-generating threshold τ increases for both the non-colluding $(M = 1)$ and colluding cases $(M > 1)$, which is also because that more interference will be generated at the eavesdroppers for a greater τ. This indicates that increasing the noise-generating threshold is also an effective way to enhance the security performance of the wireless data collection process of the IoT.

To further investigate the impact of the SIR threshold for legitimate receivers β on the secrecy outage performance, we summarize in Fig. 6 how the SOP varies with β for different collusion intensity M, when $n = 30, m = 10, \beta_e = 0.5$, and $\tau = 0.5$. It can be observed from Fig. 6 that as β increases the SOP first remain constant and then decreases. This is mainly due to the reason that there exists some threshold (e.g., about 0.2 in Fig. 6) on β. The transmission is conducted in two hops almost surely for β less than this threshold, whereas the probability that the transmission is suspended in the second hop increases as β increases beyond the threshold. Therefore, the eavesdroppers can overhear the source message in two hops at the beginning but then only in the first hop with an increasing probability as β increases.

To illustrate the inherent tradeoff between the number of relays n and the noise-generating threshold τ, we summarized in Fig. 7 the feasible (n, τ) pairs to achieve a target SOP, under

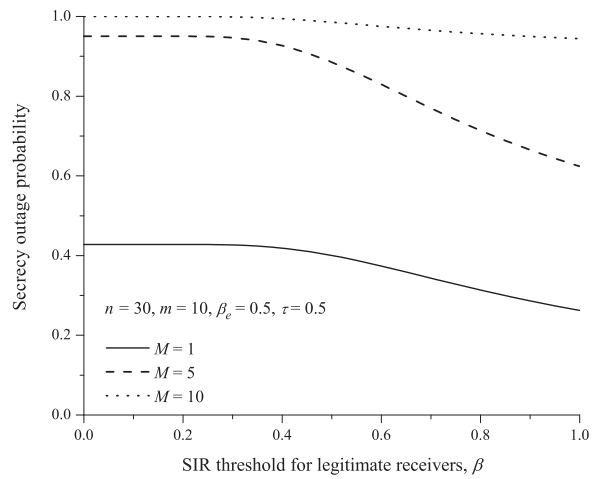

Fig. 6. SOP versus SIR threshold for legitimate receivers β for different M, with $n = 30, m = 10, \beta_e = 0.5$, and $\tau = 0.5$.

the setting of $m = 5, M = 2, \beta_e = 0.5$, and $\beta = 1.0$. It can be observed from Fig. 7 that the noise-generating threshold τ decreases with the number of relays n. This means that if more relays nodes are distributed in the network, a smaller noise-generating threshold is enough to achieve the same target SOP. A careful observation from Fig. 7 indicates that as the number of relays n increases, the SOP is more sensitive to the change of noise-generating threshold τ. For example, to decrease the SOP from 0.5 to 0.2, an increase of τ from 1.2 to 3.2 is required for $n = 10$, whereas a much smaller increase of τ (i.e., the increase from 0.378 to 0.515) is enough for $n = 30$.

To understand the gap between the SOP under the homogeneous setting and that under the heterogeneous settings, we conduct extensive simulations based on a homogeneous scenario of $(n = 30, m = 10, \beta = 1.0, \beta_e = 0.5, \tau = 0.5, P = 100)$ and three heterogeneous scenarios. In the first heterogeneous scenario, the source and each relay node arbitrarily choose a transmit power P in the range of $(0, 200]$. In the second heterogeneous scenario, each eavesdropper arbitrarily adopts a decoding threshold β_e in the range of $(0, 1.0]$. In the third heterogeneous scenario, each relay node arbitrarily adopts a de-

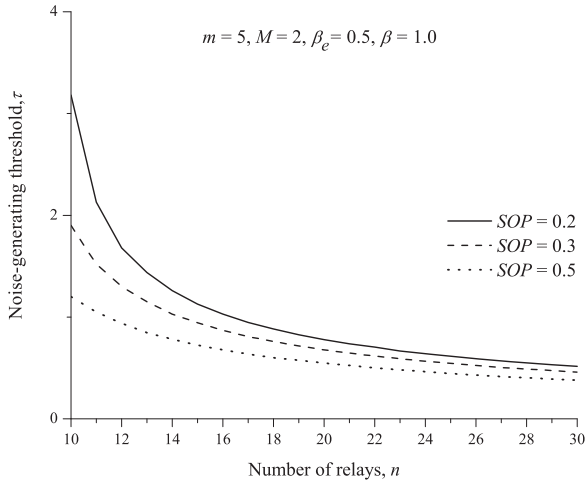

Fig. 7. Feasible (n, τ) curve under the constraint of SOP $= 0.2, 0.3$, and 0.5, for $m = 5, M = 2, \beta_e = 0.5$ and $\beta = 1.0$.

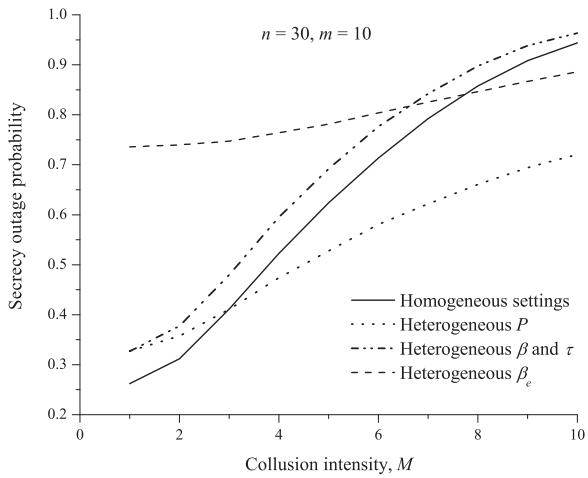

Fig. 8. SOP gap between homogeneous and heterogeneous settings for $n = 30$ and $m = 10$.

coding threshold β in the range of $(0, 2.0]$ and a noise-generating threshold τ in the range of $(0, 1.0]$. The corresponding simulation and theoretical results are summarized in Fig. 8. It can be observed from Fig. 8 that the SOP gaps caused by the heterogeneity of P and β_e are relatively large, while that caused by the heterogeneity of τ and β is much smaller. It can also be observed from Fig. 8 that the behavior of SOP is more sensitive to the heterogeneity of P and β_e, but not sensitive to the heterogeneity of τ and β.

VI. Related Works

The security performance study of wireless communications under eavesdropper collusion and physical layer security can be classified into two categories, depending on the considered network scenario.

For two-hop wireless networks, the secure connection probability, i.e., the probability that the secrecy rates in two hops are both positive, was investigated in [12] to study when a relay is needed to establish a more secure connection. The authors in [13] proposed novel relay strategies to neutralize information leakage from each user to the colluding eavesdroppers by choosing the forwarding matrix of an amplify-and-forward relay in a multi-antenna non-regenerative relay-assisted multi-carrier interference channel. For a multiuser peer-to-peer (MUP2P) relay network with multiple source-destination pairs, multiple relays and a colluding eavesdropper with multiple antennas, the authors in [14] optimized the transmit power of the source and the beamforming weights of the relays jointly to maximize the secrecy rate subject to the minimum signal-to-interference-noise-ratio constraint at each user and the individual and total power constraints. In [15], Vasudevan *et al.* considered a very similar system model with opportunistic relaying and cooperative jamming schemes as in this paper, whereas they focused on the infinite network where the number of relays tends to infinity and explored the number of eavesdroppers (i.e., eavesdropper-tolerance capability) can be present in the network.

For other wireless networks, the scaling law of secrecy capacity was examined for large-scale networks in [16], [17] as network size tends to infinity. The secrecy-constrained connectivity property of large multihop wireless networks with colluding eavesdroppers was considered in [18]. The problem of finding a secure minimum energy routing path of K hops between two nodes in an arbitrary wireless network was considered in [19], subject to constraints on the end-to-end successful eavesdropping probability and throughput over the path. In [20], the tradeoffs between resilience, security, and local-repairability of a distributed storage system in the presence of colluding eavesdroppers were investigated. The security scheme design issue and the related optimization problem under eavesdropper collusion also attracted considerable attention for various network scenarios [21], [22]. The SOP, i.e., the probability that instantaneous secrecy rate between a transmitter-receiver pair is below some threshold, was investigated for various stochastic networks [23], [24].

VII. Conclusion

This paper conducted theoretical analysis to explore the secrecy outage performance of a two-hop wireless communication for collecting data in the IoT under eavesdropper collusion, where physical layer security is adopted to counteract such attack. Two eavesdropper cases were considered, i.e., the non-colluding case where eavesdroppers operate independently and the M-colluding case where any M eavesdroppers combine their observations to conduct eavesdropping attacks. We first derived the SOP of noncolluding case and then determined the SOP for M-colluding case by jointly applying the Laplace and inverse Laplace transform, the keyhole contour integral and the Cauchy Integral Theorem. Our results indicate that eavesdropper collusion can significantly increase the possibility of secrecy outage, and thus, deteriorate the security performance of the data collection in the IoT. Another important finding of this paper is that the cooperative jamming scheme can improve the data collection security by either distributing more relays or increasing the noise-generating threshold. Although the simple cooperative jamming serves as an attractive physical layer security approach to provide security guarantee for the data collection in an IoT system, it usually involves a high cost. So another future work is to combine the cooperative jamming with other

advanced physical layer security methods (e.g., secrecy beamforming/precoding [40]) to reduce the cost for achieving the same SOP guarantee.

APPENDIX A
PROOF OF LEMMA 1

It can be seen from the definition of event A that

$$p_{A|J_1^l} = P\left(\gamma_{S,R_b} < \beta \,|J_1^l\right) = P\left(|h_{S,R_b}|^2 < \beta \sum_{j\in\mathcal{J}_1} |h_{R_j,R_b}|^2 \,|J_1^l\right) \quad (43)$$

Hence, we first need to determine the distribution of $|h_{S,R_b}|^2$. Define $min\{|h_{S,R_k}|^2, |h_{R_k,D}|^2\}$ for each relay $R_k, k = 1,\ldots,n$ by T_k and the event that relay R_k announce itself as the message relay by B_k (i.e., $b = k$). It is easy to see that

$$B_k \triangleq \bigcap_{j=1,j\neq k}^{n} (T_j \le T_k)$$

and all T_k's are i.i.d. and exponential random variables with mean $1/2$. Thus, applying the law of total probability, we obtain

$$\begin{aligned}
&P(|h_{S,R_b}|^2 < x)\\
&= \sum_{k=1}^{n} P\left(|h_{S,R_k}|^2 < x, B_k\right)\\
&= \sum_{k=1}^{n} P\left(|h_{S,R_k}|^2 < x, \bigcap_{j=1,j\neq k}^{n} (T_j \le T_k)\right)\\
&= \sum_{k=1}^{n} \int_0^{\infty} P\left(|h_{S,R_k}|^2 < x, \bigcap_{j=1,j\neq k}^{n} (T_j \le t), T_k = t\right) \mathrm{d}t\\
&= \sum_{k=1}^{n} \int_0^{\infty} P\left(|h_{S,R_k}|^2 < x, T_k = t\right)(1-e^{-2t})^{n-1}\mathrm{d}t\\
&= \int_0^{\infty} nP\left(|h_{S,R_k}|^2 < x, T_k = t\right)(1-e^{-2t})^{n-1}\mathrm{d}t. \quad (44)
\end{aligned}$$

Again, by the law of total probability, we obtain

$$\begin{aligned}
&P\left(|h_{S,R_k}|^2 < x, T_k = t\right)\\
&= \begin{cases} P\left(|h_{S,R_k}|^2 = t, |h_{R_k,D}|^2 > t\right) \\ +P\left(t < |h_{S,R_k}|^2 < x, |h_{R_k,D}|^2 = t\right), & 0 \le t \le x \\ 0, & \text{otherwise} \end{cases}\\
&= \begin{cases} e^{-t}(2e^{-t} - e^{-x}), & 0 \le t \le x \\ 0, & \text{otherwise.} \end{cases} \quad (45)
\end{aligned}$$

Hence

$$\begin{aligned}
&P(|h_{S,R_b}|^2 < x)\\
&= \int_0^x ne^{-t}(2e^{-t} - e^{-x})(1-e^{-2t})^{n-1}\mathrm{d}t\\
&= (1-e^{-2x})^n - ne^{-x}\int_0^x e^{-t}(1-e^{-2t})^{n-1}\mathrm{d}t\\
&= (1-e^{-2x})^n - ne^{-x}\sum_{k=0}^{n-1}\binom{n-1}{k}(-1)^k\int_0^x e^{-(2k+1)t}\mathrm{d}t\\
&= (1-e^{-2x})^n - \sum_{k=0}^{n-1} n\binom{n-1}{k}(-1)^k\frac{e^{-x}-e^{-2(k+1)x}}{2k+1}\\
&= (1-e^{-2x})^n + \sum_{k=0}^{n}\binom{n}{k}(-1)^k\frac{k}{2k-1}(e^{-x}-e^{-2kx})\\
&= \sum_{k=0}^{n}\binom{n}{k}(-1)^k\left[e^{-2kx} + \frac{k}{2k-1}(e^{-x}-e^{-2kx})\right]\\
&= \sum_{k=0}^{n}\binom{n}{k}(-1)^k\frac{ke^{-x} + (k-1)e^{-2kx}}{2k-1} \quad (46)
\end{aligned}$$

Next, the probability distribution of $|h_{R_j,R_b}|^2$ for any $j \in \mathcal{J}_1$ can be given by

$$f_{|h_{R_j,R_b}|^2}(x) = \begin{cases} \frac{e^{-x}}{1-e^{-\tau}}, & 0 \le x < \tau \\ 0, & x \ge \tau \end{cases} \quad (47)$$

Hence,

$$\begin{aligned}
p_{A|J_1^l} =& \mathbb{E}_{\{|h_{R_j,R_b}|^2, j\in\mathcal{J}_1\}}\\
&\left[\sum_{k=0}^{n}\binom{n}{k}(-1)^k\frac{1}{2k-1}\right.\\
&\times\left(ke^{-\beta\sum|h_{R_j,R_b}|^2}\right.\\
&\left.\left.+(k-1)e^{-2k\beta\sum|h_{R_j,R_b}|^2}\right)|J_1^l\right]\\
=& \sum_{k=0}^{n}\binom{n}{k}(-1)^k\frac{1}{2k-1}\left(k\mathbb{E}\left[e^{-\beta\sum|h_{R_j,R_b}|^2}|J_1^l\right]\right.\\
&\left.+(k-1)\mathbb{E}\left[e^{-2k\beta\sum|h_{R_j,R_b}|^2}|J_1^l\right]\right)\\
=& \sum_{k=0}^{n}\binom{n}{k}(-1)^k\frac{1}{2k-1}\left[k\left(\frac{1-e^{-(1+\beta)\tau}}{(1-e^{-\tau})(1+\beta)}\right)^l\right.\\
&\left.+(k-1)\left(\frac{1-e^{-(2k\beta+1)\tau}}{(1-e^{-\tau})(2k\beta+1)}\right)^l\right]. \quad (48)
\end{aligned}$$

REFERENCES

[1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[2] O. Vermesan *et al.*, "The internet of things strategic research roadmap," *Internet of Things: Global Techno. Societal Trends*, vol. 1, pp. 9–52, 2011.

[3] P. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Proc. IEEE FOCS*, 1994, pp. 124–134.

[4] A. Mukherjee, S. Fakoorian, J. Huang, and A. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts*, vol. 16, no. 3, pp. 1550–1573, 2014.

[5] S. Mathur, A. Reznik, C. Ye, R. Mukherjee, A. Rahman, Y. Shah, W. Trappe, and N. Mandayam, "Exploiting the physical layer for enhanced security [security and privacy in emerging wireless networks]," *IEEE Wireless Commun.*, vol. 17, no. 5, pp. 63–70, Oct. 2010.

[6] M. Baldi, M. Bianchi, N. Maturo, and F. Chiaraluce, "A physical layer secured key distribution technique for ieee 802.11g wireless networks," *IEEE Wireless Commun. Lett.*, vol. 2, no. 2, pp. 183–186, Apr. 2013.

[7] H. Taha and E. Alsusa, "A mimo precoding based physical layer security technique for key exchange encryption," in *Proc. IEEE 81st Veh. Technol. Conf.*, May 2015, pp. 1–5.

[8] X. He and A. Yener, "Two-hop secure communication using an untrusted relay," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, pp. 1–13, May 2009.

[9] L. Dong, Z. Han, A. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1875–1888, Mar. 2010.
[10] J. Huang and A. Swindlehurst, "Buffer-aided relaying for two-hop secure communication," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 152–164, Jan. 2015.
[11] M. Yu, M. Zhou, and W. Su, "A secure routing protocol against byzantine attacks for manets in adversarial environments," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 449–460, Jan. 2009.
[12] C. Cai, Y. Cai, X. Zhou, W. Yang, and W. Yang, "When does relay transmission give a more secure connection in wireless ad hoc networks?," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 4, pp. 624–632, Apr. 2014.
[13] Z. Ho, E. Jorswieck, and S. Engelmann, "Information leakage neutralization for the multi-antenna non-regenerative relay-assisted multi-carrier interference channel," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1672–1686, Sep. 2013.
[14] C. Wang, H. Wang, D. Ng, X. Xia, and C. Liu, "Joint beamforming and power allocation for secrecy in peer-to-peer relay networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3280–3293, Jun. 2015.
[15] S. Vasudevan, S. Adams, D. Goeckel, Z. Ding, D. Towsley, and K. Leung, "Multi-user diversity for secrecy in wireless networks," in *Proc. IEEE ITA*, Jan. 2010, pp. 1–9.
[16] J. Zhang, L. Fu, and X. Wang, "Asymptotic analysis on secrecy capacity in large-scale wireless networks," *IEEE/ACM Trans. Netw.*, vol. 22, no. 1, pp. 66–79, Feb. 2014.
[17] M. Mirmohseni and P. Papadimitratos, "Scaling laws for secrecy capacity in cooperative wireless networks," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 1527–1535.
[18] T. Yang, G. Mao, and W. Zhang, "Connectivity of wireless information-theoretic secure networks," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 317–323.
[19] M. Ghaderi, D. Goeckel, A. Orda, and M. Dehghan, "Minimum energy routing and jamming to thwart wireless network eavesdroppers," *IEEE Trans. Mobile Comput.*, vol. 14, no. 7, pp. 1433–1448, Jul. 2015.
[20] A. Rawat, O. Koyluoglu, N. Silberstein, and S. Vishwanath, "Optimal locally repairable and secure codes for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 212–236, Jan. 2014.
[21] M. Khandaker and K.-K. Wong, "Masked beamforming in the presence of energy-harvesting eavesdroppers," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 40–54, 2015.
[22] J. Yang, I.-M. Kim, and D. I. Kim, "Optimal cooperative jamming for multiuser broadcast channel with multiple eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2840–2852, Jun. 2013.
[23] J. Bai, X. Tao, J. Xu, and Q. Cui, "The secrecy outage probability for the ith closest legitimate user in stochastic networks," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1230–1233, July 2014.
[24] G. Geraci, S. Singh, J. Andrews, J. Yuan, and I. Collings, "Secrecy rates in broadcast channels with confidential messages and external eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2931–2943, May 2014.
[25] X. Zhou, M. McKay, B. Maham, and A. Hjrungnes, "Rethinking the secrecy outage formulation: A secure transmission design perspective," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 302–304, Mar. 2011.
[26] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, Aug. 2002.
[27] J. Laneman, D. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
[28] C. S. R. Murthy and B. Manoj, *Ad Hoc Wireless Networks: Architectures and Protocols*. Upper Saddle River, NJ, USA: Prentice-Hall PTR, 2004.
[29] A. Bletsas, S. Khisti, D. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006.
[30] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Parallel Distrib. Syst.*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.
[31] C. M. Ramsay, "The distribution of sums of certain i.i.d. Pareto variates," *Communications in Statistics - Theory and Methods*, vol. 35, no. 3, pp. 395–405, 2006.
[32] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, ser. Appl. Math.. Washington, DC, USA: U.S. Dept. Commerce, 1972.
[33] R. B. Bapat and M. I. Beg, "Order statistics for nonidentically distributed variables and permanents," *Sankhy: The Indian J. Statistics, Series A*, vol. 51, no. 1, pp. 79–93, 1989.
[34] N. Chaidee and M. Tuntapthai, "Berry-esseen bounds for random sums of non-iid random variables," *Int. Math. Forum*, vol. 4, no. 26, pp. 1281–1288, 2009.
[35] D. Goeckel, S. Vasudevan, D. Towsley, S. Adams, Z. Ding, and K. Leung, "Artificial noise generation from cooperative relays for everlasting secrecy in two-hop wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 10, pp. 2067–2076, Oct. 2011.
[36] O. Koyluoglu, C. Koksal, and H. El Mamal, "On secrecy capacity scaling in wireless networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3000–3015, May 2012.
[37] M. Schulz, A. Loch, and M. Hollick, "Practical known-plaintext attacks against physical layer security in wireless MIMO systems," in *Proc. NDSS*, 2014, pp. 1–13.
[38] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge Univ., 2005.
[39] C++ Simulator for Two-Hop Transmission With Colluding Eavesdroppers [Online]. Available: http://mdlval.blogspot.jp/
[40] Y.-W. Hong, P.-C. Lan, and C.-C. Kuo, "Enhancing physical-layer secrecy in multiantenna wireless systems: An overview of signal processing approaches," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 29–40, Sep. 2013.

Yuanyu Zhang received the B.S. degree in software engineering and M.S. degree in computer science from Xidian University, Xi'an, China, in 2011 and 2014, respectively. He is currently working toward the Ph.D. degree at the School of Systems Information Science, Future University Hakodate, Hokkaido, Japan.

His research interests include the physical layer security of wireless communications, and performance modeling and evaluation of wireless networks.

Yulong Shen received the B.S. and M.S. degrees in computer science and Ph.D. degree in cryptography from Xidian University, Xi'an, China, in 2002, 2005, and 2008, respectively.

He is currently a Professor with the School of Computer Science and Technology, Xidian University, China. He is also an Associate Director with the Shaanxi Key Laboratory of Network and System Security and a member of the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China. He has also served on the technical program committees of several international conferences, including ICEBE, INCoS, CIS and SOWN. His research interests include wireless network security and cloud computing security.

Hua Wang received the Ph.D. degree from the University of Southern Queensland, Australia.

He is a full-time Professor with Victoria University, Victoria, Australia. He was a Professor with the University of Southern Queensland before he joined Victoria University. He has more than ten years of teaching and working experience in applied informatics in both industry and academia. He has expertise in electronic commerce, business process modeling and enterprise architecture. As an Chief Investigator, three Australian Research Council (ARC) Discovery grants have been awarded since 2006, and 155 peer-reviewed scholar papers have been published. Six Ph.D. students have already graduated under his principal supervision.

Jianming Yong (M'14) received the Ph.D. degree from Swinburne University of Technology, Victoria, Australia.

He has over 30 years of experience in both IT industry and tertiary education. He is an Associate Professor with the Discipline of Information Systems, School of Management and Enterprise, Faculty of Business, Education, Law and Arts, University of Southern Queensland, Australia. He has been involved in many international competitive grants such as European Research grants, Australia- China grant. His research areas include e-learning, cloud computing, information security and privacy.

Prof. Yong is a member of the IEEE Computer Society.

Xiaohong Jiang (SM'09) received the B.S., M.S., and Ph.D. degrees from Xidian University, Xi'an, China, in 1989, 1992, and 1999 respectively.

He is currently a Full Professor with Future University Hakodate, Hokkaido, Japan. Before joining Future University, he was an Associate Professor with Tohoku University from February .2005 to March 2010. His research interests include computer communications networks, mainly wireless networks and optical networks, network security, and routers/switches design. He has published over 260 technical papers at premium international journals and conferences.

Dr. Jiang is a member of ACM and IEICE. He was the recipient of the Best Paper Award of IEEE HPCC 2014, IEEE WCNC 2012, IEEE WCNC 2008, IEEE ICC 2005-Optical Networking Symposium, and IEEE/IEICE HPSR 2002.