Information Sciences 281 (2014) 201–210

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Transmission protocol for secure big data in two-hop wireless networks with cooperative jamming

Yulong Shen, Yuanyu Zhang *

School of Computer Science and Technology, Xidian University, 710071 Xi'an, China

ARTICLE INFO

Article history:
Received 17 January 2014
Received in revised form 11 May 2014
Accepted 20 May 2014
Available online 2 June 2014

Keywords:
Big data
Security
Wireless network
Cooperative jamming
Energy balance

ABSTRACT

Wireless communications nowadays are increasingly becoming irreplaceable networking technologies for military, governmental and financial communications where the data is exploding in volume, variety and velocity, called big data. This poses great challenges to ensuring security through cryptography. Recently, cooperative jamming has been proved as a promising physical layer technique to provide the everlasting security for wireless networks. Based on this scheme, this paper proposes a two-hop transmission protocol with parameters l, k, r and τ (2HR-(l, k, r, τ)) to ensure secure and reliable big data transmissions in wireless networks with multiple eavesdroppers. We first determine the relay selection region (RSR) as the square of side-length l centered at the middle point between the source and the destination. Then one of the k best relays located in the RSR is randomly selected as the message relay. During the forwarding in both hops, the remaining relays at least distance r away from the intended receivers and with channel gain to the intended receivers less than τ are selected to generate jamming signals to confuse the eavesdroppers. The results in this paper indicate that our protocol can provide flexible control of security, reliability and the energy balance performance, which characterizes how energy consumption for forwarding message is balanced among all the relays.

1. Introduction

Recently, wireless networks have attracted increasing attentions in both academia and industry. Owing to the advantages of flexible and self-configuring, wireless networks are anticipated to play an important role in some mission-critical applications, like health care, battle command, and governmental and financial communication. Therefore, such networks are now becoming essential for ensuring the economic vitality, people safety and even national security. The rapid growth of the node complexity and communication technologies allow the organizations in above scenarios generate and aggregate more and more data, which is now termed as big data [14,12,1]. As we know, much of these data is very sensitive and needs to be kept away from intercepting and damaging of the eavesdroppers during the transfer and aggregation. The nature of big data – high volume, variety and velocity also makes it difficult to ensure data integrity. Thus, security is becoming one of the greatest concerns in the big data environment [21]. Since data size is growing exponentially, it is essential to use efficient data transfer protocols to move vast amounts of data. Tow-hop forwarding, where each packet travels at most two hops (source-relay-destination) to reach its destination, is increasingly becoming important and promising in wireless

* Corresponding author. Tel.: +81 080 3266 6646.
E-mail addresses: ylshen@mail.xidian.edu.cn (Y. Shen), yy90zhang@gmail.com (Y. Zhang).

http://dx.doi.org/10.1016/j.ins.2014.05.037
0020-0255/

communication [19] and will play an important role in big data environment. In this paper, we hence focus on the security issue of big data transmission in two-hop wireless networks.

The traditional method to provide a standard information is the cryptographic approach where a complex algorithm is developed such that any adversary with limited computing power and without the secret key cannot intercept the information. However, the everlasting secrecy cannot be achieved by such approach, since the adversary can record the transmitted messages and try any way to break them [25]. Especially, recent advances in high-performance computation (e.g. quantum computing) make it further unlikely to acquire long-lasting security via cryptographic approaches [11]. Furthermore, the need for fast data transfer in big data environment makes it very difficult and expensive to exchange and maintain secret keys among various organizations. This motivates the consideration of signaling scheme in physical layer secrecy framework to provide a strong form of security for big data, where a degraded signal at an eavesdropper is always ensured such that the original data can be hardly recovered regardless of how the signal is processed at the eavesdropper [28,27,13]. Moreover, physical layer security approaches can be used with the cryptographic method in a complementary way to augment the security achieved by cryptography.

In physical layer secrecy framework, the key to achieve the everlasting security is guaranteeing an advantage of the legitimate channel over the eavesdropper channel, which, however, cannot always be guaranteed in general. This motivates many ideas to create a better legitimate channel, like the public and error-free feedback channel [18], the space–time coding over multiple antennas for secure communication [10] and the artificial noise injection strategy in MIMO [20,8]. However, due to the cost of deploying multiple antennas and designing efficient noise, these schemes are not suitable for large scale wireless network with nodes of single antenna. That is why there is an increasing interest in exploiting cooperative jamming in large scale wireless networks to provide strong form of security [7,17,24,23,9]. The basic idea behind cooperative jamming is that the jamming signals (or artificial noises) are generated from cooperative nodes to provide secure communication for the network. In this paper, we aim to design an efficient transmission protocol to ensure secure and reliable big data transmission based on the cooperative jamming scheme.

Recently, various works have been dedicated to design the secure transmission scheme via cooperative relays in large scale wireless networks under the framework of physical layer secrecy. In the case where eavesdropper channels or locations are known, node cooperation is used to improve the performance of secure wireless communications and a few cooperative transmission protocols were proposed to jam eavesdroppers [5,6]. In the case where eavesdropper channels or locations are unknown, Goeckel et al. proposed a transmission protocol based on optimal relay selection [7]. For both one-dimensional and two-dimensional networks, a secure transmission protocol is proposed in [3]. Ding et al. considered the opportunistic use of relays and proposed two secrecy transmission protocols [4]. The two-way secrecy scheme was studied in [15,2]. Sheikholeslami et al. proposed a protocol, where the signal of a given transmitter is protected by the aggregate interference produced by the other transmitters [22]. A secure transmission protocol is presented in the case where the eavesdroppers can collude [26]. Li et al. proposed two secure transmission protocols to confound the eavesdroppers [16]. The above works mainly focus on the maximum the secrecy capacity, in which the system nodes with best link condition is always selected as information relay. Although these protocols are attractive to provides good security, they do not consider the energy balance which can measure how energy consumption to forward message from the source to the destination is balanced among all the cooperative relays. Since an imbalanced use of the nodes may cause some nodes die much earlier and thus create holes in the network, or what is worse, leave the network disconnected, providing a good energy balance performance is extremely important for energy-limited wireless networks, such as military or emergency networks. In this paper, we proposed a transmission protocol 2HR-(l,k,r,τ) in two-hop relay wireless networks with parameters to ensure secure and reliable big data communication. We first identify the RSR which is a square of side-length l centered at the middle point between the source and destination. Then one of the best k relays in the RSR is randomly selected to forward message. During the transmission, relays at least r away from intended receivers and with channel gain to the intended receivers less than *tau* are selected to transmit jamming signals to provide physical layer security. Theoretical analysis is presented to explore the behavior of the energy balance. Extensive simulation results are also provided to evaluate the performances of the 2HR-(l,k,r,τ) protocol. The results in this paper indicate that our protocol can provide flexible control of security and reliability through proper settings of these four parameters and the energy balance can be controlled by proper setting of l.

The remainder of this paper is organized as follows. Section 1 introduces the system models and presents the 2HR-(l,k,r,τ) protocol. Section 3 introduces the metrics of security and reliability and presents the theoretical analysis on the energy balance. Numerical results and discussions are provided in Sections 4 and 5 concludes this paper.

2. System models and 2HR-$(\boldsymbol{l},\boldsymbol{k},\boldsymbol{r},\tau)$ protocol

2.1. Network model

A two-hop wireless network scenario is considered where a source node S wishes to communicate securely with its destination node D with the help of multiple half-duplex relay nodes $R_1, R_2, \ldots, R_n$. Also present in the environment are m eavesdroppers $E_1, E_2, \ldots, E_m$ without knowledge of channels and locations. We assume that the eavesdroppers do not transmit in order to hide themselves and each of them attempts to decode the message based on its own observations. We consider a time-slotted system where the relay nodes and eavesdroppers choose a position independently and uniformly

in each slot and stay static for the whole slot. A snapshot of the network is illustrated in Fig. 1. Our goal here is to design a protocol to ensure the secure and reliable big data transmission from source S to destination D and provide flexible energy balance control in two-hop wireless networks.

2.2. Transmission model

Consider the transmission from a transmitter A to a receiver B, and denote the ith symbol transmitted by node A by $x_i^{(A)}$. We assume that all nodes transmit with the same power E_s and path-loss between all pairs of nodes is independent. We denote the frequency-nonselective Rayleigh fading from A to B by $h_{A,B}$. Under the condition that nodes with indices in a set $\mathcal{R}$ are generating noises, the ith signal received at node B from node A, denoted by $y_i^{(B)}$, is determined as:

$$y_i^{(B)} = \frac{h_{A,B}}{d_{A,B}^{\alpha/2}}\sqrt{E_s}x_i^{(A)} + \sum_{j\in\mathcal{R}}\frac{h_{A_j,B}}{d_{A_j,B}^{\alpha/2}}\sqrt{E_s}x_i^{(A_j)} + n_i^{(B)}$$

where $d_{A,B}$ is the distance between node A and B, $\alpha \geqslant 2$ is the path-loss exponent, $|h_{A,B}|^2$ is exponentially distributed and without loss of generality, we assume that $E[|h_{A,B}|^2] = E[|h_{B,A}|^2] = 1$. The noise $n_i^{(B)}$ at receiver B is assumed to be i.i.d complex Gaussian random variables with mean N_0. Thus, the SINR (Signal to Interference plus Noise Ratio) $SINR_{A,B}$ from A to B is then given by

$$SINR_{A,B} = \frac{E_s|h_{A,B}|^2 d_{A,B}^{-\alpha}}{\sum_{j\in\mathcal{R}}E_s|h_{A_j,B}|^2 d_{A_j,B}^{-\alpha} + N_0/2}$$

For a legitimate node and an eavesdropper, we use two separate SINR thresholds γ_R and γ_E to define the minimum SINR required to recover the transmitted messages for legitimate node and eavesdropper, respectively. Therefore, a system node (the selected relay or destination) is able to decode a packet if and only if its received SINR is greater than γ_R, whereas each eavesdropper try to achieve target SINR γ_E to successfully recover the transmitted message. In order to simplify the analysis of energy consumption balance for the relay selection, we denote the energy consumption for forwarding message in one transmission by I_0.

2.3. 2HR-(l, k, r, τ) protocol

In this subsection, we introduces the transmission protocol 2HR-(l, k, r, τ) which can provide control of energy balance. To describe and evaluate our 2HR-(l, k, r, τ) protocol, we map the network in Fig. 1 to a unit square with coordinate system $[-0.5, 0.5] \times [-0.5, 0.5]$ in Fig. 2, where S is located at $(-0.5, 0)$ and D is located at $(0.5, 0)$. The protocol works as follows.

1. ***Relay selection region determination***: The square of side-length l centered at the origin is determined as the relay selection region (RSR).
2. ***Channel measurement***: The source S and destination D broadcast a pilot signal to allow each relay to measure the channel from S and D to itself. The relays, which receive the pilot signal, can accurately extract the information of $(h_{S,R_j}, d_{S,R_j}),\ j = 1, 2, \ldots, n$ and $(h_{R_j,D}, d_{R_j,D}),\ j = 1, 2, \ldots, n$.
3. ***Candidate relay selection***: For each relay, we use

$$h_j = min\left(\frac{|h_{S,R_j}|^2}{d_{S,R_j}^{\alpha}}, \frac{|h_{R_j,D}|^2}{d_{R_j,D}^{\alpha}}\right)$$

to represent its object function used in relay selection. The relays with the first k largest h_j in the RSR form the relay selection group $\mathfrak{R}$. Notice that if there are less than or equal to k relays in the RSR, $\mathfrak{R}$ is the set of all nodes in the RSR.
4. ***Relay selection***: The relay, indexed by j^*, is randomly selected from the relay selection group $\mathfrak{R}$. Analogous to Step 2, each of the other relays $R_j,\ j = 1, 2, \ldots, n,\ j \neq j^*$ in network exactly knows $\left(h_{R_j,R_{j^*}}, d_{R_j,R_{j^*}}\right)$.

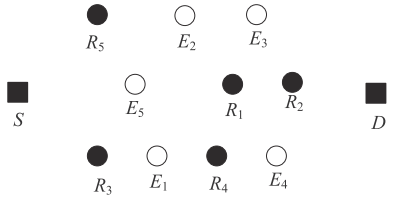

Fig. 1. System scenario: Source S wishes to communicate securely with destination D with the assistance of finite relays $R_1, R_2, \ldots, R_n$ ($n = 5$ in the figure) in the presence of passive eavesdroppers $E_1, E_2, \ldots, E_m$ ($m = 5$ in the figure). Cooperative relay scheme is used in the two-hop transmission.

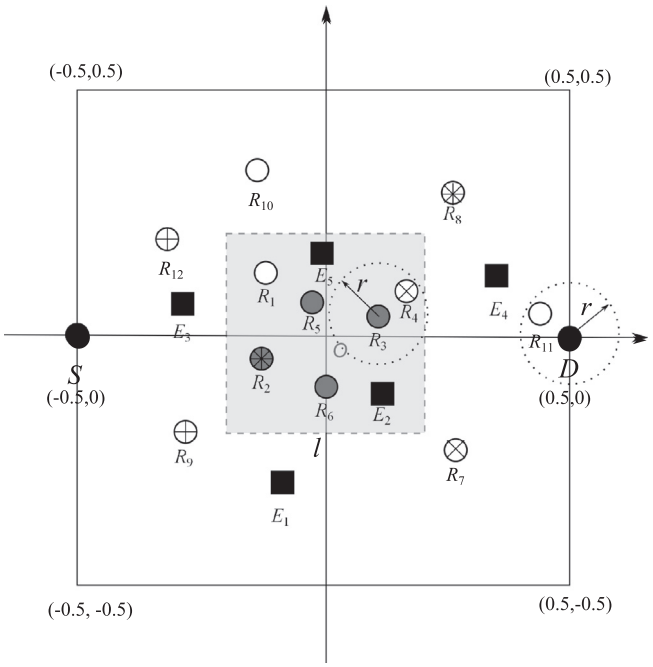

Fig. 2. Coordinate system for the scenario.

5. ***Transmission from*** S ***to*** R_{j^*}: In this step, the source S transmits the message to R_{j^*}. In order to reduce the interference at the intended receiver and to confuse the eavesdroppers, the relay in $\mathcal{R}_1 = \left\{ j | \frac{|h_{R_j,R_{j^*}}|^2}{d^2_{R_j,R_{j^*}}} < \tau,\ d_{R_j,R_{j^*}} < r,\ j \neq j^* \right\}$ transmit jamming signals to confuse the eavesdroppers.
6. ***Transmission from*** R_{j^*} ***to*** D: Similarly, the relay R_{j^*} in this step then forward the source message to the destination D. At the same time, the relay nodes whose indices are in $\mathcal{R}_2 = \left\{ j | \frac{|h_{R_j,D}|^2}{d^2_{R_j,D}} < \tau, d_{R_j,D} < r, j \neq j^* \right\}$ transmit jamming signals to assist the legitimate transmission.

We assume that only one end-to-end transmission can be conducted in one time slot. It is notable that in our protocol the energy balance can be flexibly controlled by a proper setting of relay selection region side-length l and the outage probabilities can be controlled by $l,\ k,\ r$ and τ. The parameter r here indicates that the jammers must be at least distance r away from the intended receivers and thus the interferences at the intended receivers are controlled.

To understand the protocol, we consider an example with $l = 0.4,\ k = 4$ and $r = 0.1$ in Fig. 2. First, the relay selection region is determined where 6 relays $R_1,\ R_2,\ R_3,\ R_4,\ R_5,\ R_6$ are located. The relays with first 4 largest h_j form the relay selection group $\Re$ which consists of $R_2,\ R_3,\ R_5,\ R_6$. Then one of the relays in $\Re$ (i.e., R_3 in Fig. 2) is randomly selected to forward message for S. For the first hop transmission, relay $R_2,\ R_8,\ R_9$ and R_{12} are selected as the jammers according to τ. Similarly, relay $R_2,\ R_4,\ R_7$ and R_8 are selected as jammers in the second hop.

3. Theoretical analysis on 2HR-(l, k, r, τ) protocol

In this section, we first introduce the metrics (secrecy outage probability and transmission outage probability) used in this paper to evaluate our protocol and then provide some theoretical analysis to determine the energy balance of the protocol.

3.1. Outage probabilities

We first introduce the concepts of transmission outage and secrecy outage of the transmission from S to D and then define the transmission outage probability and secrecy outage probability.

For a transmission from the source S to the destination D, we say transmission outage happens if.

1. there are no relays located in the RSR which implies that the transmission cannot be conducted;
2. under the condition that the transmission is conducted, either the message relay R_{j^*} or the destination D cannot decode the transmitted packet. That is, $SINR_{S,R_{j^*}} < \gamma$ or $SINR_{R_{j^*},D} < \gamma$.

We say secrecy outage happen if under the condition that the transmission is conducted, at least one of the eavesdroppers, say E_i, can decode the transmitted message in either hop (that is $SINR_{S,E_i} \geqslant \gamma_E$ or $SINR_{R_{j^*},E_i} \geqslant \gamma_E$). Thus, we can give the following definitions:

- **Transmission outage probability** P_{to}: the probability that the transmission outage happens for a transmission from S to D under the 2HR-(l,k,r,τ) protocol.
- **Secrecy outage probability** P_{so}: the probability that the secrecy outage happens for a transmission from S to D under the 2HR-(l,k,r,τ) protocol.

Based on the above definitions, P_{to} and P_{so} can be formulated as

$$P_{to} = (1-l^2)^n + (1-(1-l^2)^n)\,P(SINR_{S,R_{j^*}} < \gamma \cup SINR_{R_{j^*},D} < \gamma) = (1-l^2)^n + (1-(1-l^2)^n)[1 - P(SINR_{S,R_{j^*}} \geqslant \gamma, SINR_{R_{j^*},D} \geqslant \gamma)] = (1-l^2)^n + (1-(1-l^2)^n)\left[1 - P\left(\frac{|h_{S,R_{j^*}}|^2 d_{S,R_{j^*}}^{-\alpha}}{\sum_{j\in\mathcal{R}_1}|h_{R_j,R_{j^*}}|^2 d_{R_j,R_{j^*}}^{-\alpha} + N_0/2E_s} \geqslant \gamma, \frac{|h_{R_{j^*},D}|^2 d_{R_{j^*},D}^{-\alpha}}{\sum_{j\in\mathcal{R}_2}|h_{R_j,D}|^2 d_{R_j,D}^{-\alpha} + N_0/2E_s} \geqslant \gamma\right)\right]$$

and

$$P_{so} = (1-(1-l^2)^n)P\left(\bigcup_{i=1}^{m}\{SINR_{S,E_i} \geqslant \gamma_E \cup SINR_{S,E_i} \geqslant \gamma_E\}\right) = (1-(1-l^2)^n)\left[1 - P\left(\bigcap_{i=1}^{m}\{SINR_{S,E_i} < \gamma_E, SINR_{S,E_i} < \gamma_E\}\right)\right]$$
$$= (1-(1-l^2)^n)\left[1 - P\left(\bigcap_{i=1}^{m}\left\{\frac{|h_{S,E_i}|^2 d_{S,E_i}^{-\alpha}}{\sum_{j\in\mathcal{R}_1}|h_{R_j,E_i}|^2 d_{R_j,E_i}^{-\alpha} + N_0/2E_s} < \gamma_E, \frac{|h_{R_{j^*},E_i}|^2 d_{R_{j^*},E_i}^{-\alpha}}{\sum_{j\in\mathcal{R}_2}|h_{R_j,E_i}|^2 d_{R_j,E_i}^{-\alpha} + N_0/2E_s} < \gamma_E\right\}\right)\right]$$

In this paper, we use transmission outage probability and secrecy outage probability as two main metrics to evaluate the performance of our protocol.

3.2. Energy balance analysis

The energy balance characterizes how the energy used for forwarding message is balanced among n relays. In this section, we provide some theoretical analysis to explore the behavior of the energy balance of the protocol.

We define the energy consumption of relay $R_j,\ j = 1,2,\ldots,n$ in the ith transmission by X_j^i and its total energy consumption in N transmissions by C_j. Thus, we have $C_j = \sum_{t=1}^{N} X_j^i$. If all the C_js are identical, we can say that the energy is best balanced. We define the energy balance of our protocol by L hereafter. However, if $max_{j=1}^n\{C_j\}$ is quite larger than $min_{j=1}^n\{C_j\}$, the energy balance performance is very poor. Therefore, we can use the ratio of the energy consumption range $max_{j=1}^n\{C_j\} - min_{j=1}^n\{C_j\}$ to the total energy consumption $\sum_{j=1}^n C_j$ to represent the energy balance. That is,

$$L = \frac{max_{j=1}^n\{C_j\} - min_{j=1}^n\{C_j\}}{\sum_{j=1}^n C_j}.$$

Since the variance of C_j is a quick way to estimate the energy consumption range and the total energy consumption can be measured by $E[C_j]$, we adopt the ratio $\frac{Var[C_j]}{E[C_j]}$ to explore the behavior of the energy balance. That is,

$$L \propto \frac{Var[C_j]}{E[C_j]}$$

Theorem 1. *Considering the network scenario in Fig. 2, the energy balance of 2HR-(l,k,r,τ) Protocol is*

$$L \propto \frac{I_0(n-1+(1-l^2)^n)}{n}$$

Proof. To derive the variance of C_j, we need to derive the variance of X_j^i and thus the probability distribution of X_j^i. First, we define the event that relay $R_j,\ j = 1,2,\ldots,n$ is selected as the message relay in one end-to-end transmission by A_j. Besides, the event that there are t relays except R_j in the relay selection region is represented by B_t^j. Next, we define the probability density function of h_j by $f(h)$, the cumulative distribution function by $F(h)$. Thus, we have,

$$P(A_j) = l^2\sum_{t=0}^{n-1} P\left(A_j|B_t^j\right)P\left(B_t^j\right) = l^2\left[\sum_{t=0}^{k-1}\frac{P\left(B_t^j\right)}{t+1} + \sum_{t=k}^{n-1} P\left(A_j|B_t^j\right)P\left(B_t^j\right)\right]$$
$$= l^2\left[\sum_{t=0}^{k-1}\binom{n-1}{t}(l^2)^t(1-l^2)^{n-1-t}\frac{1}{t+1} + \sum_{t=k}^{n-1}\binom{n-1}{t}(l^2)^t(1-l^2)^{n-1-t}P\left(A_j|B_t^j\right)\right] \quad (1)$$

Next,

$$\begin{aligned}
P\left(A_j|B_t^j\right) &= \frac{1}{k}\int_0^\infty P\left(A_j|B_t^j, h_j = h\right) f(h)dh \\
&= \frac{1}{k}\int_0^\infty \sum_{i=0}^{k-1}\binom{t}{i}(1-F(h))^i F(h)^{t-i} f(h)dh \\
&= \frac{1}{k}\sum_{i=0}^{k-1}\binom{t}{i}\int_0^\infty (1-F(h))^i F(h)^{t-i} dF(h) \\
&\overset{F(h)=u}{=} \frac{1}{k}\sum_{i=0}^{k-1}\binom{t}{i}\int_{F(0)}^{F(\infty)} (1-u)^i u^{t-i} du \\
&= \frac{1}{k}\sum_{i=0}^{k-1}\binom{t}{i}\int_0^1 (1-u)^i u^{t-i} du \\
&= \frac{1}{k}\sum_{i=0}^{k-1}\binom{t}{i} B(t-i+1, i+1) \\
&= \frac{1}{k}\sum_{i=0}^{k-1}\binom{t}{i}\frac{(t-i!)i!}{(t+1)!} \\
&= \frac{1}{t+1}
\end{aligned} \tag{2}$$

where $B(t-i+1, i+1)$ is the beta function with parameters $t-i+1$ and $i+1$. Substituting (2) into (1), we have

$$P(A_j) = l^2\sum_{t=0}^{n-1}\binom{n-1}{t}(l^2)^t(1-l^2)^{n-1-t}\frac{1}{t+1} = \sum_{i=1}^{n}\binom{n-1}{i-1}(l^2)^i(1-l^2)^{n-i}\frac{1}{i} = \frac{1}{n}\sum_{i=1}^{n}\binom{n}{i}(l^2)^i(1-l^2)^{n-i} = \frac{1-(1-l^2)^n}{n} \tag{3}$$

Therefore,

$$P\left(X_j^i = I_0\right) = \frac{1-(1-l^2)^n}{n}$$

and

$$P\left(X_j^i = 0\right) = 1 - \frac{1-(1-l^2)^n}{n}$$

The mean and variance of X_j^i can be given by

$$E\left[X_j^i\right] = \frac{I_0(1-(1-l^2)^n)}{n}$$

and

$$Var\left[X_j^i\right] = \frac{I_0^2(1-(1-l^2)^n)(n-1+(1-l^2)^n)}{n^2}$$

Since $X_j^i,\ i = 1, 2, \ldots, N$ are independent and identically distributed, we have

$$E[C_j] = \frac{NI_0(1-(1-l^2)^n)}{n}$$

$$Var[C_j] = \frac{NI_0^2(1-(1-l^2)^n)(n-1+(1-l^2)^n)}{n^2}$$

Therefore,

$$L \propto \frac{I_0(n-1+(1-l^2)^n)}{n} \qquad \square$$

4. Numerical results and discussions

In order to evaluate the performances of the 2HR-(l,k,r,τ) protocol, we provide extensive numerical results to present the behaviors of the secrecy outage probability, transmission outage probability and energy balance as the system parameters vary.

4.1. Simulation settings

We developed a simulator in C++ to simulate the end-to-end message transmission based on the 2HR-(l,k,r,τ) protocol in this paper, which is now available at [29]. We consider a network scenario where there are $n = 500$ relays, $m = 5$ eavesdroppers, the transmit power is 10 and the noise is fixed as 1.0, the SINR threshold for legitimate receivers is fixed as $\gamma_R = 0.05$ and the SINR threshold for eavesdroppers is fixed as $\gamma_E = 0.04$. In the simulation of transmission outage probability and secrecy outage probability, we conducted 100,000 end-to-end transmissions and the secrecy (transmission) outage probability is calculated as the ratio of the number of transmissions with secrecy (transmission) outage to the number of total transmissions 100,000. In the simulation of energy balance, we conducted 10,000 end-to-end transmissions for 100 times and got a sample result each time. The energy balance is calculated by averaging all the sample results.

4.2. Outage probabilities

To explore the impact of jamming threshold τ and distance constraint from the jammers to the intended receivers r on the transmission outage probability P_{to} and secrecy outage probability P_{so}, we show in Fig. 3 how P_{to} and P_{so} vary with τ and r. For both cures, we consider a network scenario of $n = 500,\ m = 5,\ k = 1,\ l = 0.3,\ \gamma_R = 0.05$ and $\gamma_E = 0.04$. The setting that γ_R is greater than γ_E implies that the eavesdroppers have a better decoding ability than the intended receivers. The solid line shows the behavior of P_{to} and P_{so} as τ increases from 0.0 to 8.0 with $r = 0.1$ and the dashed line represents the tendency of P_{to} and P_{so} as r decreases from 1.0 to 0.0 with $\tau = 2.0$. It can be observed from the solid line that the transmission outage probability P_{to} increases as τ increases, while secrecy outage probability P_{so} decreases as τ increases. This is because that increasing the jamming threshold for given n and r would increases the number of jammers and thus more interference can be generated at the receivers. Thus, it is more difficult for them to decode the message. We can see from the dashed line that the transmission outage probability P_{to} increases as r decreases, while the secrecy outage probability decreases as r decreases. This is due to the reason that more relays can be selected as jammers if a less stringent constraint r on the distance from relays to the intended receivers is adopted and thus more interference can be generated at both the intended receivers and eavesdroppers.

To find out how the relay selection group k and side-length of RSR l affect the transmission outage probability and secrecy outage probability, we summarize the simulation results in Figs. 4 and 5 to show the behavior of P_{to} and P_{so} with k and l respectively. It can be observed from Fig. 4 that the transmission outage probability P_{to} increases as k increases for a specific l. This is because that the quality of the channel $S - R_{j^*}$ and channel $R_{j^*} - D$ becomes worse as the relay selection group k increases. A careful observation from Fig. 4 indicates that for any $l < 1$ the transmission outage probability will stay constant when k is larger than some value. This is because that when k is large enough, the probability that there are less than k relays in the RSR approaches 1 and the random relay selection has no impact on the channel condition. We can also observe from Fig. 4 that for a specific k, the transmission outage probability P_{to} first decreases, then increases and at last decreases with l. This interesting phenomenon is due to the reason that at the beginning the probability that there are no relays in the RSR dominates the P_{to} and will decrease as l increases. Then as l increases, the probability there are no relays in the RSR is

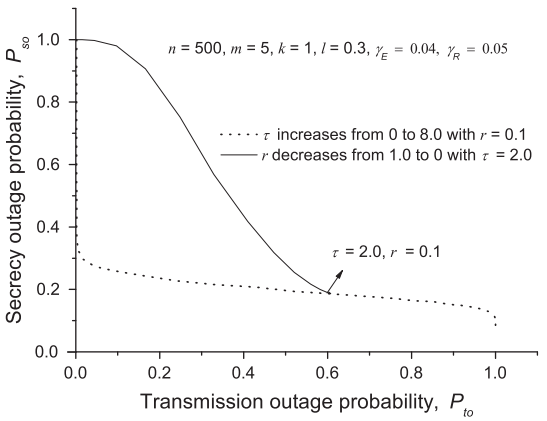

Fig. 3. Secrecy outage probability and transmission outage probability vary with jamming threshold τ and distance constraint r, when $n = 500,\ m = 5,\ k = 1,\ l = 0.3,\ \gamma_R = 0.05$ and $\gamma_E = 0.04$.

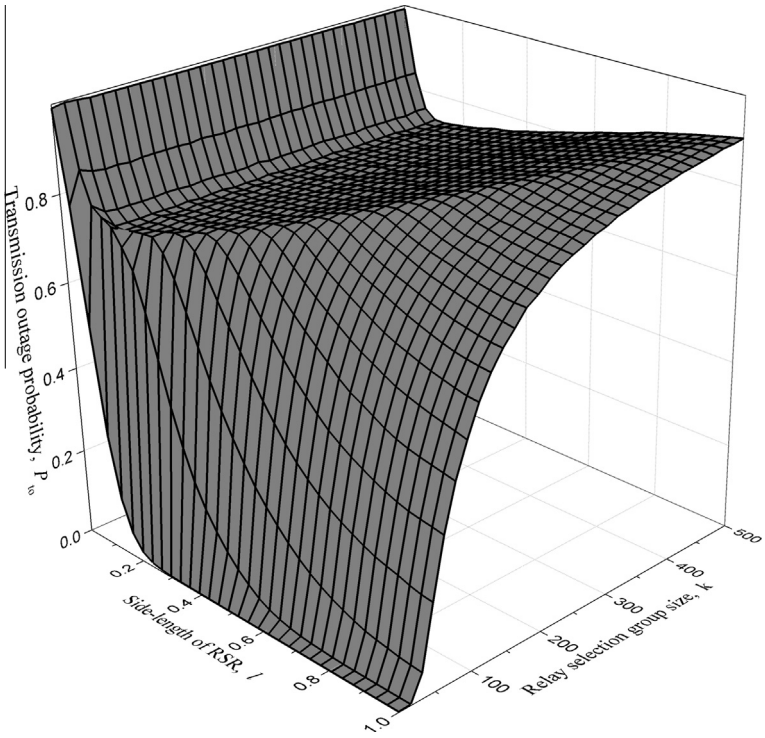

Fig. 4. Transmission outage probability vs. k and l, when $n = 500,\ m = 5,\ \tau = 1.0,\ r = 0.1,\ \gamma_R = 0.05$ and $\gamma = 0.04$.

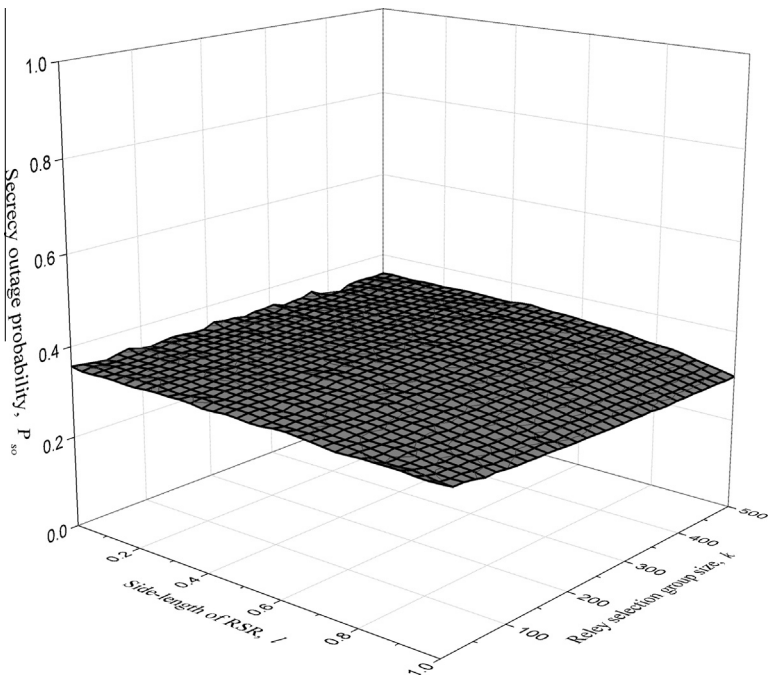

Fig. 5. Secrecy outage probability vs. k and l, when $n = 500$, $m = 5,\ \tau = 1.0,\ r = 0.1,\ \gamma_R = 0.05$ and $\gamma = 0.04$.

becoming negligible but the number of relays in the RSR is still less than k. In this case. the probability that the decoding outage happens increases since the relay is randomly selected and either of the distance of $S - R_{j^*}$ and that of $R_{j^*} - D$ increases. At last, when l is large enough such that there are more than k relays in the RSR with high probability, the relay

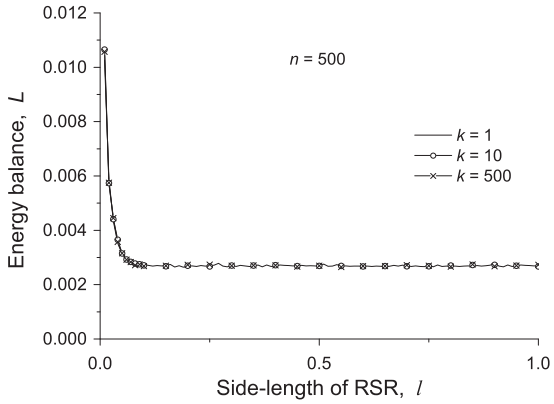

Fig. 6. Energy balance vs. side-length of RSR l for $n = 500$.

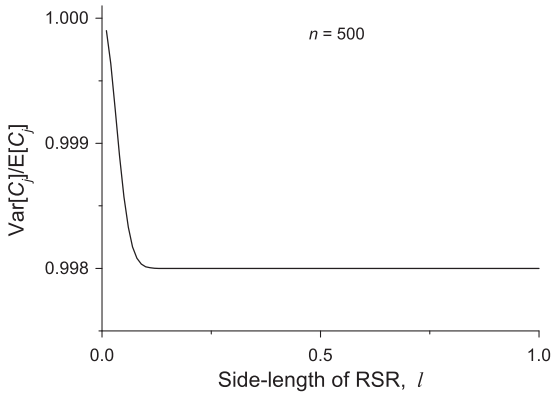

Fig. 7. $Var[C_j]/E[C_j]$ vs. side-length of RSR l for when $n = 500$.

is sub-optimally selected which would increase the conditions of both channels and thus the transmission outage probability decreases. It is notable that the case $l = 1,\ k = 1$ corresponds to global optimal selection and the case $l = 1,\ k = 500$ corresponds to global random selection. We can see obvious alternation from random selection to sub-optimal selection or optimal selection no matter l or k increases. Therefore, the transmission outage probability can be flexibly controlled by the relay selection scheme.

In Fig. 5, we summarize the simulation results to present the secrecy outage probability P_{so} vs. the relay selection group k and the side-length of RSR l. It can be observed from Fig. 5 that the secrecy outage probability decreases very slightly with the side-length of the RSR l for large enough l and the relay selection group k has no impact on the secrecy outage probability. This suggests that the impact of the relay selection scheme in our 2HR-(l, k, r, τ) on the secrecy outage probability can be neglected.

4.3. Energy balance

To evaluate how the performance of energy balance varies with the side-length of RSR l and the relay selection group size k, we summarize the simulation results in Fig. 6 for l increasing from 0.01 to 1 with $k = 1,\ k = 10$ and $k = n = 500$. Moreover, to verify the efficiency of our theoretical analysis for energy balance, we provide theoretical results of $Var[C_j]/E[C_j]$ vs. side-length of RSR l in Fig. 7.

It can be observed from Fig. 6 that the energy balance L decreases dramatically with l at the beginning and then remain constant. This suggests that the energy balance can be controlled by changing the side-length of RSR l for small l. We also can observe from the three curves that the relay selection group size has no impact on the energy balance, which agrees with Theorem 1. It can be observed from Fig. 7 that $Var[C_j]/E[C_j]$ has the similar behavior as the energy balance, as l increases from 0.01 to 1.0, which implies that our theoretical result in Theorem 1 is efficient in the evaluation of energy balance.

5. Conclusion

This paper considers the problem of securing big data communication in a two-hop wireless network with eavesdroppers of unknown channel and location information, and proposed a 2HR-(l, k, r, τ) protocol to achieve secure and reliable

communication. Cooperative jamming is adopted to provide the physical layer security for big data. The results in this paper indicate that our protocol can provide flexible control of security and reliability through proper settings of these four parameters and the energy balance can be controlled by proper setting of the side-length of RSR l. The results also reveal that the relay selection scheme can be alternated between random and optimal by changing either the side-length of RSR l or the relay selection group size k.

References

[1] R. Bergelt, M. Vodel, W. Hardt, Energy efficient handling of big data in embedded, wireless sensor networks, in: Sensors Applications Symposium (SAS), IEEE, February 2014, pp. 53–58.
[2] C. Capar, D. Goeckel, Network coding for facilitating secrecy in large wireless networks, in: 46th Annual Conference on Information Sciences and Systems (CISS), March 2012, pp. 1–6.
[3] C. Capar, D. Goeckel, B. Liu, D. Towsley, Secret communication in large wireless networks without eavesdropper location information, in: Proceedings IEEE INFOCOM, 2012, pp. 1152–1160.
[4] Z. Ding, K. Leung, D. Goeckel, D. Towsley, Opportunistic relaying for secrecy communications: cooperative jamming vs. relay chatting, IEEE Trans. Wirel. Commun. 10 (6) (2011) 1725–1729.
[5] L. Dong, Z. Han, A. Petropulu, H. Poor, Secure wireless communications via cooperation, in: 46th Annual Allerton Conference on Communication, Control, and Computing, September 2008, pp. 1132–1138.
[6] L. Dong, Z. Han, A. Petropulu, H.V. Poor, Improving wireless physical layer security via cooperating relays, IEEE Trans. Signal Process. 58 (2010) 1875–1888.
[7] D. Goeckel, S. Vasudevan, D. Towsley, S. Adams, Z. Ding, K. Leung, Artificial noise generation from cooperative relays for everlasting secrecy in two-hop wireless networks, IEEE J. Sel. Areas Commun. 29 (2011) 2067–2076.
[8] S. Goel, R. Negi, Secret communication in presence of colluding eavesdroppers, in: IEEE Military Communications Conference, 2005, pp. 1501–1506.
[9] B. Han, J. Li, J. Su, M. Guo, B. Zhao, Secrecy capacity optimization via cooperative relaying and jamming for WANETs, IEEE Trans. Parall. Distrib. Syst. (99) (2014). pp. 1–1.
[10] A. Hero, Secure space-time communication, IEEE Trans. Inform. Theory 49 (12) (2003) 3235–3249.
[11] A. Joux, A tutorial on high performance computing applied to cryptanalysis, in: EUROCRYPT'12, 2012, pp. 1–7.
[12] A. Katal, M. Wazid, R. Goudar, Big data: issues, challenges, tools and good practices, in: Sixth International Conference on Contemporary Computing (IC3), August 2013, pp. 404–409.
[13] O. Koyluoglu, C. Koksal, H. El Mamal, On secrecy capacity scaling in wireless networks, IEEE Trans. Inform. Theory 58 (5) (2012) 3000–3015.
[14] J.K. Laurila, D. Gatica-Perez, I. Aad, J. Blom, O. Bornet, T.-M.-T. Do, O. Dousse, J. Eberle, M. Miettinen, The mobile data challenge: big data for mobile computing research, in: Mobile Data Challenge by Nokia Workshop, in Conjunction with Int. Conf. on Pervasive Computing, Newcastle, UK, 158, 2012.
[15] C. Leow, C. Capar, D. Goeckel, K. Leung, Two-way secrecy schemes for the broadcast channel with internal eavesdroppers, in: Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), November 2011, pp. 1840–1844.
[16] J. Li, A. Petropulu, S. Weber, On cooperative relaying schemes for wireless physical layer security, IEEE Trans. Signal Process. 59 (10) (2011) 4985–4997.
[17] Y. Liu, J. Li, A. Petropulu, Destination assisted cooperative jamming for wireless physical-layer security, IEEE Trans. Inform. Forensics Secur. 8 (4) (2013) 682–694.
[18] U. Maurer, Secret key agreement by public discussion from common information, IEEE Trans. Inform. Theory 39 (3) (1993) 733–742.
[19] S. Narayanan, P. University, Two-hop Forwarding in Wireless Networks, Polytechnic University, 2006.
[20] R. Negi, S. Goel, Secret communication using artificial noise, in: IEEE Vehicular Technology Conference, 2005, pp. 1906–1910.
[21] R. Schell, Security – a big question for big data, in: IEEE International Conference on Big Data, October 2013, pp. 5–5.
[22] A. Sheikholeslami, D. Goeckel, H. Pishro-Nik, D. Towsley, Physical layer security from inter-session interference in large wireless networks, in: Proceeding of IEEE INFOCOM, 2012, pp. 1179–1187.
[23] Y. Shen, X. Jiang, J. Ma, Flexible relay selection for secure communication in two-hop wireless networks, in: 11th International Symposium on Modeling Optimization in Mobile, Ad Hoc Wireless Networks (WiOpt), 2013, pp. 648–651.
[24] Y. Shen, Y. Zhang, Exploring relay cooperation for secure and reliable transmission in two-hopwireless networks, EAI Endorsed Trans. Scalable Inform. Syst. 2 (2014) e2.
[25] J. Talbot, D. Welsh, Complexity and Cryptography: An Introduction, Cambridge University Press, 2006.
[26] S. Vasudevan, S. Adams, D. Goeckel, Z. Ding, D. Towsley, K. Leung, Multi-user diversity for secrecy in wireless networks, in: Information Theory and Applications Workshop (ITA), January 2010, pp. 1–9.
[27] S. Vasudevan, D. Goeckel, D. Towsley, Security-capacity trade-off in large wireless networks using keyless secrecy, in: ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2010, pp. 21–30.
[28] A.D. Wyner, The wire-tap channel, Bell Syst. Tech. J. 54 (8) (1975) 1355–1387.
[29] Y. Zhang, C++ Simulator for the 2hr-(l,k,r,tau) Protocol, 2014. <http://mdlval.blogspot.jp/>.