sensors

MDPI

Article

Physical Layer Secret-Key Generation Scheme for Transportation Security Sensor Network

Bin Yang and Jianfeng Zhang *

College of Information Engineering, Northwest A&F University, Xianyang 712100, China; b_yang@nwsuaf.edu.cn

* Correspondence: zjf@nwsuaf.edu.cn; Tel.: +86-298-709-2338; Fax: +86-298-709-2315

Received: 13 April 2017; Accepted: 24 June 2017; Published: 28 June 2017

Abstract: Wireless Sensor Networks (WSNs) are widely used in different disciplines, including transportation systems, agriculture field environment monitoring, healthcare systems, and industrial monitoring. The security challenge of the wireless communication link between sensor nodes is critical in WSNs. In this paper, we propose a new physical layer secret-key generation scheme for transportation security sensor network. The scheme is based on the cooperation of all the sensor nodes, thus avoiding the key distribution process, which increases the security of the system. Different passive and active attack models are analyzed in this paper. We also prove that when the cooperative node number is large enough, even when the eavesdropper is equipped with multiple antennas, the secret-key is still secure. Numerical results are performed to show the efficiency of the proposed scheme.

Keywords: transportation security; secret-key generation; physical layer security; wireless sensor network

1. Introduction

Wireless sensors are widely used in transportation systems to ensure the system security, lower the fuel consumption and increase system efficiency etc. [1–3]. Cargo shipments security is a critical challenge for shippers, every year cargo theft costs billions of dollars. A Transportation Security Sensor Network (TSSN) architecture is developed to realize the vision of trusted corridors [3]. In TSSN the cargo security is monitored with active and battery-powered container seals (sensors) to report security seal events timely. Any unauthenticated attempts to unlock the container will be reported to the operations center through mobile network and internet. A security Seal Interrogation Transceiver (SIT) is designed to communicate with the container seals over a wireless network, which should be secured with secret-key, or else an adversary could interfere the wireless communication link to disable the alert system.

The security issue in TSSN is similar of that in Wireless Sensor Networks (WSNs). There are two main challenges on the secure communication: 1. the low cost of wireless nodes leads to severe resource constraints such as limited battery power, memory and low computation capability; 2. the open nature of the wireless link makes it easy to be eavesdropped. The normally used public cryptography approaches are not suitable for TSSN because sensor nodes (container seals) are source constrained devices that cannot afford for public key cryptography. Therefore, symmetric key-based schemes are widely used in WSNs, because of the advantages of low cost in power comsumption, time execution and code size [4,5]. The main challenge for WSNs to implement symmetric key-based scheme is the secret-key distribution in the network.

Physical layer key generation schemes could offer a solution of the issue [6]. The main advantage

of physical layer key generation scheme is that the key is directly generated in physical layer, and there is not any key distribution process. In an ideal situation the eavesdropper could not get any information about the key.

The theoretical aspects of secrecy extraction from correlated random source have been firstly studied by [6,7]. It is shown that correlated observations of random sources could be used to distill secret-keys by discussing over a public channel, while the information rate leaked to the eavesdropper can be arbitrarily low. The supremum of achievable secret-key rate is called secret-key capacity. In recent years, significant interests in developing practical approaches to generate secret-key between multiple users have been attracted [8–34]. It is shown in [9,11] that there is a trade-off between the secret-key rate and the public communication rate in the key agreement protocols.

One of the main issues of these schemes is how to find proper random sources for secret-key generation. Such sources should create correlated randomness between the legitimate users, would have high level of randomness, and should be difficult for the eavesdropper to observe.

In this paper we propose a method to create artificial correlated random sources for wireless sensors to generate secret-key in a cooperative TSSN. The random sources are created by multiple nodes in the system, when the cooperative helpers send independent symbols simultaneously, different channel vectors result in different receiving signals, which prevents the eavesdropper from getting a copy of the legitimate users' signals. Since the random source is artificially generated, even when CSI (Channel State Information) of the wireless channels changes slowly, high secret-key rate can still be achievable.

In the proposed scheme, the helpers have no idea of the receiving signals of the users. When the eavesdropper is equipped with multiple antennas or there are multiple eavesdroppers, it is possible for them to get what the helpers send. However, since the eavesdroppers have no idea about the legitimate nodes' CSI, they still cannot get what the legitimate users get. It is proven that, from computational complexity security point of view, the proposed scheme is secure with enough helpers even when the antenna number of the eavesdropper is unlimited.

In [8], the random source is also artificial signals, secret-key is generated by opportunistic transmission over the quasi-static fading channel by sending signals when the channel condition of the legitimate users are better than the eavesdropper's. However, this approach is based on certain assumptions that are hard to be realized in practice, while the proposal in this paper does not have. In [27,33], authors have investigated the impact of cooperative relay nodes on the secret-key generation, these algorithms are based on the wireless channel reciprocity of the users and relays. Since the achievable secret-key rate scales linearly with the number of relays, when there are large number of available relays, better system performance could be achieved. While the proposed scheme in this paper can generate secret-key artificially, our proposal is more suitable for the system with relatively small number of relays. The scheme of [34] is based on the knowledge of the eavesdropper and the communication capacity of the nodes is unlimited, that is not practical in a real system.

The organization of the paper is as follows. Section 1 introduces the proposed scheme of the system. Section 2 presents secret-key rate analysis. Section 3 studies the system security with different thread models. Numerical result is presented in Section 4. Finally, we conclude the paper in Section 5.

2. Proposed Scheme

The system is shown in Figure 1, which is a wireless sensor network equipped in a rail-borne cargo. There are two types of nodes in the network. One is Cargo Sensor (CS), which is placed on the cargo container. Normally CS is battery-powered and designed to monitor security seal events or the status of the cargo. The other type of node in the network is Monitor Center (MC), which could be installed in the cab of the locomotive. CS reports sensor data to MC or receives commands from MC through secure wireless link, which is encrypted by symmetric cryptographic scheme. In this paper, we propose a novel key agreement or key distribution scheme in TSSN.

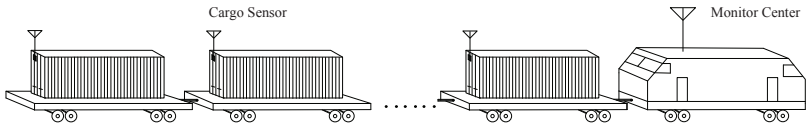

Figure 1. System Configure. In the system, cargo sensors (CS) are placed on the cargo containers to monitor or secure the containers. The monitor center (MC) is normally placed in locomotive. MC communicate with CS through wireless link, which should be secured by cryptography algorithm.

The proposed scheme is shown in Figures 2 and 3. MC tries to update the secret-key of one of the CSs which is marked as A (the red container in Figure 3). All the communication parties can communicate with each other through a public channel. Other CSs can help MC and A to achieve the goal, we call them helpers. Note that helpers are not fixed, A could also play as a helper when MC wants to update the key of another CS.

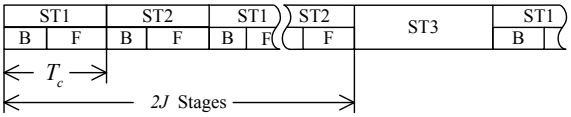

Figure 2. Communication frame structure.

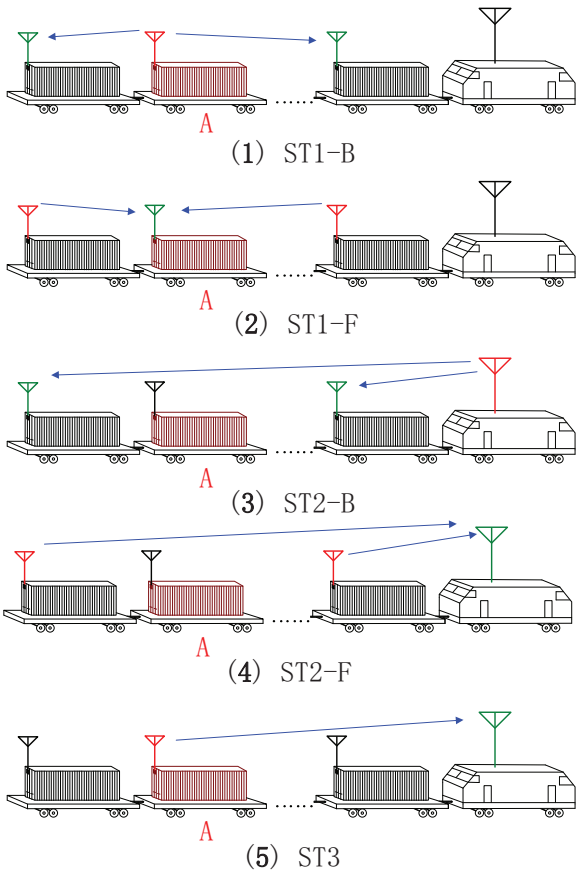

Figure 3. Algorithm steps. (**1**) A sends channel estimating sequence to helpers; (**2**) helpers transmit random signals simultaneously to A; (**3**) MC sends channel estimating sequence to helpers; (**4**) helpers send reversed random signals simultaneously to MC; (**5**) A communicates with MC to get agreement on the secret-key. Note that in the figure when sending signals the antenna is colored in red, when receiving signals the antenna is colored in green.

Here we assume the number of the helpers is $N, N \geq 3$. A passive eavesdropper is located at somewhere trying to crack the secret-key, he can access the signals from all the communication partners. The system is a narrow band system, the wireless channels are assumed to be block fading with the coherent time T_c. Note that all the nodes in the system are equipped with single antenna.

The railway radio propagation environment is significantly impacted by the railway structures, such as cuttings, viaducts, and tunnels etc., so the wireless channel's propagating characteristics changes while the train traveling from one site to another. Since the system is a narrow band system, the channel can be modeled as a complex random variable. The complex random channel gain is considered to keep constant within the coherent time T_c. Then the nodes in the system have to measure and renew the channel gain when the transmitting time is longer than T_c.

There are three stages in the proposed scheme.

In Stage 1 (ST1) within T_c, all the helpers synchronously send random symbols to A to create stochastic signals at A. There are two sub-stages in ST1: the first is the backward signal transmission sub-stage (ST1-B) from partner A to the helpers. In this sub-stage, User A sends channel estimation sequence to the helpers, then they can estimate the channel coefficients of the links from A to helpers. Due to the radio propagation reciprocity, the helpers then know the channel gains of the links from themselves to A. On the other hand, the eavesdropper can also get the channel estimation sequence from A, then he can estimate the channel gain between himself and A. Because user A and helpers do not send any messages about the channel information, the eavesdropper cannot get any information about the helpers' links directly.

The second sub-stage in ST1 is forward signal transmission sub-stage (ST1-F) from the helpers to user A. All the helpers individually send K random symbols to user A, then the symbols that user A receives are

$$x(k) = \sum_{i=1}^{N} h_i s_i(k) + n_1(k) \tag{1}$$

where $h_i,\ i = 1, \ldots, N$ denotes the complex channel gains between the helpers and A, $s_i(k),\ i = 1, \ldots, N$ denotes complex zero-mean Gaussian random symbols sent by the helpers which are independent of each other, and $n_1(k)$ denotes the receiving noise of user A. Here we assume that $|h_i|^2 \neq 0$, which is a reasonable assumption for a practical system. Then the transmitting power of the helpers is $P_i = E(|s_i|^2),\ i = 1, \ldots, N$. In the proposed scheme, we set all the transmitting power to be the same, that is $P_i = P,\ i = 1, \ldots, N$.

During Stage 2 (ST2), all the helpers repeat the K symbols sent in the first stage multiplied with weight factors $w_i,\ i = 1, \ldots, N$. There are also two sub-stages in this stage. The first is the backward sub-stage (ST2-B) for channel estimation. In this sub-stage, MC send channel estimation sequence to the helpers to estimate the complex channel gain. In addition, the eavesdropper also can only get the channel information between MC and himself. Next is the forward sub-stage (ST2-F) for the helpers to repeat the K symbols, then the symbols that MC receives are

$$y(k) = \sum_{i=1}^{N} g_i w_i s_i(k) + n_2(k), \tag{2}$$

where $g_i,\ i = 1, \ldots, N$ denotes the complex channel gains between the helpers and MC, and $n_2(k)$ is the noise of MC. We also assume that $|g_i|^2 \neq 0$. To create correlated sources for the partners, we set

$$w_i = \sqrt{\rho}\frac{h_i}{g_i},\ i = 1, \ldots, N, \tag{3}$$

where ρ is power factor to adjust the total transmitting power in the second stage. Note that w_i is only determined by the helper's own CSI and a global factor ρ, no other global information should be shared between the helpers. Then the receiving symbols of MC are

$$y(k) = \sqrt{\rho}\sum_{i=1}^{N} h_i s_i(k) + n_2(k). \tag{4}$$

Then $y(k) = \sqrt{\rho}x(k) - \sqrt{\rho}n_1(k) + n_2(k)$, when signal power is high enough, the symbols received by MC and A are almost the same.

Note that if $|g_i|$ is a small low value, which means the link between MC and helper i is very weak. Then the transmit power of helper i during ST2-F has to be very high, and possibly exceeds the maximum transmit power of the helper. If the helper repeat the random symbols only with maximum power, MC's receiving signals will not be the same as A's without considering noises. There are three solutions: One is that MC can use antenna with high antenna gain, for example, using directional antenna, then usually $|g_i|$ will be lower $|h_i|$. The other solution is that helpers can adjust the global power factor ρ to make sure all the helper's transmit power in ST2 is lower than the upper limit. However, this solution has the risk to leak some information of the legitimate user's channel information. The last one is that the helper i is mute in ST2-F when the transmit power should be higher than the maximum power of himself. In an additional stage, helper i broadcast his signals sent in ST1, and tells every one that he is mute in ST2. Then A can remove what the helper i sends in his receiving signals with the check-sum information from MC in ST3. This additional stage will not influence the model and analysis in the paper.

In Stage 2, we assume perfect channel estimation of h_i and g_i, which is not true in a practical system. When considering channel estimation errors, the receiving signals of MC will be

$$\begin{aligned} y'(k) &= \sqrt{\rho}\sum_{i=1}^{N} \frac{h_i + \Delta h_i}{g_i + \Delta g_i} g_i s_i(k) + n_2(k) \\ &= \sqrt{\rho}\epsilon \sum_{i=1}^{N} h_i s_i(k) + \sqrt{\rho} n_\Delta(k) + n_2(k), \end{aligned} \tag{5}$$

where Δh_i and Δg_i are the channel estimation errors, and

$$\epsilon = \frac{\sum_{i=1}^{N} \frac{h_i+\Delta h_i}{g_i+\Delta g_i} g_i h_i^*}{\sum_{i=1}^{N} |h_i|^2}, \quad n_\Delta(k) = \sum_{i=1}^{N} \left[\frac{h_i + \Delta h_i}{g_i + \Delta g_i} g_i - \frac{\sum_{i=1}^{N} \frac{h_i+\Delta h_i}{g_i+\Delta g_i} g_i h_i^*}{\sum_{i=1}^{N} |h_i|^2} h_i \right] s_i(k). \tag{6}$$

It is easy to know that

$$E_{s_i}\left(n_\Delta(k) \left(\sum_{i=1}^{N} h_i s_i(k) \right)^* \right) = 0, \tag{7}$$

which means $n_\Delta(k)$ can be considered as additional Gaussian noise. Then channel estimation errors will result in an additional noise to the receiving signals of MC, and in turn decrease the performance of the system. In this paper, we mainly focus on the upper bound of the system performance, the further discussion of the model in a real system is left for future work.

During the above two stages, the eavesdropper gets the signals as

$$v_1(k) = \sum_{i=1}^{N} e_i s_i(k) + n_{E1}(k), \; v_2(k) = \sqrt{\rho}\sum_{i=1}^{N} e_i \frac{h_i}{g_i} s_i(k) + n_{E2}(k), \tag{8}$$

where e_i, $i = 1, \ldots, N$ denotes the complex channel gains between the helpers and the eavesdropper, $n_{E1}(k)$ and $n_{E2}(k)$ are the noises of the eavesdropper in the two stages respectively.

Note that all the noise terms n_1, n_2, n_{E1} and n_{E2} are zero-mean white independent complex

Gaussian random variables with the variance σ_n^2.

After $2J$ stages of transmission, there should be a stage of hand-shaking, which is Stage 3 (ST3). During ST3, A and MC exchange information to distill a common secret-key. In this paper, we concentrate on the class of key agreement protocols in which only A sends messages to MC. A computes a secret-key $\mathcal{M}$ and sending message $\mathcal{C}$ from x, then sends $\mathcal{C}$ to MC over the public channel, the total length of these sequences is JK. MC then computes the key $\mathcal{M}'$ from y and $\mathcal{C}$. A secret-key rate R_k is achievable if for any $\epsilon > 0$ and all sufficiently large number JK, there is a secret-key agreement such that

$$\frac{H(\mathcal{M})}{JK} \geq R_k - \epsilon, \tag{9}$$

$$Pr\{\mathcal{M} \neq \mathcal{M}'\} \leq \epsilon, \text{ (Reliability Condition)} \tag{10}$$

$$\frac{I(\mathcal{M}; v_1, v_2, \mathcal{C})}{JK} \leq \epsilon. \text{ (Secrecy Condition)} \tag{11}$$

Then secret-key capacity is defined as the supremum of secret-key rates achievable for the model.

Since the length of $\mathcal{C}$ is relatively long, the transmission of $\mathcal{C}$ may be divided into several frames, which depends on the coherent time T_c. Here we do not show any details on ST3, for the information about the link between A and B is public, and does not affect the key generation process in this paper.

In general, a closed-form expression of the secret-key capacity is still an open problem. Nevertheless, in [6,7], upper bound and lower bound of the secret-key capacity are shown. Since we focus on the design of the correlated sources for secret-key generation in this paper, a tight upper bound is good enough for demonstrating the effectiveness of the proposal.

3. Secret-Key Rate Analysis

Security Capacity of the Scheme

The model in this paper is a typical source model (Section 4.1 in [35]) of secret-key agreement, which represents a situation in which the two parties in the system observe the realizations of a random source to generate secret-key. A closed-form expression for the secret-key capacity for a general source model remains elusive. Then from the results in [6,7,35], with considering the one-way public communication between the legitimate users, the secret-key rate of our model is bounded as

$$I(x;y) - I(x;\mathbf{v}) \leq C_{sk} \leq I(x;y|\mathbf{v}), \tag{12}$$

where $I(\cdot;\cdot)$ is mutual information, and $\mathbf{v} \doteq (v_1, v_2)^T$. Here we define

$$R^- \doteq I(x;y) - I(x;\mathbf{v}), \quad R^+ \doteq I(x;y|\mathbf{v}). \tag{13}$$

We will show that the bounds are tight in our model when the total transmitting power of the helpers is large enough, then the upper bound could be a proper substitution of the secret-key capacity to show the performance of the proposed algorithms.

We know

$$I(x;y|\mathbf{v}) = H(x|\mathbf{v}) - H(x|y\mathbf{v}), \tag{14}$$

then we have

$$H(x|\mathbf{v}) = \log|\mathbf{Q}_{x\mathbf{v}}| - \log|\mathbf{Q}_{\mathbf{v}}| + \log(\pi e), \tag{15}$$

$$H(x|y\mathbf{v}) = \log|\mathbf{Q}_{xy\mathbf{v}}| - \log|\mathbf{Q}_{y\mathbf{v}}| + \log(\pi e), \tag{16}$$

where $\mathbf{Q}_{x\mathbf{v}}$, $\mathbf{Q}_{\mathbf{v}}$, $\mathbf{Q}_{xy\mathbf{v}}$ and $\mathbf{Q}_{y\mathbf{v}}$ are all covariance matrices, which are given as

$$\mathbf{Q}_{x\mathbf{v}} \doteq E\left(\begin{bmatrix} x \\ \mathbf{v} \end{bmatrix}[x^*, \mathbf{v}^H]\right),$$

$$\mathbf{Q}_{\mathbf{v}} \doteq E(\mathbf{v}\mathbf{v}^H),$$

$$\mathbf{Q}_{xy\mathbf{v}} \doteq E\left(\begin{bmatrix} x \\ y \\ \mathbf{v} \end{bmatrix}[x^*, y^*, \mathbf{v}^H]\right),$$
$$\mathbf{Q}_{y\mathbf{v}} \doteq E\left(\begin{bmatrix} y \\ \mathbf{v} \end{bmatrix}[y^*, \mathbf{v}^H]\right).$$

Then we have

$$|\mathbf{Q}_{x\mathbf{v}}| = \sum_{i>j>k} \alpha_{ijk} P^3 + \sigma_n^2 \left(\sum_{i>j} \beta_{ij} P^2 + \sigma_n^2 \left(\sum_i \gamma_i P + \sigma_n^2\right)\right), \tag{17}$$

$$|\mathbf{Q}_{\mathbf{v}}| = \sum_{i>j} \mu_{ij} P^2 + \sigma_n^2 \left(\sum_{i=1}^{N} \eta_i P + \sigma_n^2\right), \tag{18}$$

$$\left|\mathbf{Q}_{y\mathbf{v}}\right| = \rho \left|\mathbf{Q}_{x\mathbf{v}}\right| + (1-\rho)\sigma_n^2 \left|\mathbf{Q}_{\mathbf{v}}\right|, \tag{19}$$

$$\left|\mathbf{Q}_{xy\mathbf{v}}\right| = (1+\rho)\sigma_n^2 \left|\mathbf{Q}_{x\mathbf{v}}\right| - \rho\sigma_n^4 \left|\mathbf{Q}_{\mathbf{v}}\right|, \tag{20}$$

where

$$\alpha_{ijk} = \rho \left| e_j h_i h_k \left(\frac{e_i}{g_i} - \frac{e_k}{g_k}\right) - e_i h_j h_k \left(\frac{e_j}{g_j} - \frac{e_k}{g_k}\right) - e_k h_i h_j \left(\frac{e_i}{g_i} - \frac{e_j}{g_j}\right) \right|^2, \tag{21}$$

$$\beta_{ij} = \left|h_i e_j - h_j e_i\right|^2 + \rho \left| \frac{h_i h_j e_j}{g_j} - \frac{h_j h_i e_i}{g_i} \right|^2 + \rho \left| \frac{e_i h_j e_j}{g_j} - \frac{e_j h_i e_i}{g_i} \right|^2, \tag{22}$$

$$\gamma_i = |h_i|^2 + |e_i|^2 + \rho \left| \frac{e_i h_i}{g_i} \right|^2, \tag{23}$$

$$\mu_{ij} = \rho |e_i e_j|^2 \left| \frac{h_i}{g_i} - \frac{h_j}{g_j} \right|^2, \tag{24}$$

$$\eta_i = |e_i|^2 \left(1 + \rho \left| \frac{h_i}{g_i} \right|^2\right). \tag{25}$$

All the above coefficients are positive, and because

$$\beta_{ij} - \mu_{ij} = \left|h_i e_j - h_j e_i\right|^2 + \rho \left| \frac{h_i h_j e_j}{g_j} - \frac{h_j h_i e_i}{g_i} \right|^2 > 0, \tag{26}$$

$$\gamma_i - \eta_i = |h_i|^2 > 0, \tag{27}$$

the coefficients of $\left|\mathbf{Q}_{y\mathbf{v}}\right|$ and $\left|\mathbf{Q}_{xy\mathbf{v}}\right|$ are all positive too.

We also have

$$I(x;y) - I(x;\mathbf{v}) \quad = \quad H(x|\mathbf{v}) - H(x|y), \tag{28}$$

where

$$H(x|y) = \log\left(\frac{(1+\rho)\sigma_n^2\sum_{i=1}^{N}|h_i|^2P + \sigma_n^4}{\rho\sum_{i=1}^{N}|h_i|^2P + \sigma_n^2}\right) + \log(\pi e). \tag{29}$$

Then we have the following theorem:

Theorem 1. *When there are at least three nodes sending signals, we have* $\lim_{P\to\infty}(R^+ - R^-) = 0$.

Proof. Please see Appendix A. □

The result of Theorem 1 is simple, and can be explained from another point of view: When the signal power is infinite, x and y are almost the same, then x, y and $\mathbf{v}$ almost form a Markov chain. Consequently, from the conclusion in [7], the bounds are tight.

Theorem 1 limits the number of helpers to be at least three. Because the eavesdropper has two inputs during the first two stages, he can then be considered as a receiver equipped with two antennas. When there are only one or two helpers sending signals, the eavesdropper can decode the signals especially when SNR is high. If we want to achieve positive secret-key rate, there should be at least three helpers who send signals with large enough transmitting power in the system.

From Theorem 1, we know that when the transmitting power is large enough the lower bound and upper bound of the secret-key rate of the proposed scheme are almost the same, that is, the bounds of the secret-key capacity are tight. In the following parts of the paper, we just analyze the upper bound of the secret-key capacity.

If we set the total transmit power of the second stage to be the same as that of the first stage, namely

$$\rho\sum_{i=1}^{N}\left|\frac{h_i}{g_i}\right|^2 P = NP, \tag{30}$$

then we have $\rho = \frac{N}{\sum_{i=1}^{N}\left|\frac{h_i}{g_i}\right|^2}$, when P goes to infinite, we have

$$\begin{aligned}&\lim_{P\to\infty} R^+\\ &= \lim_{P\to\infty}\log NP + \log\left(\frac{\sum_{i>j>k}\alpha_{ijk}}{N\sum_{i>j}\beta_{ij}}\right) - \log\left(1 + \frac{\sigma_n^2}{N}\sum_{i=1}^{N}\left|\frac{h_i}{g_i}\right|^2\right).\end{aligned} \tag{31}$$

Actually, $\sum\alpha_{ijk}$ is impossible to be zero in a practical system. In high SNR region, R^+ increases linearly with the total power $\log(P)$. This means such a simple algorithm can help to get infinite secret-key rate by increasing the total signal power without considering any information of the eavesdropper.

4. Threat Model Analysis

The above analysis on secret-key rate is based on the information about the eavesdropper, we can only get a theoretical upper bound of system performance. Actually, in a real system, we have no idea of the potential eavesdropper, so the information-theoretic security is not available. In this section, we discuss the practical challenges on the system without assumptions on the eavesdropper.

4.1. Passive Attacks

Since a passive eavesdropper does not send any signal, it is hard to estimate how much information is leaked to such an attacker. Different stage of the process has different level of the possibility to leak the confidential messages.

ST1-B and ST2-B: During these sub-stages, the partner A and MC send channel probing signals to the helpers, then they can estimate the channel gains of themselves. For a passive eavesdropper, he can also get the channel probing signals, then the channels state information of his own is known by the eavesdropper. However, the legitimate system links is still unknown by the eavesdropper. In a poor multipath scattering environment, the eavesdropper may have a strong correlation in measurements of the wireless channels. However, since there are many helpers in the system, the eavesdropper could not have correlation with all the helpers. So when the helper number is large enough, the leaking information to the eavesdropper is limited.

ST1-F and ST2-F: In these sub-stages, all the helpers send signals simultaneously to partner A. The eavesdropper can also get the signals. However, the two signals are different linear combination of $s_i(k), i = 1, \ldots, N$. When the node number N goes to infinite, the eavesdropper's signals is independent of what node A and MC receive. More node number can help the system to achieve higher security level.

When the eavesdropper is placed very near A or MC, or else the whole system works in a poor multipath scattering environment, the eavesdropper's channels might have a strong correlation with the legitimate users. In these cases, the phase of the complex channel coefficients is usually still independent of each others, while the modulo of the channel coefficients or the RSS (received signal strength) are correlated to each others. The proposed scheme is even very sensitive of the phases, which is demonstrated in Figure 4.

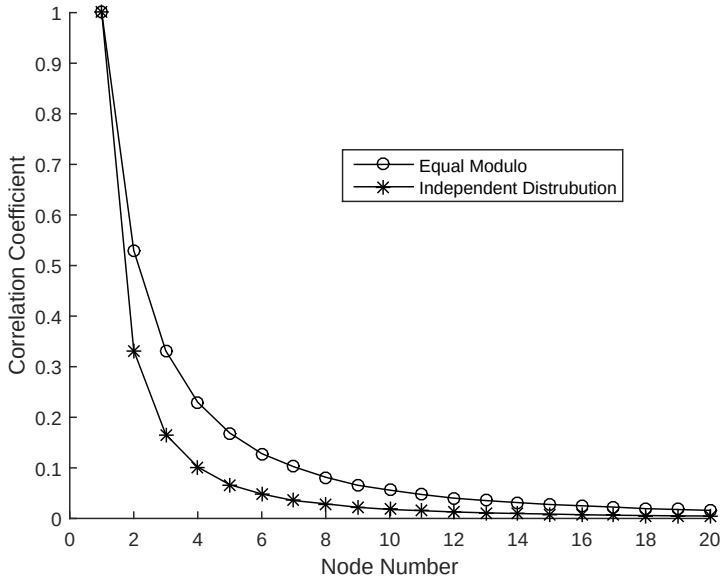

Figure 4. Average correlation coefficients of $x(k)$ and $v_1(k)$ in ST1 vs. helper number. Equal Modulo means $|h_i| = |e_i|, i = 1, \ldots, N$, with $h_i, i = 1, \ldots, N$ are independent Rayleigh fading channels, and $e_i = |h_i|e^{j\phi_i}$ where $\phi_i, i = 1, \ldots, N$ are independent random phases uniformly distributed within $[0, 2\pi)$. Independent Distribution means h_i is independent of e_i, with $h_i, i = 1, \ldots, N$ and $e_i, i = 1, \ldots, N$ are all independent Rayleigh fading channels. We do independent experiments 10,000 times to get the average values without considering the receiving noises.

In Figure 4, the worst and best cases are demonstrated. The best situation is that the eavesdropper's channels are independent of the legitimate users'. The worst case is that the eavesdropper's channels are highly related with the legitimate users', even the extreme case is $|h_i| = |e_i|$, that is the eavesdropper's channels have the same modulo of A's. We consider the average correlation coefficient of x and v_1 to show how the helper number improve the system security. The average correlation coefficient is defined as

$$\zeta = E_{\mathbf{h},\mathbf{e}} \left(\frac{|Cov(x, v_1)|}{\sqrt{Var(x)Var(v_1)}} \right). \tag{32}$$

It is shown in Figure 4 that even the channels are highly correlated, the signals intercepted by the passive eavesdropper is almost independent of the legitimate receiver's when the helper number is large enough, which means the eavesdropper almost could not get any useful information from the receiving signals.

ST3: ST3 is usually an error correction process. Due to error of channel measuring and noise, the extracted bits at A and MC sides are usually not identical. During this stage, parity bit information may be transmitted openly to correct errors. The eavesdropper can get the messages in this stage. In order to eliminating the eavesdropper's partial information about the key, there will be a privacy amplification process [6]. In the privacy amplification phase, both legitimate parts compress the information to their "real entropy". However, in a practical system, the information leaked to the eavesdropper is hard to estimated, then the "real entropy" is hard to be decided. A possible solution is to get average or maximum leaking information by experiments.

Multi-Antenna attack: The most threatening passive attack is multi-antenna attack, that is, the eavesdropper is equipped with multiple antennas to intercept the signals. In this case, the eavesdropper performs like a MIMO (Multiple-In-Multiple-Out) system . In an ideal situation, when the antenna number of the eavesdropper is unlimited, the eavesdropper could know exactly the signals from every transmitting antenna. Then the receiving signals of the eavesdropper can be considered as $s_i, i = 1, \ldots, N$. It is easy to know that $I(x; y|s_1, \ldots, s_N) = 0$, then the secret-key rate of the system is down to zero, that is, the system cannot get information theory secrecy any more. However we will prove in a following subsection that the system is still secure with enough number of helpers.

4.2. Active Attacks

In [36], the author classifies the now existing active attacks into three types: disruptive jamming attack, manipulative jamming and channel manipulation attack.

Disruptive jamming: The purpose of disruptive jamming attacks is to minimize the key generation rate between legitimate users. The jamming signals can be injects in every stage of the proposed scheme. Most harmful behavior is to disrupt the channel probing process, without accurate channel estimation, the secret-key rate of the system will be dramatically deduced. A possible solution of this issue is proposed in [36], random probing signals is used to hide the channel state information, which is also suitable of the scheme of this paper.

Manipulative jamming: In [37], a manipulative attack is proposed (Man-In-The-Middle Attack) to control the channel measurements at legitimate users. In our proposal, it does not work, any misleading of the channel measurements will cause failure on the key generation. Manipulative jamming on channel probing process will deduce the secret-key rate instead of compromising the generated key.

A possible way of manipulative jamming attack is that the attacker can transmit signals with high enough power in substage ST1-F and ST2-F, then the receiving signals of A and MC are mainly controlled by the attacker, thus the generated key is compromised by the malicious third-party. Since the attacker acts just like a normal helper, this type of attack is hard to be defended against.

One possible solution to address the issue is power detecting: After generating a new key, all the helpers can report through the open channel about the average power during the period they transmitted to MC (This action will leak part of the channel information to the eavesdropper, which should be considered in ST3). Then MC can compare the average power that he has received and the messages helpers reports. If there be an attacker, the receiving signal power will be higher than the sum of all the reported power, which means the generated key is possibly manipulated. The more attacker controls, the more easy he would be detected.

Channel manipulation attack: Because the key is not generated from the channel information, channel manipulation attack can not influence the key generation process of this paper. On the other hand, since the system is equipped in a train which travels through a long distance, channel manipulation is almost impossible for any potential attacker.

4.3. Multi-Antenna Attack

From the scheme, we know that the eavesdropper cannot get any information of the channel gain $h_i, i = 1, 2, \ldots, N$ from his receiving signals v_1 and v_2. When information theory secrecy is not achievable, the unknown of h_i can still help the system to achieve computationally secure secrecy.

The term of secret-key rate is based on information theory secrecy, or unconditionally secure secrecy. When the eavesdropper is equipped with multiple antennas, or else there are multiple cooperative eavesdroppers, the performance of the system will be lower. If number of the antennas or the eavesdroppers is infinite, the eavesdropper could possible almost know what the helpers send. In this case the system cannot achieve unconditional secrecy any more. However, the system is still computationally secure, which means cracking the secret-key is equivalent to the solution of some problem known to be laborious.

The analysis of this case is valuable, because in a practical system, we cannot limit the number of the eavesdropper's antenna. The proposed scheme is computationally secure when there are infinite eavesdroppers. The reason is that in our system what the helpers send is different from what the users receive, multiple antennas can help the eavesdropper to get the sending signals but not the receiving signals. Without knowledge of the legitimate user's channels, the eavesdropper still cannot crack the secret-key.

We assume that when there are infinite eavesdroppers, they could know exactly what the helpers send during the first two stages in an ideal situation. The eavesdroppers know the symbols $s_i, i = 1, 2, \ldots, N$ and $\sqrt{\rho}\frac{h_i}{g_i}$, i =1, …, N, but they have no idea about $h_i, i = 1, 2, \ldots, N$. The method for the eavesdroppers to crack the secret-key is just guess. Then the problem is what is the probability for the eavesdropper to crack the key for one trial.

Since the receiving signals of the two legitimate users and the eavesdropper are all Gaussian signals, we consider

$$s = \sum_{i=1}^{N} h_i s_i, \tag{33}$$

as the effective signals, then user A, B and the eavesdropper all get a noisy version of s. We re-model the signals at the legitimate users as

$$x = s + n_1,\ y = \sqrt{\rho}s + n_2. \tag{34}$$

When the eavesdropper tries to estimate $h_i, i = 1, \ldots, N$ with $\hat{\boldsymbol{h}}$, where $\hat{\boldsymbol{h}} \doteq (\hat{h}_1, \hat{h}_2, \ldots, \hat{h}_N)^T$ denotes random selected complex vector as the estimation of the CSI, then the estimated signals can be written as

$$\hat{v} = \sum_{i=1}^{N} \hat{h}_i s_i = \delta s + n_e, \tag{35}$$

where n_e denotes the equivalent noise of the estimation and is independent of s, then we have

$$\begin{aligned} \delta &= \frac{\sum_{i=1}^{N} h_i^* \hat{h}_i}{\sum_{i=1}^{N} |h_i|^2}, \\ E(|n_e|^2) &= P \frac{\sum_{i=1}^{N} |h_i|^2 \sum_{i=1}^{N} |\hat{h}_i|^2 - \left|\sum_{i=1}^{N} h_i \hat{h}_i^*\right|^2}{\sum_{i=1}^{N} |h_i|^2} \doteq \sigma_e^2. \end{aligned}$$

The SNRs of the legitimate users and the eavesdropper are

$$\eta_1 = \frac{P}{\sigma_n^2} \sum_{i=1}^{N} |h_i|^2, \eta_2 = \frac{NP \sum_{i=1}^{N} |h_i|^2}{\sigma_n^2 \sum_{i=1}^{N} \left|\frac{h_i}{g_i}\right|^2}, \tag{36}$$

$$\eta_e = \frac{\left|\sum_{i=1}^{N} h_i \hat{h}_i^*\right|^2}{\sum_{i=1}^{N} |h_i|^2 \sum_{i=1}^{N} |\hat{h}_i|^2 - \left|\sum_{i=1}^{N} h_i \hat{h}_i^*\right|^2}. \tag{37}$$

If we consider the trial of the eavesdropper as an observation of the random source s, we can compute a secret-key rate of the system, $\hat{R}_k$, where $I(x;y) - I(x;\hat{v}) \leq \hat{R}_k \leq I(x;y|\hat{v})$. When the legitimate users want to generate secret-key with rate R_0, any secret-key generation process of our model with R_0 lower than $\hat{R}_k$ could be secure. In addition, if the secret-key generating rate is higher than $\hat{R}_k$, the system is no longer secure, or else, we can say in this case the eavesdropper can crack the key. Then we have the following theorem:

Theorem 2. *If the system channel gain* $\mathbf{h}$ *is statistically independent complex Gaussian random variables with the same variance* σ^2, $\hat{\boldsymbol{h}}$ *is the channel estimation vector, the legitimate users tend to generate secret-key with rate* R_0, *the average probability of one trial for the eavesdroppers to crack the key is*

$$\begin{aligned} & E\left(Pr\left\{\hat{R}_k < R_0\right\}\right) \\ = & \left(\sqrt{(\frac{\eta_1+\eta_2}{2})^2 + \frac{\eta_1\eta_2}{e^{R_0}-1}} - (\frac{\eta_1+\eta_2}{2})\right)^{1-N}. \end{aligned} \tag{38}$$

Proof. Please see Appendix B. □

Theorem 2 means when we want to achieve secret-key rate of R_0, if the SNRs at the legitimate users, η_1 or η_2, are large enough, and the number of the helpers is also large enough, the probability for the eavesdropper to crack the key can be arbitrary small.

For example, when η_1 and η_2 are both 20 dB and we want to achieve the secret-key rate of 1 nats/symbol, and there are 20 helpers. Then the probability is about 10^{-27}, this means if the eavesdropper wants to ensure 90% probability to get the proper secret-key, he have to do about 2×10^{27} independent trials, which is almost impossible to be done.

Note that Theorem 2 does not limit the probability density function of $\hat{\boldsymbol{h}}$, which means the eavesdropper can guess the channel gains in any way as he will, and the this will not affect the average probability for him to get the secret-key.

5. Numerical Results

In this section, we demonstrate the performance of the proposed schemes numerically. We perform the simulations with three types of configure: fixed channels, Rayleigh fading channels and line-of-sight (LOS) channels. In this section, all the notion of secret-key rate is actually the upper bound of the secret-key capacity.

We randomly generate some channel gains shown in Table 1, the noise power σ_n^2 is 0 dBm. We compare the R^- and R^+ in Figure 5 . We can see that R^- and R^+ are very close to each other, especially when the transmitting power is high, they are almost the same. Figure 5 shows that R^+ is a tight upper bound of the secret-key capacity, and verifies the result of Theorem 1.

Table 1. CSI list for simulation.

Number	Channel Gain
N = 4	$\mathbf{h} = \{0.045 + 1.645i; 2.416 - 0.454i; -0.309 + 1.008i; 0.187 + 2.049i\}$ $\mathbf{g} = \{0.531 + 1.371i; 2.134 + 0.245i; 0.354 + 0.118i; 0.231 + 0.384i\}$ $\mathbf{e} = \{0.807 + 0.637i; -1.027 - 0.404i; 1.294 - 0.403i; 0.014 + 0.084i\}$
N = 6	$\mathbf{h} = \{0.947 + 0.602i; -0.525 + 0.017i; -1.115 - 1.610i; -1.592 + 1.238i; 1.174 + 0.683i;$ $0.485 - 0.780i\}$ $\mathbf{g} = \{1.288 - 0.070i; -0.013 - 0.578i; -1.333 + 0.469i; -0.556 + 1.299i; 0.755 + 1.634i;$ $-0.911 - 0.702i\}$ $\mathbf{e} = \{0.218 - 0.435i; 1.713 - 0.562i; -2.078 + 0.878i; 0.112 - 0.814i; -1.086 - 0.258i;$ $-1.558 + 0.493i\}$
N = 8	$\mathbf{h} = \{-0.236 - 1.674i; -0.105 - 0.768i; 2.682 - 1.610i; -1.530 - 0.405i; -0.001 - 0.530i;$ $0.773 + 0.315i; 0.458 + 1.547i; -0.331 + 0.611i\}$ $\mathbf{g} = \{-1.213 - 0.691i; -0.274 + 0.767i; -1.414 - 1.255i; -0.713 - 0.016i; 1.081 - 1.448i;$ $-0.291 + 0.249i; 1.597 + 0.715i; 0.085 - 1.031i\}$ $\mathbf{e} = \{0.579 + 0.484i; -0.642 + 0.787i; -0.211 + 0.426i; -1.332 - 0.176i; -0.868 - 0.136i;$ $-1.469 - 0.955i; -1.562 - 0.097i; 0.127 + 0.407i\}$

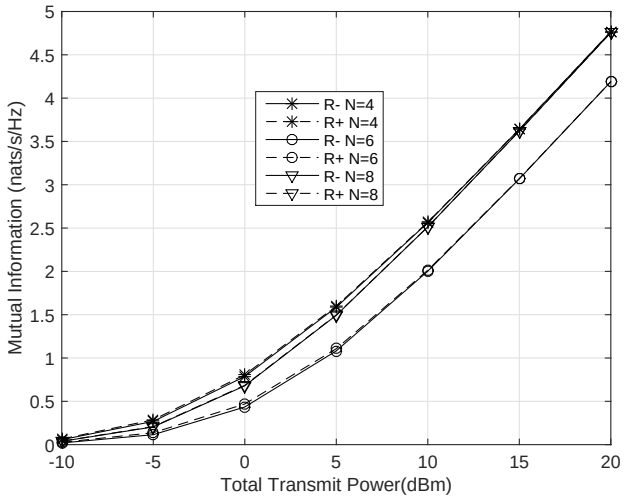

Figure 5. Upper and lower bound of secret-key capacity with the channels listed in Table 1. In the figure, R^+ and R^- are close to each other, which verifies the result of Theorem 1.

In Figure 6 we consider wireless communication system in fading environment. The channels of the users and the eavesdropper are all Rayleigh fading channels which are independent of each other, and follow the unit variance zero mean complex Gaussian distribution. The noise power σ_n^2 is also set as 0 dBm. We compare the results with different numbers of the helpers and different algorithms, 1000 times of experiments are performed to get average secret-key rate, which are shown in Figure 6. The secret-key rates increase linearly with the total transmitting power, as shown in (31), and more helpers result in better performance. Even when we have no idea about the eavesdropper, the performance of the algorithm is fairly good.

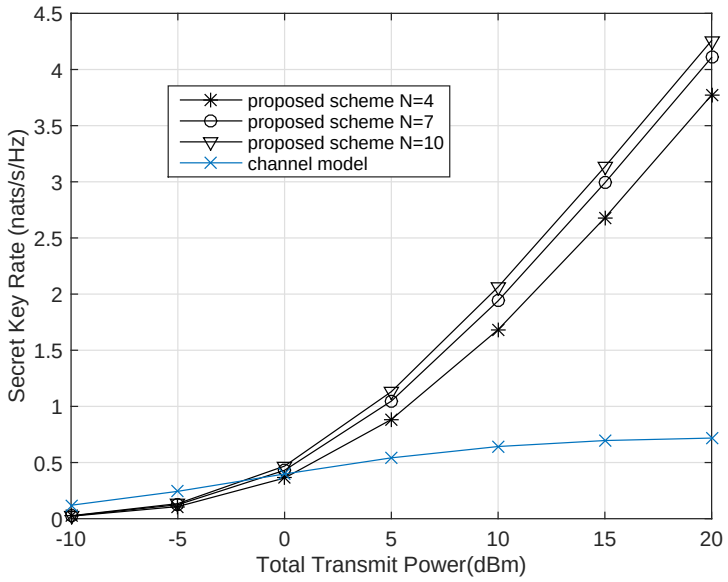

Figure 6. Average secret-key rate versus transmitting power for Rayleigh fading channels.

Figure 6 also shows the performance of a simple key generation scheme of channel model (Section 4.1 in [35]). In this model, the common randomness is from user A, that is, user A sends random signals to MC, and then the two users try to generate a secret-key from these random signals by public discussion. This is a typical channel model for secret-key generation. We implement the same transmitting power configure as the proposed schemes. It is shown in Figure 6 that the performance of channel model saturates with the increasing of the transmitting power, the secret-key rate is much lower than the proposed schemes in high SNR region.

Secret-key generation process is not a common communication process, whose transmitting rate increases with the transmitting power. Consider the lower bound of secret-key capacity, $R^- = I(x;y) - I(x;\mathbf{v})$. The first part of R^- is the mutual information of x and y, which increases with the transmitting power. In addition, the second part of R^- will also increase with the transmitting power. This means when antenna power increases, the mutual information of x and y is larger, and the leaked information to the eavesdropper is also larger. Then if the correlated random source observed by the two legitimate parties is some types of radio signals, higher sending power would not benefit the system performance much. However, the proposal in this paper performs just like a common communication system. The reason is that the leaked information to the eavesdropper is almost fixed, which is mainly determined by the correlation coefficient of x, y and $\mathbf{v}$. The eavesdropper does not get more information when the signal power is higher, that is, $I(x;\mathbf{v})$ does not increase with the signal power.

Figure 7 illustrates the secret-key rate of the system versus helper number with the same configure as in Figure 6. It is shown that more helpers result in better performance in average, but the secret-key rate increases more slowly when the helper number becomes large. The reason is when the number of helpers is large enough, the receiving signals of the eavesdropper are almost independent of the legitimate users' (as shown in Figure 4), the secret-key rate is almost saturated to its upper bound $I(x;y)$. Then more helpers benefit a little to the system performance when the helper number is large. In a practical system, more helpers cause higher system complexity, then there would be a trade-off between system complexity and secret-key rate.

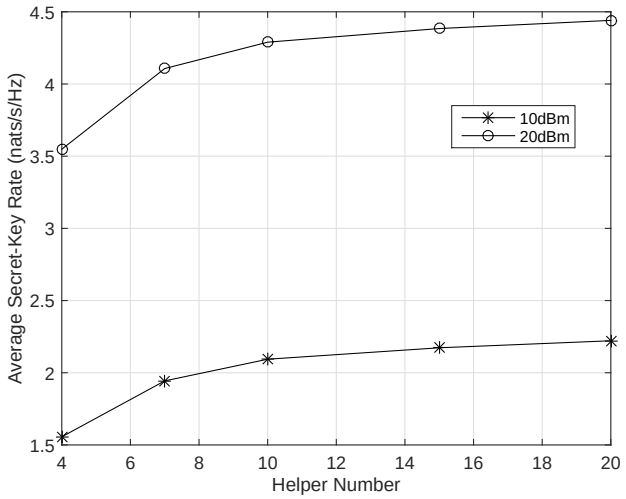

Figure 7. Average secret-key rate versus helper number for Rayleigh fading channels with total transmitting power of 10 dBm and 20 dBm.

Figure 8 compares the proposed schemes and the channel model key generation scheme in LOS channel. In the experiment, the eavesdropper moves along the horizonal line between A and MC. Channels between any two nodes are modeled by a simple line-of-sight channel model including the path loss effect and a random phase: $h = d^{-c/2}e^{j\theta}$ where d is the distance between any two nodes, $c = 3.5$ is the path loss exponent, θ is the random phase uniformly distributed within $[0, 2\pi)$. There are four helpers placed along the horizonal line, with equal spacing of 20 m. The noise power is $\sigma_n^2 = -60$ dBm. It is shown that the secret-key rate of channel model is almost zero when the eavesdropper is placed close to user A, while the proposed schemes model can still achieve fairly good performance near the two users. In this case, the channels of the system are highly correlated to each other, most of the physical layer secret-key generation schemes could not achieve fairly good performance except for the algorithm in this paper. This is ascribed to the sensitivity of the proposal.

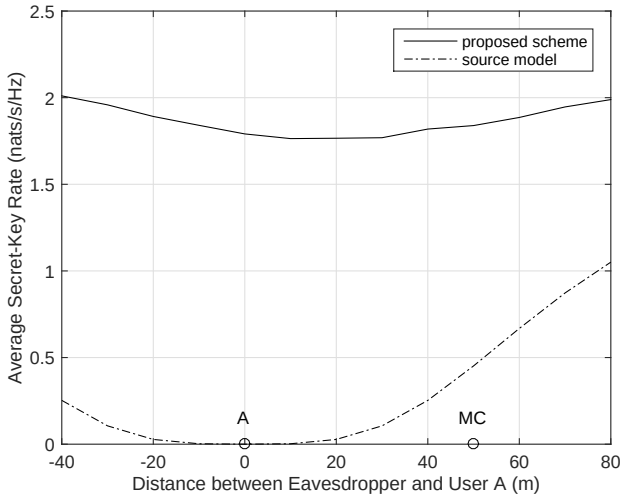

Figure 8. Average secret-key rate of LOS channel when the eavesdropper moves along the horizonal line.

We do not compare the proposed schemes with other CSI based key generation algorithms of [13–33] in numerical simulations, because the system configures are different. All the CSI based

algorithms depend on the coherence time and bandwidth of the wireless channel models, while the performance our scheme mainly depends on the symbol rate of the helpers. Typically, coherent time of a wireless communication system would be longer than 10 milliseconds, then for a narrow band system, there will be up to several thousands of bits secret-key generated by the system per second in high SNR region with small number of relays. In addition, for the scheme in this paper, hundreds of thousands of bits secret-key could be generated per second with only 10 kHz bandwidth in high SNR region.

6. Conclusions

In this paper, we have investigated the design of the correlated Gaussian sources with multiple cooperative helpers for physical layer secret-key generation for TSSN. The proposed scheme can help to update the secret-key of the wireless sensors in the system dynamically and securely.

In traditional distributed physical layer security communication systems, cooperative helpers are used as relays or interference sources. The basic idea of these algorithms is beam-forming. The main difference between the proposed schemes and the traditional algorithms is that the helpers send independent random signals in our schemes. What the helpers send is different from what the users receive, even the helpers themselves have no idea of the receiving signals of the users. This helps to create spatial differences between the legitimate users and the eavesdropper.

The proposed scheme provides an artificial random source for secret-key generation, then it is possible to get high secret-key rate by increasing the symbol rate of the helpers. Traditional CSI based secret-key generation schemes can only achieve up to several hundreds of bits secret-key per second for narrow band system, and highly depend on the coherence time of the channels. The proposed scheme, by contrast, can achieve hundreds of thousands of bits per second for narrow band system, and could possibly generate several mega bits secret-key per second for wideband system.

Note that when there are too many helpers, the synchronization is difficult for a practical system. The estimation of the channel CSI could not be accurate, the estimation errors accumulate with the increasing of the helper number. These facts will destroy the correlationship of the legitimate users, and suffer the performance of the system. How to find a balance or an optimal solution for a practical system with these issues is left for future work.

Acknowledgments: The authors would like to thank the associate editor and anonymous reviewers for their valuable comments. This work was supported by the PhD Start-up Funds of Northwest A&F University (No. Z109021613), the Fundamental Research Funds for the Central Universities, Northwest A&F University (No. 2014YB068), and the major agricultural science and technology extension service project of Shaanxi Province (Grant No.2016XXPT-00).

Author Contributions: The work presented here was carried out in collaboration between all authors. Bin Yang contributed to the experiment conduction, data handling, and paper writing; Jianfeng Zhang gave some good suggestions to the paper structure and revised the English language and grammar.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Theorem 1

Proof. Because $|h_i|^2 > 0, i = 1, 2, \ldots, N$

$$\begin{aligned} &\lim_{P\to\infty} H(x|y) \\ =&\lim_{P\to\infty} \log\left(\frac{(1+\rho)\sigma_n^2\sum_{i=1}^{N}|h_i|^2P+\sigma_n^4}{\rho\sum_{i=1}^{N}|h_i|^2P+\sigma_n^2}\right)+\log(\pi e) \\ =&\quad \log\left((1+1/\rho)\sigma_n^2\right)+\log(\pi e), \end{aligned} \tag{A1}$$

then from Equations (17) and (18) we have

$$\begin{aligned}&\lim_{P\to\infty} |\mathbf{Q}_{\mathbf{v}}|/|\mathbf{Q}_{x\mathbf{v}}| \\ =&\lim_{P\to\infty} \frac{\sum \mu_{ij}P^2+\sigma_n^2\sum\eta_i P+\sigma_n^4}{\sum\alpha_{ijk}P^3+\sigma_n^2\sum\beta_{ij}P^2+\sigma_n^4\sum\gamma_i P+\sigma_n^6}\end{aligned} \tag{A2}$$

There are at least three helpers sending signals, which means $\sum \alpha_{ijk} > 0$, then we have

$$\lim_{P\to\infty} |\mathbf{Q}_{\mathbf{v}}|/|\mathbf{Q}_{x\mathbf{v}}| = 0, \tag{A3}$$

then

$$\begin{aligned}&\lim_{P\to\infty} H(x|y\mathbf{v}) \\ =&\lim_{P\to\infty} \log\left(\frac{(1+\rho)\sigma_n^2|\mathbf{Q}_{x\mathbf{v}}|-\rho\sigma_n^4|\mathbf{Q}_{\mathbf{v}}|}{\rho|\mathbf{Q}_{x\mathbf{v}}|+(1-\rho)\sigma_n^2|\mathbf{Q}_{\mathbf{v}}|}\right)+\log(\pi e) \\ =&\log\left((1+1/\rho)\sigma_n^2\right)+\log(\pi e).\end{aligned} \tag{A4}$$

Finally from Equations (14) and (28) we know

$$\lim_{P\to\infty}\left(R^+-R^-\right)=\lim_{P\to\infty}\left(H(x|y\mathbf{v})-H(x|y)\right)=0. \tag{A5}$$

□

Appendix B. Proof of Theorem 2

Proof. Assuming that the channel estimation vector $\hat{\boldsymbol{h}}$ follows the probability density function of L($\hat{\boldsymbol{h}}$). Because $C_{sk} \le I(x;y|\hat{v})$, for the proposed algorithm we have

$$\begin{aligned}R^+ &= I(x;y|\hat{v}) \\ &= H(x|\hat{v})-H(x|y\hat{v}) \\ &= \log\frac{|\mathbf{Q}_{x\hat{v}}||\mathbf{Q}_{y\hat{v}}|}{|\mathbf{Q}_{\hat{v}}||\mathbf{Q}_{xy\hat{v}}|}.\end{aligned} \tag{A6}$$

We assume that the power of signal $s(k)$ is P_s, that is,

$$P_s \doteq E\left(ss^*\right). \tag{A7}$$

From Equations (34) and (35) we have

$$\eta_1=\frac{P_s}{\sigma_n^2},\ \eta_1=\frac{\rho P_s}{\sigma_n^2},\ \eta_e=\frac{|\delta|^2P_s}{\sigma_e^2}. \tag{A8}$$

Then we have

$$\begin{aligned}|\mathbf{Q}_{x\hat{v}}| &= \begin{vmatrix} P_s+\sigma_n^2, & \delta^*P_s \\ \delta^*P_s, & |\delta|^2P_s+\sigma_e^2 \end{vmatrix} \\ &= |\delta|^2P_s\sigma_n^2+P_s\sigma_e^2+\sigma_n^2\sigma_e^2 \\ &= \sigma_n^2\sigma_e^2\left(\eta_1+\eta_e+1\right).\end{aligned} \tag{A9}$$

Following almost the same steps, we have

$$|\mathbf{Q}_{y\hat{v}}| = \sigma_n^2\sigma_e^2\left(\eta_2+\eta_e+1\right), \tag{A10}$$
$$|\mathbf{Q}_{xy\hat{v}}| = \sigma_n^4\sigma_e^2\left(\eta_1+\eta_2+\eta_e+1\right), \tag{A11}$$
$$|\mathbf{Q}_{\hat{v}}| = \sigma_e^2\left(\eta_e+1\right). \tag{A12}$$

Then we have

$$R^+ = I(x;y|\hat{v}) = \log\frac{(\eta_1+\eta_e+1)(\eta_2+\eta_e+1)}{(\eta_e+1)(\eta_1+\eta_2+\eta_e+1)}. \tag{A13}$$

Actually, R^+ is a decreasing function of η_e. Because $\eta_e > 0$, then for given $R^+ = R_0$, η_1 and η_2, if

$$\eta_e > \sqrt{(\frac{\eta_1+\eta_2}{2})^2 + \frac{\eta_1\eta_2}{e^{R_0}-1}} - (\frac{\eta_1+\eta_2}{2}) - 1 \doteq \eta_0, \tag{A14}$$

We have $\hat{R}_k \leq R_0$. Then from Equation (37) we have

$$\frac{\left|\sum_{i=1}^N h_i\hat{h}_i^*\right|^2}{\sum_{i=1}^N|h_i|^2\sum_{i=1}^N|\hat{h}_i|^2} > \frac{\eta_0}{1+\eta_0}. \tag{A15}$$

To compute the probability of the trial, we convert all the complex numbers to real numbers. We define the real channel gain vector and the channel gain estimation vector as

$$\mathbf{x} \doteq (h_1^{(r)}, h_1^{(i)}, \ldots, h_N^{(r)}, h_N^{(i)})^T, \hat{\mathbf{x}} \doteq (\hat{h}_1^{(r)}, \hat{h}_1^{(i)}, \ldots, \hat{h}_N^{(r)}, \hat{h}_N^{(i)})^T, \tag{A16}$$

where $h_i = h_i^{(r)} + jh_i^{(i)}$, $\hat{h}_i = \hat{h}_i^{(r)} + j\hat{h}_i^{(i)}$, $i = 1 \ldots N$. Then the condition (A15) is

$$\frac{\left(\mathbf{x}^T\hat{\mathbf{x}}\right)^2 + \left(\mathbf{x}^T\mathbf{R_0}\hat{\mathbf{x}}\right)^2}{\|\mathbf{x}\|\|\hat{\mathbf{x}}\|} > \frac{\eta_0}{1+\eta_0} \tag{A17}$$

where

$$\mathbf{R_0} \doteq \begin{pmatrix} 0, & -1, & \ldots & 0, & 0 \\ 1, & 0, & \ldots & 0, & 0 \\ \vdots & & \ddots & \vdots & \\ 0, & 0, & \ldots & 0, & -1 \\ 0, & 0, & \ldots & 1, & 0 \end{pmatrix} \tag{A18}$$

It is clear that for any vector $\mathbf{x} \in \mathbb{R}^{2N}$, we have $\mathbf{x}^T\mathbf{R_0}\mathbf{x} = 0$. The probability of the eavesdropper successing to have a estimation of certain $\mathbf{h}$ is

$$\begin{aligned} P(\mathbf{x}) &= Pr\left\{\frac{\left(\mathbf{x}^T\hat{\mathbf{x}}\right)^2 + \left(\mathbf{x}^T\mathbf{R_0}\hat{\mathbf{x}}\right)^2}{\|\mathbf{x}\|\|\hat{\mathbf{x}}\|} > \frac{\eta_0}{1+\eta_0}\right\} \\ &= \underbrace{\int\cdots\int}_{\mathbf{y}} L(\mathbf{y})J(\mathbf{y},\mathbf{x})d\mathbf{y}, \end{aligned} \tag{A19}$$

where

$$J(\mathbf{y},\mathbf{x}) \doteq \begin{cases} 1, & \frac{\left(\mathbf{x}^T\mathbf{y}\right)^2+\left(\mathbf{x}^T\mathbf{R_0}\mathbf{y}\right)^2}{\|\mathbf{x}\|\|\mathbf{y}\|} > \frac{\eta_0}{1+\eta_0} \\ 0, & others \end{cases}. \tag{A20}$$

Then the average probability of one test to crack the key is

$$E\left(P(\mathbf{x})\right) = \underbrace{\int\cdots\int}_{\mathbf{y}} L(\mathbf{y}) \left(\underbrace{\int\cdots\int}_{\mathbf{x}} J(\mathbf{y},\mathbf{x})\mathcal{N}(\mathbf{x})d\mathbf{x} \right) d\mathbf{y}, \tag{A21}$$

where $\mathcal{N}(\mathbf{x}) \doteq \frac{1}{(2\pi\sigma^2)^N} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(h_i^{(r)^2} + h_i^{(i)^2}\right)\right)$. Then we have

$$\begin{aligned} &\underbrace{\int\cdots\int}_{\mathbf{x}} J(\mathbf{y},\mathbf{x})\mathcal{N}(\mathbf{x})d\mathbf{x} \\ =& \underbrace{\int\cdots\int}_{\substack{\left(\mathbf{u}^T\hat{\mathbf{u}}\right)^2+\left(\mathbf{u}^T\mathbf{R_0}\hat{\mathbf{u}}\right)^2>\frac{\eta_0}{1+\eta_0} \\ \|\mathbf{u}\|=1}} \int_0^\infty \frac{\exp\left(-\frac{r^2}{2\sigma^2}\right)}{(2\pi\sigma^2)^N} r^{2N-1}\, dr d\mathbf{u}, \end{aligned} \tag{A22}$$

where Equation (A22) is derived by transforming the axis with

$$r = \sqrt{\|\mathbf{x}\|},\ \mathbf{u} = \frac{\mathbf{x}}{\sqrt{\|\mathbf{x}\|}},\ \hat{\mathbf{u}} = \frac{\mathbf{y}}{\sqrt{\|\mathbf{y}\|}}. \tag{A23}$$

We rotate the axis with a matrix $\mathbf{A}$, which satisfies $\mathbf{A}\hat{\mathbf{u}} = \mathbf{l_0}$, $\mathbf{A}\mathbf{R_0}\mathbf{A}^T = \mathbf{R}_0^T$, $\mathbf{A}^T\mathbf{A} = \mathbf{I}_{2N}$, where $\mathbf{l_0} \doteq (1,0,\ldots,0)^T$. Then we have $\mathbf{u}^T\hat{\mathbf{u}} = v_1$, $\mathbf{u}^T\mathbf{R_0}\hat{\mathbf{u}} = -v_2$.

Because $\int_0^\infty \frac{1}{(2\pi\sigma^2)^N} \exp\left(-\frac{r^2}{2\sigma^2}\right) r^{2N-1}\, dr = \frac{(N-1)!}{2\pi^N}$, we have

$$(\text{A22}) = \underbrace{\int\cdots\int}_{\substack{v_1^2+v_2^2>\frac{\eta_0}{1+\eta_0} \\ \|\mathbf{v}\|=1}} \frac{(N-1)!}{2\pi^N} |\mathbf{A}|\, d\mathbf{v} = \frac{1}{(1+\eta_0)^{N-1}}. \tag{A24}$$

Since $\underbrace{\int\cdots\int}_{\mathbf{y}} L(\mathbf{y})d\mathbf{y} = 1$, then finally we have $E\left(P(\mathbf{x})\right) = \frac{1}{(1+\eta_0)^{N-1}}$. □

References

1. Sa, J.; Choi, Y.; Chung, Y.; Kim, H.-Y.; Park, D.; Yoon, S. Replacement condition detection of railway point machines using an electric current sensor. *Sensors* **2017**, *17*, 263.
2. Li, J.; Dridi, M.; El-Moudni, A. A cooperative traffic control of vehicle-intersection (CTCVI) for the reduction of traffic delays and fuel consumption. *Sensors* **2016**, *16*, 2175.
3. Fokm, D.T.; Frost, V.S.; Kuehnhausen, M.; DePardo, D.; Oguna, A.N.; Searl, L.S.; Komp, E.; Zeets, M.; Deavours, D.D.; Evans, J.B.; et al. An open-system transportation security sensor network: Field-trial experiences. *IEEE Trans. Veh. Technol.* **2010**, *59*, 3942–3955.
4. Jung, J.; Kim, J.; Choi Y.; Won, D. An Anonymous User Authentication and Key Agreement Scheme Based on a Symmetric Cryptosystem in Wireless Sensor Networks. *Sensors* **2016**, *16*, 1299.
5. Elgenaidi, W.; Newe, T.; O'Connell, E.; Toal, D.; Dooly, G. Secure and Efficient Key Coordination Algorithm for Line Topology Network Maintenance for Use in Maritime Wireless Sensor Networks. *Sensors* **2016**, *16*, 2204.
6. Maurer, U.M. Secret key agreement by public discussion from common information. *IEEE Trans. Inf. Theory* **1993**, *39*, 733–742.

7. Ahlswede, R.; Csiszár, I. Common randomness in information theory and cryptography part I: Secret sharing. *IEEE Trans. Inf. Theory* **1993**, *39*, 1121–1132.
8. Bloch, M.; Barros, J.; Rodrigues, M.R.D.; McLaughlin, S.W. Wireless information-theoretic security. *IEEE Trans. Inf. Theory* **2008**, *54*, 2515–2534.
9. Chou, T.; Draper, S.; Sayeed, A. Key generation using external source excitation: Capacity, reliability, and secrecy exponent. *IEEE Trans. Inf. Theory* **2012**, *58*, 2455–2474.
10. Watanabe, S.; Oohama, Y. Secret key agreement from correlated gaussian sources by rate limited public communication. *IEICE Trans. Fundam.* **2010**, *93* , 1976–1983.
11. Watanabe, S.; Oohama, Y. Secret key Agreement from Vector Gaussian Sources by Rate Limited Public Communication. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 541–550.
12. Nitinawarat, S.; Narayan, P. Secret key generation for correlated gaussian sources. *IEEE Trans. Inf. Theory* **2012**, *58*, 3373–3391.
13. Shimizu, T.; Iwai, H.; Sasaoka, H. Physical-layer secret key agreement in two-way wireless relaying systems. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 650–660.
14. Zhou, H.; Huie, L.M.; Lai, L. Secret key generation in the two-way relay channel with active attackers. *IEEE Trans. Inf. Forensics Secur.* **2014**, *9*, 476–488.
15. Chou, T.-H.; Draper, S.C.; Sayeed, A.M. Secret key generation from sparse wireless channels: Ergodic capacity and secrecy outage. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 1751–1764.
16. Wilhelm, M.; Martinovic, I.; Schmitt, J.B. Secure key generation in sensor networks based on frequency-selective channels. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 1779–1790.
17. Prabhakaran, V.M.; Eswaran, K.; Ramchandran, K. Secrecy via sources and channels. *IEEE Trans. Inf. Theory* **2012**, *58*, 6747–6765.
18. Prabhakaran, V.M.; Eswaran, K.; Ramchandran, K. Secrecy via sources and channels—A secret key-Secret message rate tradeoff region. In Proceedings of the IEEE International Symposium on Information Theory, Toronto, ON, Canada , 6–11 July 2008; pp. 1010–1014.
19. Khisti, A.; Diggavi, S.; Wornell, G. Secret key agreement using asymmetry in channel state knowledge. In Proceedings of the IEEE International Symposium on Information Theory, Seoul, Korea, 28 June–3 July 2009; pp. 2286–2290.
20. Khisti, A.; Diggavi, S.; Wornell, G. Secret-key agreement with channel state information at the transmitter. *IEEE Trans. Inf. Forens. Secur.* **2011**, *6*, 672–681.
21. Wilson, R.; Tse, D.; Scholtz, R.A. Channel identification: Secret sharing using reciprocity in ultrawideband channels. *IEEE Trans. Inf. Forens. Secur.* **2007**, *2*, 364–375.
22. Ye, C.; Mathur, S.; Reznik, A.; Shah, Y.; Trappe, W.; Mandayam, N. Information-theoretically secret key generation for fading wireless channels. *IEEE Trans. Inf. Forens. Secur.* **2010**, *5*, 240–254.
23. Wallace, J.W.; Chen, C.; Jensen, M.A. Key generation exploiting MIMO channel evolution: Algorithms and theoretical limits. In Proceedings of the 3rd European Conference on Antennas and Propagation, Berlin, Germany, 23–27 March 2009; pp. 1499–1503.
24. Patwari, N.; Croft, J.; Jana, S.; Kasera, S.K. High-rate uncorrelated bit extraction for shared secret key generation from channel measurements. *IEEE Trans. Mob. Comput.* **2010**, *9*, 17–30.
25. Aono, T.; Higuchi, K.; Ohira, T.; Komiyama, B.; Sasaoka, H. Wireless secret key generation exploiting reactance-domain scalar response of multipath fading channels. *IEEE Trans. Antennas Propag.* **2005**, *53*, 3776–3784.
26. Sayeed, A.; Perrig, A. Secure wireless communications: Secret keys through multipath. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 3013–3784.
27. Lai, L.; Liang, Y.; Du, W. Cooperative key generation in wireless networks. *IEEE J. Sel. Areas Commun.* **2012**, *30*, 1578–1588.
28. Zeng, K.; Wu, D.; Chan, A.; Mohapatra, P. Exploiting multipleantenna diversity for shared key generation in wireless networks. In Proceedings of the IEEE Conference on Computer Communications (INFOCOM), San Diego, CA, USA, 15–19 March 2010.
29. Liu, Y.; Draper, S.C.; Sayeed, A.M. Exploiting Channel Diversity in Secret Key Generation From Multipath Fading Randomness. *IEEE Trans. Inf. Forens. Secur.* **2012**, *7*, 1484–1497.

30. Zan, B.; Gruteser, M.; Hu, F. Key agreement algorithms for vehicular communication networks based on reciprocity and diversity theorems. *IEEE Trans. Veh. Tech.* **2013**, *62*, 4020–4027.
31. Primak, S.; Liu, K.; Wang, X. Secret key generation using physical channels with imperfect CSI. In Proceedings of the IEEE 80th Vehicular Technology Conference, Vancouver, BC, Canada, 14–17 September 2014; pp. 1–5.
32. Gungor, O.; Chen, F.; Koksal, C.E. Secret key generation via localization and mobility. *IEEE Trans. Veh. Tech.* **2015**, *64*, 2214–2230.
33. Wang, Q.; Xu, K.; Ren, K. Cooperative Secret Key Generation from Phase Estimation in Narrowband Fading Channels. *IEEE J. Sel. Areas Commun.* **2012**, *30*, 1666–1674.
34. Yang, B.; Wang, W.; Yin, Q. Secret key generation from multiple cooperative helpers by rate unlimited public communication. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 8183–8187.
35. Bloch, M.; Barros, J. *Physical-Layer Security: From Information Theory to Security Engineering*; Cambridge University Press: New York, NY, USA, 2011.
36. Zeng, K. Physical layer key generation in wireless networks: Challenges and opportunities. *IEEE Commun. Mag.* **2015**, *53*, 33–39.
37. Eberz, S.; Strohmeler, M.; Wilhelm, M.; Martinovie, I. A practical man-in-the-middle attack on signal-based key generation protocols. In Proceedings of the European Symposium on Research in Computer Security, Pisa, Italy, 10–14 September 2012; pp. 235–252.