SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •
Special Focus on Machine-Type Communications

October 2017, Vol. 60 100307:1–100307:14
doi: 10.1007/s11432-017-9195-3

Secret key generation based on private pilot under man-in-the-middle attack

Yu HUANG[1], Liang JIN[1]*, Na LI[2], Zhou ZHONG[1] & Xiaoming XU[1]

[1]*National Digital Switching System Engineering and Technological R&D Center, Zhengzhou* 450002*, China;*
[2]*National Engineering Laboratory for Mobile Network Technologies,*
Beijing University of Posts and Telecommunications, Beijing 100876*, China*

Received April 21, 2017; accepted July 25, 2017; published online September 1, 2017

Abstract Given the openness and invariance of public pilot, secret key generation (SKG) based on wireless channels is vulnerable to active attacks. In this paper, we explore man-in-the-middle (MITM) attacks, where the attacker acts as a transparent relay to intercept channel state information and deduce the generated keys. To prevent this type of attacks, a dynamic private pilot is generated, where legitimate nodes first consider the information authenticated between them as seed information for the private pilot, and then generate the private pilot based on this seed information according to the pilot requirements. Then, both the new seed information and secret keys are dynamicaally derived from wireless channels that are estimated with the private pilot instead of a public pilot. The proposed private pilot encrypts and authenticates wireless channels, allowing an SKG rate close to that without attackers. Analysis and simulation results show that the proposed SKG approach can effectively withstand an MITM attack.

Keywords private pilot, man-in-the-middle attack, secret key generation, physical layer security

Citation Huang Y, Jin L, Li N, et al. Secret key generation based on private pilot under man-in-the-middle attack. Sci China Inf Sci, 2017, 60(10): 100307, doi: 10.1007/s11432-017-9195-3

1 Introduction

With the development of the Internet of things, the coexistence of machine-to-machine (M2M) and human-to-human (H2H) communications will increase, meanwhile machine-type communications (MTC) will become essential in wireless communications. As one of the methods to ensure MTC security, secret key generation (SKG) based on wireless fading channels can derive a symmetric cryptographic key between two endpoints by exploiting the natural uniqueness, privacy, randomness and reciprocity of the wireless channels. SKG does not require expensive computation and has the potential to achieve information-theoretic security [1–6]. Moreover, the security of the generated key does not depend on the difficulty of computational problems but on the physical properties of the wireless fading channels, thus enabling universality. Given these advantages, SKG based on wireless channels has gained considerable attention [7, 8].

* Corresponding author (email: Liangjin@ndsc.com.cn)

Studies on SKG mainly focus on passive eavesdropping, where an eavesdropper located more than half-wave length away from legitimate users will only obtain uncorrelated channel measurements, and thus cannot acquire any information about the generated key [8]. Unfortunately, SKG is vulnerable to active attacks [9–12] because the eavesdropper not only acquires but actively transmits information to legitimate nodes and contaminates the uniqueness and privacy of the wireless channels, thus intercepting more channel state information (CSI). Kapetanovic et al. [9] presented an overview on passive eavesdropping and active attacks. The attacker sends identical pilots to legitimate receivers during the channel training phase, so that the estimated channels at the legitimate receivers are a linear combination of the legitimate and eavesdropping links [10]. Zhou and Lauren [11] achieved a nonzero SKG rate (SKGR) in a two-way relay channel under the optimal attacker strategy. The disruptive jamming attack is explored in [12], in which the attacker transmits jamming signals to disrupt the SKG process and reduce the SKGR of legitimate nodes.

The man-in-the-middle (MITM) attack is one of the most well-known active attacks in computer security, where a malicious third party secretly manipulates the communication link between two endpoints to intercept the transmitted information [13–20]. One of the earliest mention to MITM attacks is attributed to Bellovin et al. in [13], with reference to [14]. Afterwards, MITM has become a reference attack in the security community, with an increasing number of studies and reference every year. To mention a few, in Verizon's data investigation report [15, 16], researchers showed that the MITM attack is one of the most common type of security attacks. Likewise, Frankel et al. [17] described the MITM attack as one of the major threats against network security. Such publications along with previous awareness clearly show the increasing importance and widespread use of MITM attacks, which in principle are able to affect any online interaction. Aiming at the Hole 196 vulnerability, Mayank et al. [18] combined stealth MITM and wireless denial-of-service attacks to inject spoofed broadcast/multicast frames in Wi-Fi networks enabled with Wi-Fi Protected Access II. To prevent MITM attacks in satellite communications, In and Yong [19] improved the key exchange protocol by exchanging keys through cookie-based user authentication, which was a method designed by the European Telecommunication Standards Institute. In computer security, the MITM attack is a severe threat to SKG due to the openness of the wireless channels and public pilot, which gives the attacker ease of access and the capability to remain unnoticed. However, to the best of our knowledge, there is no study about MITM attacks regarding SKG.

In this paper, we first investigate the effect of the MITM attack on traditional SKG schemes by determining the SKGR under MITM transparent forwarding attacks. The SKGR is zero under this MITM attack given the openness and invariance of a public pilot. To prevent the MITM attack and ensure the security of the generated key, we derive a dynamic private pilot, where the legitimate nodes initially take the information authenticated between them as seed information of the private pilot. Then, these nodes generate a private pilot based on the seed information and the requirements for the private pilot. Finally, both the new seed information and secret key are dynamically derived from different characteristics of wireless channels estimated with the private pilot instead of a public pilot. Compared with schemes based on a public pilot, the proposed scheme cannot only encrypt wireless channels between the transmitter and receiver but also authenticate wireless channels, which makes the SKGR of the proposed scheme equal to that without attacks. Simulation results show that the proposed private pilot helps to effectively withstand MITM attacks, passive eavesdropping and impersonation attacks.

The main contributions presented in this paper are summarized as follows:

• We first investigate the threat of MITM attacks against existing SKG schemes and find that the SKGR only depends on wireless channels that do not pass through the MITM attacker, which means the attacker can intercept all the CSI passing through it when secret keys are generated with a public pilot.

• We then propose a novel method where both the secret key and private pilot are dynamically derived from wireless channels to prevent MITM attacks. The generated private pilot encrypts and authenticates the wireless channels.

• Our proposed SKG scheme results in an SKGR equal to that without attackers or eavesdroppers. Moreover, it can effectively withstand MITM attacks, passive eavesdropping, and impersonation attacks.

The remainder of this paper is organized as follows: Section 2 introduces the system model and problem

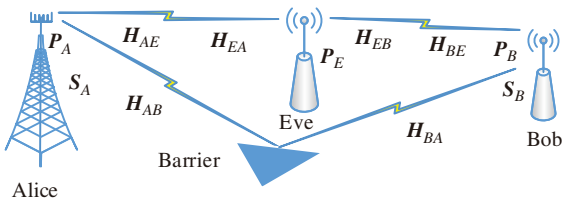

Figure 1 (Color online) System model.

description. The proposed SKG scheme based on private pilot under an MITM attack is presented in Section 3. Section 4 presents the performance analysis of the proposed scheme. The numerical results and discussion are presented in Section 5. Section 6 concludes the paper.

2 System model and problem description

2.1 System model

As shown in Figure 1, a time-division duplexing system considers a narrowband block fading point-to-point (e.g., M2M, H2H, device-to-device) wireless communication. The communication link is operated in a discrete memoryless Rayleigh flat-fading environment, where legitimate users Alice and Bob want to generate a shared secret key through the wireless channel by using reverse-link channel training and channel estimation. The communication has the presence of MITM transparent forwarding attacker Eve. Eve works in a full-duplex mode with simultaneous information reception and relaying, and wants to intercept the CSI and infer the generated secret key. Moreover, Eve always follows the transmission agreements of the system as a transparent relay without destroying the communication. Alice, Bob, and Eve have N_A, N_B, and N_E antennas, respectively. Two scenarios are explored: in scenario 1, all communication paths pass through MITM attacker Eve, whereas in scenario 2, there are some other paths available between Alice and Bob besides those via Eve. We assume that all the channels among Alice, Bob, and Eve are independent, remain in the same state in a coherence time, and change after every coherence time. In addition, at the beginning of the communication, none of the users or the attacker knows the CSI.

2.2 Problem of the existing SKG scheme

Reverse-link channel training and channel estimation is divided into 4 phases when an MITM attacker is considered. Assume that Alice transmits a pilot and Bob receives information in the first 2 phases, whereas Bob transmits a pilot and Alice receives information in the last 2 phases. As the process of the first 2 phases is almost the same as the last 2 phases, we only discuss the first 2 phases and provide the corresponding results of the last 2 phases. All the paths between Alice and Bob are equivalent to two uncorrelated paths, with one passing through Eve. In addition, Eve does not have information on the channels not passing through it.

Phase 1: For each time slot i, each Alice's antenna transmits a private pilot signal $\boldsymbol{S}_{Ai}$, $i = 1, 2, \ldots, N_A$ to Bob during N_A time slots. Hence, column i corresponds the private pilot signal transmitted in time slot i. For simplicity, we assume that $\boldsymbol{S}_A = \mathrm{diag}\{\boldsymbol{S}_{Ai}\}$. The received signals at Eve are given by

$$\boldsymbol{Y}_{AE} = \boldsymbol{H}_{AE}\sqrt{\mu}\boldsymbol{S}_A + \boldsymbol{W}_{AE}, \quad \boldsymbol{Y}_{BE} = \boldsymbol{H}_{BE}\sqrt{\mu}\boldsymbol{S}_B + \boldsymbol{W}_{BE}, \tag{1}$$

where $\mu \in [0, 1]$ denotes the power allocation factor; $\boldsymbol{H}_{AE}$ and $\boldsymbol{H}_{BE}$ are the channel gains from Alice and Bob to Eve, respectively; $\boldsymbol{W}_{AE}$ and $\boldsymbol{W}_{BE}$ are the corresponding receiver noise signals at Eve; $\boldsymbol{Y}_{AE}$ and $\boldsymbol{Y}_{BE}$ are the received signals at Eve from Alice and Bob, respectively. These subscripts are also used in the following with the same meaning.

Phase 2: For time slot i of this phase, each Eve's antennas relays the corresponding received signals in time slot i of phase 1 to Bob. The received signals at Bob and Alice are represented by

$$\begin{aligned}\boldsymbol{Y}_B &= \boldsymbol{H'}_{AB}\sqrt{1-\mu}\boldsymbol{S}_A + \alpha\boldsymbol{H}_{EB}(\boldsymbol{H}_{AE}\sqrt{\mu}\boldsymbol{S}_A + \boldsymbol{W}_{AE}) + \boldsymbol{W}_B,\\ \boldsymbol{Y}_A &= \boldsymbol{H'}_{BA}\sqrt{1-\mu}\boldsymbol{S}_B + \beta\boldsymbol{H}_{EA}(\boldsymbol{H}_{BE}\sqrt{\mu}\boldsymbol{S}_B + \boldsymbol{W}_{BE}) + \boldsymbol{W}_A,\end{aligned} \tag{2}$$

where $\alpha = \sqrt{P_E(\mu\|\boldsymbol{H}_{AE}\boldsymbol{S}_A\|_2^2 + \sigma_E^2)^{-1}}$ and $\beta = \sqrt{P_E(\mu\|\boldsymbol{H}_{BE}\boldsymbol{S}_B\|_2^2 + \sigma_E^2)^{-1}}$ are the amplification factors at Eve, $\boldsymbol{H'}_{AB}$ and $\boldsymbol{H'}_{BA}$ are the channel gains of the other paths available between Alice and Bob.

Thus, Alice and Bob can estimate the channels gains by

$$\begin{aligned}\tilde{\boldsymbol{H}}_{BA} &= \boldsymbol{Y}_A\boldsymbol{S}_B^{\mathrm{T}}\boldsymbol{P}_B^{-1} = \sqrt{(1-\mu)}\boldsymbol{H'}_{AB} + \alpha\sqrt{\mu}\boldsymbol{H}_{AB} + W_b\boldsymbol{S}_B^{\mathrm{T}}\boldsymbol{P}_B^{-1},\\ \tilde{\boldsymbol{H}}_{AB} &= \boldsymbol{Y}_B\boldsymbol{S}_A^{\mathrm{T}}\boldsymbol{P}_A^{-1} = \sqrt{(1-\mu)}\boldsymbol{H'}_{BA} + \beta\sqrt{\mu}\boldsymbol{H}_{BA} + W_a\boldsymbol{S}_A^{\mathrm{T}}\boldsymbol{P}_A^{-1},\end{aligned} \tag{3}$$

where $\boldsymbol{H}_{AB} = \boldsymbol{H}_{EB}\boldsymbol{H}_{AE}$, $\boldsymbol{H}_{BA} = \boldsymbol{H}_{EA}\boldsymbol{H}_{BE}$, $\boldsymbol{W}_b = \alpha\boldsymbol{H}_{EB}\boldsymbol{W}_{AE} + \boldsymbol{W}_B$, and $\boldsymbol{W}_a = \beta\boldsymbol{H}_{EA}\boldsymbol{W}_{BE} + \boldsymbol{W}_A$. Similarly, Eve can estimate $\boldsymbol{H}_{AE}$ and $\boldsymbol{H}_{BE}$ by

$$\begin{aligned}\tilde{\boldsymbol{H}}_{AE} &= \boldsymbol{Y}_{AE}\boldsymbol{S}_A^{\mathrm{T}}\boldsymbol{P}_A^{-1} = \sqrt{\mu}\boldsymbol{H}_{AE} + \boldsymbol{W}_{AE}\boldsymbol{S}_A^{\mathrm{T}}\boldsymbol{P}_A^{-1},\\ \tilde{\boldsymbol{H}}_{BE} &= \boldsymbol{Y}_{BE}\boldsymbol{S}_B^{\mathrm{T}}\boldsymbol{P}_B^{-1} = \sqrt{\mu}\boldsymbol{H}_{BE} + \boldsymbol{W}_{BE}\boldsymbol{S}_B^{\mathrm{T}}\boldsymbol{P}_B^{-1}.\end{aligned} \tag{4}$$

The maximum number of information bits extracted from this process is given by the mutual information (MI) of the measured channels, e.g., $I_K = I(\tilde{\boldsymbol{H}}_{AB};\tilde{\boldsymbol{H}}_{BA})$, which is called SKGR without eavesdroppers. Assuming that $\tilde{\boldsymbol{H}}_{AB}$ and $\tilde{\boldsymbol{H}}_{BA}$ are zero-mean complex Gaussian random matrices, the MI I_{AB} between Alice and Bob can be written as [21]

$$I_{k(AB)} = h(\tilde{\boldsymbol{H}}_{AB}) + h(\tilde{\boldsymbol{H}}_{BA}) - h(\tilde{\boldsymbol{H}}_{AB},\tilde{\boldsymbol{H}}_{BA}) = \log_2\frac{|\tilde{\boldsymbol{R}}_{AB,AB}||\tilde{\boldsymbol{R}}_{BA,BA}|}{|\tilde{\boldsymbol{C}}_{AB,BA}|}, \tag{5}$$

where $\tilde{\boldsymbol{R}}_{AB,AB}$ and $\tilde{\boldsymbol{R}}_{BA,BA}$ are the covariance matrices of $\tilde{\boldsymbol{H}}_{AB}$ and $\tilde{\boldsymbol{H}}_{BA}$, resectively, obtained from

$$\boldsymbol{R}_{xy} = E\{\boldsymbol{H}_x\boldsymbol{H}_y^{\mathrm{H}}\}\tilde{R}_{xy} = E\{\tilde{\boldsymbol{H}}_x\tilde{\boldsymbol{H}}_y^{\mathrm{H}}\}, \tag{6}$$

and $\tilde{\boldsymbol{C}}_{AB,AB}$ is the covariance matrix of vectors $\tilde{\boldsymbol{H}}_{AB}$ and $\tilde{\boldsymbol{H}}_{BA}$, determined by

$$\boldsymbol{C}_{xy} = E\left[\begin{pmatrix}\boldsymbol{H}_{\mathrm{x}}\\ \boldsymbol{H}_y\end{pmatrix}(\boldsymbol{H}_x^{\mathrm{H}},\boldsymbol{H}_y^{\mathrm{H}})\right]\tilde{\boldsymbol{C}}_{AB} = E\left[\begin{pmatrix}\tilde{\boldsymbol{H}}_x\\ \tilde{\boldsymbol{H}}_y\end{pmatrix}\left(\tilde{\boldsymbol{H}}_x^{\mathrm{H}}\tilde{\boldsymbol{H}}_y^{\mathrm{H}}\right)\right]. \tag{7}$$

We assume that all channel gains follow independent zero-mean Gaussian distributions; $\boldsymbol{W}_A$, $\boldsymbol{W}_B$, $\boldsymbol{W}_{AE}$, and $\boldsymbol{W}_{BE}$ are independent Gaussian noises with zero mean; and additive noises and channel gains are uncorrelated. According to the SKG agreements, the SKGR can be given by [22]

$$I_{\mathrm{SK}} = I(\tilde{\boldsymbol{H}}_{AB};\,\tilde{\boldsymbol{H}}_{BA}\Big|\,\tilde{\boldsymbol{H}}_{AE},\tilde{\boldsymbol{H}}_{BE}). \tag{8}$$

The maximum number of information bits that eavesdropper Eve can intercept is called the information leaked rate (ILR). From [21], we have

$$I_{\mathrm{ILR}} = I_{AB} - I_{\mathrm{SK}}\,. \tag{9}$$

Theorem 1. The SKGR with public pilot between Alice and Bob is zero when all paths pass through MITM attacker Eve.

Proof. This theorem corresponds to scenario 1, where all the paths between Alice and Bob pass through MITM attacker Eve, which means $\mu = 1$. According to (3), Eve can manipulate the channel measurements

by power amplification factors α and β to intercept the CSI and derive some or all of the generated secret keys. Combining (3) and (4), SKGR $I_{\rm SK}$ with public pilot between Alice and Bob can be determined by

$$\begin{aligned} I_{\rm SK} &= I(\tilde{\boldsymbol{H}}_{AB};\tilde{\boldsymbol{H}}_{BA}\Big|\,\tilde{\boldsymbol{H}}_{AE},\tilde{\boldsymbol{H}}_{BE})\\ &= I(\alpha\boldsymbol{H}_{EB}\tilde{\boldsymbol{H}}_{AE}+\boldsymbol{W}_B\boldsymbol{S}_A^{\rm T}P_A^{-1};\beta\boldsymbol{H}_{EA}\tilde{\boldsymbol{H}}_{BE}+\boldsymbol{W}_A\boldsymbol{S}_B^{\rm T}P_B^{-1}\Big|\,\tilde{\boldsymbol{H}}_{AE},\tilde{\boldsymbol{H}}_{BE})\\ &= I(\alpha\boldsymbol{H}_{EB}+\boldsymbol{W}_B\boldsymbol{S}_A^{\rm T}P_A^{-1}\tilde{\boldsymbol{H}}_{AE}^{-1};\beta\boldsymbol{H}_{EA}+\boldsymbol{W}_A\boldsymbol{S}_B^{\rm T}P_B^{-1}\tilde{\boldsymbol{H}}_{BE}^{-1}\Big|\,\tilde{\boldsymbol{H}}_{AE},\tilde{\boldsymbol{H}}_{BE})\\ &= I(\alpha^{-1}\boldsymbol{W}_B\boldsymbol{S}_A^{\rm T}P_A^{-1}\tilde{\boldsymbol{H}}_{AE}^{-1}-\boldsymbol{W}_{AE}\boldsymbol{S}_B^{\rm T}P_B^{-1};\beta^{-1}\boldsymbol{W}_A\boldsymbol{S}_B^{\rm T}P_B^{-1}\tilde{\boldsymbol{H}}_{BE}^{-1}-\boldsymbol{W}_{BE}\boldsymbol{S}_A^{\rm T}P_A^{-1}). \end{aligned} \tag{10}$$

Let

$$\boldsymbol{A}=\alpha^{-1}\boldsymbol{W}_B\boldsymbol{S}_A^{\rm T}P_A^{-1}\tilde{\boldsymbol{H}}_{AE}^{-1}-\boldsymbol{W}_{AE}\boldsymbol{S}_B^{\rm T}P_B^{-1},\quad \boldsymbol{B}=\beta^{-1}\boldsymbol{W}_A\boldsymbol{S}_B^{\rm T}P_B^{-1}\tilde{\boldsymbol{H}}_{BE}^{-1}-\boldsymbol{W}_{BE}\boldsymbol{S}_A^{\rm T}P_A^{-1}.$$

Given that parameters $\boldsymbol{S}_A^{\rm T}P_A^{-1}\tilde{\boldsymbol{H}}_{AE}^{-1}$, $\boldsymbol{S}_B^{\rm T}P_B^{-1}$, $\boldsymbol{S}_B^{\rm T}P_B^{-1}\tilde{\boldsymbol{H}}_{BE}^{-1}$, and $\boldsymbol{S}_A^{\rm T}P_A^{-1}$ are known by Alice and Bob, $\boldsymbol{W}_A$, $\boldsymbol{W}_B$, $\boldsymbol{W}_{AE}$, and $\boldsymbol{W}_{BE}$ are zero-mean independent Gaussian noises, and according to (7), covariance matrix $\boldsymbol{C}_{AB}$ of vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, is given as

$$\boldsymbol{C}_{AB}=E\{(\boldsymbol{A},\boldsymbol{B})^{\rm H}(\boldsymbol{A},\boldsymbol{B})\}=\begin{bmatrix}\boldsymbol{R}_{AA}&0\\0&\boldsymbol{R}_{BB}\end{bmatrix}. \tag{11}$$

Combining (5), (6), and (7), SKGR $I_{\rm SK}$ can be written as

$$I_{\rm SK}=I(\tilde{\boldsymbol{H}}_{AB};\tilde{\boldsymbol{H}}_{BA}\Big|\,\tilde{\boldsymbol{H}}_{AE},\tilde{\boldsymbol{H}}_{BE})=\log_2\frac{|\boldsymbol{R}_{AA}|\,|\boldsymbol{R}_{BB}|}{|\boldsymbol{C}_{AB}|}=\log_2\frac{|\boldsymbol{R}_{AA}|\,|\boldsymbol{R}_{BB}|}{|\boldsymbol{R}_{AA}|\,|\boldsymbol{R}_{BB}|}=0. \tag{12}$$

The results of Theorem 1 provide the following interpretation. When Alice and Bob undergo an MITM attack and have no other commnication paths available between them, the SKGR with public pilot is zero, which means that MITM attacker Eve can intercept all the channel information during the reverse-link channel training.

Theorem 2. When there are other paths available between Alice and Bob besides the paths via MITM attacker Eve, the SKGR with public pilot is given by

$$\begin{aligned} I_{\rm SK} &= I(\tilde{\boldsymbol{H}}_{AB};\tilde{\boldsymbol{H}}_{BA}\Big|\,\tilde{\boldsymbol{H}}_{AE},\tilde{\boldsymbol{H}}_{BE})=I(\tilde{\boldsymbol{H}'}_{AB};\tilde{\boldsymbol{H}'}_{BA})\\ &= -\sum_{m=1}^{N_A}\log_2\left(1-\frac{(1-\mu)b_m}{(1-\mu)b_m+\sigma_B^2P_B^{-1}}\frac{(1-\mu)b_m}{(1-\mu)b_m+\sigma_A^2P_A^{-1}}\right), \end{aligned} \tag{13}$$

where $b_m=\sum_{n=1}^{N_B}|\boldsymbol{H}_{AmBn}|^2$.

Proof. This theorem corresponds to scenario 2, where there are other paths available between Alice and Bob besides those via Eve. When $\mu=1$, scenario 2 is equivalent to scenario 1, thus the latter is a special case of the former. Theorem 1 tells us that Eve can intercept all channel information, whereas it has no information about other paths available between Alice and Bob. Thus, the SKGR is only achieved from these other paths.

According to (3), we have

$$\tilde{\boldsymbol{H}}_{AB}=\boldsymbol{Y}_B\boldsymbol{S}_B^{\rm T}P_B^{-1}=\sqrt{1-\mu}\boldsymbol{H'}_{BA}+\boldsymbol{W}_{\rm B}\boldsymbol{S}_B^{\rm T}P_B^{-1}=\tilde{\boldsymbol{H}'}_{AB}, \tag{14}$$

$$\tilde{\boldsymbol{H}}_{BA}=\boldsymbol{Y}_A\boldsymbol{S}_A^{\rm T}P_A^{-1}=\sqrt{1-\mu}\boldsymbol{H'}_{AB}+\boldsymbol{W}_{\rm A}\boldsymbol{S}_A^{\rm T}P_A^{-1}=\tilde{\boldsymbol{H}'}_{BA}, \tag{15}$$

considering (5), and we can achieve (13).

Theorem 2 shows that the SKGR is only obtained from the CSI that does not pass through MITM attacker Eve.

Table 1 The proposed SKG scheme based on private pilot under MITM attack

Scheme of the proposed SKG
Step 1: Initialization:
The legitimate users initialize the seed sequence of the private pilot with initial authentication key $\boldsymbol{X}^K$.
The legitimate users generate private pilot $\boldsymbol{S}^K$ from seed sequence $\boldsymbol{X}^K$ according to the channel estimation algorithm and the pilot characteristics.
Step 2: Secret key generation:
The legitimate users measure wireless channels $\boldsymbol{H}_K$ with private pilot $\boldsymbol{S}^K$, and obtain channel estimation values $\tilde{\boldsymbol{H}}_K$.
The legitimate users generate secret key $\boldsymbol{K}_K$ based on channel estimation values $\tilde{\boldsymbol{H}}_K$.
Step 3: New private pilot generation:
The legitimate users generate new seed sequence $\boldsymbol{X}^{K+1}$ from different characteristics of channel estimation values $\tilde{\boldsymbol{H}}_K$.
The legitimate users generate private pilot $\boldsymbol{S}^{K+1} = \sqrt{E}(\boldsymbol{S_\Sigma}^K / \left\|\boldsymbol{S_\Sigma}^K\right\|)$ according to seed sequence $\boldsymbol{X}^{K+1}$ and the condition of the proposed private pilot generation $\min\limits_{\mathrm{tr}\{\boldsymbol{S}^{K+1}(\boldsymbol{S}^{K+1})^{\mathrm{H}}\}\leqslant E} \left\|\boldsymbol{S}^{K+1} - \boldsymbol{S}_\Sigma^K\right\|_F$.
The legitimate users take new private pilot $\boldsymbol{S}^{K+1}$ to replace private pilot $\boldsymbol{S}^K$ from step 1.
Step 4: Repeat step 2 and step 3.

3 Proposed SKG scheme based on private pilot under MITM attack

3.1 Generation of dynamic private pilot

We assume wireless channels $\boldsymbol{H} \in \mathbb{C}^{N_B \times N_A}$ (or $\boldsymbol{H} \in \mathbb{C}^{N_A \times N_B}$) are zero means with covariance matrix $\boldsymbol{R}$, and the noises at each receiver have zero mean with covariance matrix $\boldsymbol{P}$. Our main goal is to design private pilot $\boldsymbol{S}$ to achieve an accurate estimation of channels $\boldsymbol{H}$. To accomplish this goal, we consider the minimization of the mean-square error (MSE) for the channel estimation. When the minimum mean-square error (MMSE) algorithm is used to estimate the instantaneous channel coefficients, the MMSE of the channel estimate can be formulated as [23, 24]

$$\mathrm{MSE} = \mathrm{tr}\left\{(\boldsymbol{R}^{-1} + (\boldsymbol{S} \otimes \boldsymbol{I}_{N_B})^{\mathrm{H}} \boldsymbol{P}^{-1} (\boldsymbol{S} \otimes \boldsymbol{I}_{N_B}))^{-1}\right\} \quad \text{subject to} \quad \mathrm{tr}(\boldsymbol{S}\boldsymbol{S}^{\mathrm{H}}) \leqslant E, \tag{16}$$

where E denotes the total transmitting power of $\boldsymbol{S}$.

According to Mojtaba et al. [25], we are able to achieve the MMSE when private pilot $\boldsymbol{S}$ satisfies the following condition:

$$\min_{\mathrm{tr}\{S^{K+1}(S^{K+1})^{\mathrm{H}}\}\leqslant E} \left\|\boldsymbol{S}^{(K+1)} - \boldsymbol{S}_\Sigma^K\right\|_F, \tag{17}$$

where $\boldsymbol{S}_\Sigma^K(m,n) = \sum_{l=1}^{N_B} S'_{m,n}(l)$ for all $m \in \{0, 1, \ldots, L-1\}$, $n \in \{0, 1, \ldots, N_A - 1\}$, which is achieved by optimizing private pilot $\boldsymbol{S}^K$. Given that the initialization uses the authenticated information between Alice and Bob, it remains private to third parties.

To obtain a specific solution for the above problem, private pilot $\boldsymbol{S}^{K+1}$ can be obtained from

$$\boldsymbol{S}^{K+1} = \sqrt{E}(\boldsymbol{S}_\Sigma^K / \left\|\boldsymbol{S}_\Sigma^K\right\|). \tag{18}$$

3.2 Proposed SKG scheme based on private pilot

The proposed SKG scheme based on a private pilot under an MITM attack is described in Table 1.

4 Performance analysis of the proposed SKG scheme

In this section, we discuss the performance and advantages of the proposed SKG scheme under an MITM transparent forwarding attack for the two above mentioned scenarios.

4.1 Scenario 1: all paths pass through MITM attacker

We first explore scenario 1, where all paths between Alice and Bob pass through MITM attacker Eve, which means that $\mu = 1$.

Theorem 3. The SKGR with private pilot when all paths between Alice and Bob pass through MITM attacker Eve is given as

$$I_{\mathrm{SK}} = I_{AB} - I_{\mathrm{ILR}} = -\sum_{m=1}^{N_A}\left[\log_2\left(1-\frac{\alpha^2 a_m}{\alpha^2 a_m + P_A^{-1}(\alpha^2 c+d)}\frac{\beta^2 a_m}{\beta^2 a_m + P_B^{-1}(\beta^2 f_m + e_m)}\right) - \log_2\left(1-\frac{gP_B}{gP_B+\sigma_E^2}\frac{g\beta^2}{g\beta^2+\sigma_B^2(q_m P_A+\sigma_E^2)^{-1}}\right)\right], \tag{19}$$

where

$$a_m = \sum_{n=1}^{N_B}\sum_{i=1}^{N_E}(|\boldsymbol{H}_{AmE}|^2|\boldsymbol{H}_{BnE_i}|^2), \quad c = \sum_{n=1}^{N_B}\sum_{l=1}^{L}\sum_{i=1}^{N_E}(|\boldsymbol{H}_{E_iB}|^2\sigma_{E_{i,l}}^2), \quad d = \sum_{n=1}^{N_B}\sum_{l=1}^{L}\sigma_{Bn,l}^2,$$
$$e_m = \sum_{l=1}^{L}\sigma_{Am,l}^2, \quad f_m = \sum_{l=1}^{L}\sum_{i=1}^{N_E}(|\boldsymbol{H}_{E_iAm}|^2\sigma_{E_{i,l}}^2), \quad g = \sum_{n=1}^{N_B}\sum_{i=1}^{N_E}|\boldsymbol{H}_{BnE_i}|^2, \quad q_m = \sum_{i=1}^{N_E}|\boldsymbol{H}_{AmEi}|^2.$$

Proof. When Alice and Bob generate a secret key with the proposed scheme, MITM attacker Eve can only acquire information $\boldsymbol{Y}_{AE}$ and $\boldsymbol{Y}_{BE}$, and cannot estimate channel $\boldsymbol{H}_{AE}$ and $\boldsymbol{H}_{BE}$ because of the private pilot security. Given that channels $\tilde{\boldsymbol{H}}_{AB}$ and $\tilde{\boldsymbol{H}}_{BA}$ are multivariate Gaussian and considering (5), the results from [26–28] can be used to write MI $\boldsymbol{I}_{AB}$ between $\tilde{\boldsymbol{H}}_{AB}$ and $\tilde{\boldsymbol{H}}_{BA}$ as

$$I_{AB} = I(\tilde{\boldsymbol{H}}_{AB};\tilde{\boldsymbol{H}}_{BA}) = \log_2\frac{|\tilde{\boldsymbol{R}}_{AB,AB}||\tilde{\boldsymbol{R}}_{BA,BA}|}{|\tilde{\boldsymbol{C}}_{AB,BA}|}. \tag{20}$$

From (7), covariance matrix $\tilde{\boldsymbol{C}}_{AB,BA}$ is given by

$$\tilde{\boldsymbol{C}}_{AB,BA} = E\left\{[\tilde{\boldsymbol{H}}_{AB},\tilde{\boldsymbol{H}}_{BA}]^{\mathrm{H}}[\tilde{\boldsymbol{H}}_{AB},\tilde{\boldsymbol{H}}_{BA}]\right\} = \begin{bmatrix}\boldsymbol{R}_{AB,AB} & \boldsymbol{R}_{AB,BA}\\ \boldsymbol{R}_{BA,AB} & \boldsymbol{R}_{BA,BA}\end{bmatrix}. \tag{21}$$

According to (20) and (21), MI I_{AB} can be written as

$$I_{AB} = -\log_2\left(\left|I - \boldsymbol{R}_{\hat{\boldsymbol{H}}_{BA},\hat{\boldsymbol{H}}_{AB}}\boldsymbol{R}^{-1}_{\hat{\boldsymbol{H}}_{AB},\hat{\boldsymbol{H}}_{AB}}\boldsymbol{R}_{\hat{\boldsymbol{H}}_{AB},\hat{\boldsymbol{H}}_{BA}}\boldsymbol{R}^{-1}_{\hat{\boldsymbol{H}}_{BA},\hat{\boldsymbol{H}}_{BA}}\right|\right). \tag{22}$$

Next, according to (3) and (6), channel $\tilde{\boldsymbol{H}}_{AB}$ and $\tilde{\boldsymbol{R}}_{AB,AB}$ can be written as

$$\begin{aligned}\tilde{\boldsymbol{H}}_{AB} &= \alpha\boldsymbol{H}_{EB}\boldsymbol{H}_{AE} + (\alpha\boldsymbol{H}_{EB}n_E + n_B)\boldsymbol{S}_A^{\mathrm{T}}P_{\mathrm{A}}^{-1}\\ &= \begin{bmatrix}\alpha\sum_{n=1}^{N_B}\sum_{i=1}^{N_E}\boldsymbol{H}_{E_iBn}\boldsymbol{H}_{A1E_i} + P_A^{-1}\sum_{l=1}^{L}\sum_{n=1}^{N_B}(\alpha\sum_{i=1}^{N_E}\boldsymbol{H}_{E_iBn}n_{E_{i,l}} + n_{Bn,l})\boldsymbol{S}_{A1,l}\\ \alpha\sum_{n=1}^{N_B}\sum_{i=1}^{N_E}\boldsymbol{H}_{E_iBn}\boldsymbol{H}_{A2E_i} + P_A^{-1}\sum_{l=1}^{L}\sum_{n=1}^{N_B}(\alpha\sum_{i=1}^{N_E}\boldsymbol{H}_{E_iBn}n_{E_{i,l}} + n_{Bn,l})\boldsymbol{S}_{A2,l}\\ \vdots\\ \alpha\sum_{n=1}^{N_B}\sum_{i=1}^{N_E}\boldsymbol{H}_{E_iBn}\boldsymbol{H}_{AN_AE_i} + P_A^{-1}\sum_{l=1}^{L}\sum_{n=1}^{N_B}(\alpha\sum_{i=1}^{N_E}\boldsymbol{H}_{E_iBn}n_{E_{i,l}} + n_{Bn,l})\boldsymbol{S}_{AN_A,l}\end{bmatrix}^{\mathrm{T}},\end{aligned} \tag{23}$$

$$\begin{aligned}\boldsymbol{R}_{AB,AB} &= E\{\tilde{\boldsymbol{H}}_{AB}\tilde{\boldsymbol{H}}_{AB}^{\mathrm{H}}\}\\ &= \begin{bmatrix}\alpha^2\sum_{i=1}^{N_E}\|\boldsymbol{H}_{E_iB}\|^2\|\boldsymbol{H}_{A1E_i}\|^2 & & \\ +P_A^{-1}\sum_{l=1}^{L}(\alpha^2\sum_{i=1}^{N_E}\|\boldsymbol{H}_{E_iB}\|^2\sigma_{E_{i,l}}^2+\sigma_{B,l}^2) & \cdots & 0\\ 0 & \cdots & 0\\ \vdots & \cdots & \vdots\\ 0 & \cdots & \alpha^2\sum_{i=1}^{N_E}\|\boldsymbol{H}_{E_iB}\|^2\|\boldsymbol{H}_{AN_AE_i}\|^2\\ & & +P_A^{-1}\sum_{l=1}^{L}(\alpha^2\sum_{i=1}^{N_E}\|\boldsymbol{H}_{E_iB}\|^2\sigma_{E_{i,l}}^2+\sigma_{B,l}^2)\end{bmatrix}.\end{aligned} \tag{24}$$

We can obtain $\tilde{\boldsymbol{R}}_{BA,BA}$, $\tilde{\boldsymbol{R}}_{BA,AB}$, and $\tilde{\boldsymbol{R}}_{AB,BA}$ in an analogous way. Hence, MI I_{AB} can be written as

$$\begin{aligned} I_{AB} &= \log_2 \frac{|\tilde{\boldsymbol{R}}_{AB,AB}||\tilde{\boldsymbol{R}}_{BA,BA}|}{|\hat{\mathbf{C}}_{AB,BA}|} = -\log_2 \left(\left| I - \boldsymbol{R}_{BA,AB}\boldsymbol{R}_{AB,AB}^{-1}\boldsymbol{R}_{AB,BA}\boldsymbol{R}_{BA,BA}^{-1} \right| \right) \\ &= -\sum_{m=1}^{N_A} \log_2 \left(1 - \frac{\alpha^2 a_m}{\alpha^2 a_m + P_A^{-1}(\alpha^2 c + d)} \frac{\beta^2 a_m}{\beta^2 a_m + P_B^{-1}(\beta^2 f_m + e_m)} \right). \end{aligned} \tag{25}$$

The ILR that MITM attacker Eve can intercept from wireless channels $\boldsymbol{H}$ and estimated by Alice and Bob using the proposed private pilot is given by [29]

$$\begin{aligned} I_{\mathrm{ILR}} &= I(\boldsymbol{Y}_{AE}, \boldsymbol{Y}_{BE}; \tilde{\boldsymbol{H}}_{AB}) = I(\boldsymbol{Y}_{AE}; \tilde{\boldsymbol{H}}_{AB}) + I(\boldsymbol{Y}_{BE}; \tilde{\boldsymbol{H}}_{AB} \Big| \boldsymbol{Y}_{AE}) = I(\boldsymbol{Y}_{BE}; \tilde{\boldsymbol{H}}_{AB} \Big| \boldsymbol{Y}_{AE}) \\ &= I(\boldsymbol{H}_{BE}\boldsymbol{S}_B + \boldsymbol{W}_{BE}; \beta \boldsymbol{H}_{EB}\boldsymbol{S}_A^{\mathrm{T}} P_A^{-1} + \boldsymbol{W}_B \boldsymbol{Y}_{AE}^{-1} \boldsymbol{S}_A^{\mathrm{T}} P_A^{-1}) \\ &= -\sum_{m=1}^{N_A} \log_2 \left(1 - \frac{gP_B}{gP_B + \sigma_E^2} \frac{g\beta^2}{g\beta^2 + \sigma_B^2 (q_m P_A + \sigma_E^2)^{-1}} \right). \end{aligned} \tag{26}$$

Thus, according to (25) and (26), we can obtain (19).

4.2 Scenario 2: other paths are available besides those passing through MITM attacker

We investigate scenario 2 in this subsection, where other paths are available besides those passing through Eve, and we assume that Eve has no CSI of the other available paths.

Theorem 4. The SKGR with private pilot, when there are other paths available besides those via MITM attacker Eve is given by

$$\begin{aligned} I_{\mathrm{SK}} &= I_{AB} - I_{\mathrm{ILR}} \\ &= -\sum_{m=1}^{N_A} \left[\log_2 \left(1 - \frac{\alpha^2 \mu a_m + (1-\mu) b_m}{\alpha^2 \mu a_m + (1-\mu) b_m + P_A^{-1}(\alpha^2 \mu c + d)} \frac{\beta^2 a_m + (1-\mu) b_m}{\beta^2 \mu a_m + (1-\mu) b_m + P_B^{-1}(\beta^2 \mu f_m + e_m)} \right) \right. \\ &\quad \left. -\log_2 \left(1 - \frac{gP_B}{gP_B + \sigma_E^2} \frac{g\beta^2 \mu}{g\beta^2 \mu + \sigma_B^2 (q_m P_A + \sigma_E^2)^{-1}} \right) \right], \end{aligned} \tag{27}$$

where

$$\begin{aligned} &a_m = \sum_{n=1}^{N_B} \sum_{i=1}^{N_E} (|\boldsymbol{H}_{AmE}|^2 |\boldsymbol{H}_{BnE_i}|^2), \quad b_m = \sum_{n=1}^{N_B} |\boldsymbol{H}_{AmBn}|^2, \quad c = \sum_{n=1}^{N_B} \sum_{l=1}^{L} \sum_{i=1}^{N_E} (|\boldsymbol{H}_{E_iB}|^2 \sigma_{E_{i,l}}^2), \\ &d = \sum_{n=1}^{N_B} \sum_{l=1}^{L} \sigma_{Bn,l}^2, \quad e_m = \sum_{l=1}^{L} \sigma_{Am,l}^2, \quad f_m = \sum_{l=1}^{L} \sum_{i=1}^{N_E} (|\boldsymbol{H}_{E_iAm}|^2 \sigma_{E_{i,l}}^2), \\ &g = \sum_{n=1}^{N_B} \sum_{i=1}^{N_E} |\boldsymbol{H}_{BnE_i}|^2, \quad q_m = \sum_{i=1}^{N_E} |\boldsymbol{H}_{AmEi}|^2. \end{aligned}$$

Proof. The proof follows arguments similar to those in the proof of Theorem 3 for scenario 1.

Theorem 4 is equivalent to Theorem 3 when $\mu = 1$, which indicates that scenario 1 is a special case of scenario 2. In addition, Theorem 4 is equivalent to Theorem 2 when Eve intercepts all channel information.

Compared with the traditional SKG scheme, the proposed SKG scheme based on private pilot employs the wireless channels to dynamically generate seed information and secret keys, which can ensure the security of the CSI and the generated secret key under an MITM attack. However, our scheme must first acquire the authenticated information between the transmitter and receiver by steps such as pre-installation, dynamic generation, and secure transmission, and depends on the algorithm for private pilot generation. The computational complexity of the traditional SKG scheme is mainly given by the channel

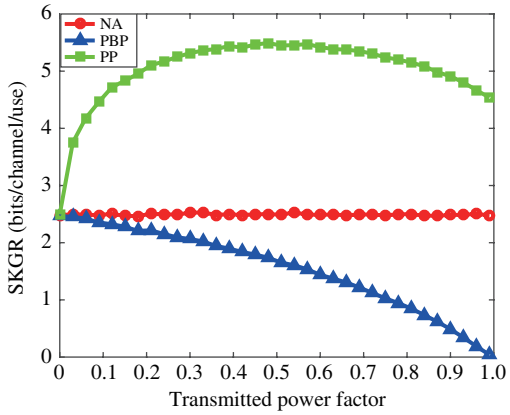

Figure 2 (Color online) SKGR according to power allocation factor for the transmitter.

estimation as $\boldsymbol{O}((N_A \times N_B \times L)^2)$, where L is the length of the private pilot, when the channels are estimated with MMSE algorithm.

In our scheme, the computational complexity of the SKG is $2 \times \boldsymbol{O}((N_A \times N_B \times L)^2)$, which consists the private pilot generation and the channel estimation. As mentioned in [25], the process for the private pilot generation can be computed as MMSE $= \sum_{k=1}^{L} \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \left| \boldsymbol{S}^{K+1}(k,i) - \boldsymbol{S}_{\Sigma}^{K}(j) \right|^2$. Thus, the computational complexity of the private pilot generation from the seed information is $\boldsymbol{O}((N_A \times N_B \times L)^2)$, which is the tradeoff to obtain a higher security level.

5 Numerical results and discussion

To verify the performance and suitability of the proposed SKG scheme, we compared the SKGR of the proposed scheme based on private pilot with the method described in [7], which is based on a public pilot and consider different eavesdropping models using Monte Carlo simulations. We used the MATLAB information theoretical estimators toolbox developed by Szabo to evaluate the Shannon MI [30]. We used the signal-noise-ratio (SNR) to assess the signal quality of the wireless communication, and the SKGR to determine the security of the SKG scheme. We evaluated the effect of both the private and public pilots on the SKGR according to the variation of the SNR. We assumed that all the channels were Rayleigh fading channels and the entries obeyed a Gaussian distribution with zero mean and unit variance. In addition, we considered the receiver noises as white Gaussian noises with zero mean and unit variance, and that Eve could estimate channels $\tilde{\boldsymbol{H}}_{AE}$ and $\tilde{\boldsymbol{H}}_{BE}$ during the SKG with the public pilot.

5.1 MITM transparent forwarding attack

We first discuss the effect of power allocation factor μ on the SKGR. Then, we investigate the following four cases considering the best μ value for the SKGR: 1) $N_A = N_B = 1$, and $N_E = 1$ or $N_E = 2$, which corresponds to a SISO case; 2) $N_A = 4$, $N_B = 1$, and $N_E = 1$ or $N_E = 2$, which corresponds to a MISO case, and where the above two cases consider either a single or multi-antenna MITM attack; 3) $N_A = 4$, $N_B = 2$, and $N_E = 2$, which corresponds a MIMO case under a multi-antenna MITM attack; and 4) $N_A = 4$, $N_B = N_E = 1$, where we also fixed the legitimate users SNR to 10 dB, and investigated the variation of SKGR according to the SNR at Eve. These four cases included our investigated two scenarios, i.e., with no paths available (NP) between Alice and Bob avoiding Eve, and with other paths available (OP) between Alice and Bob besides those passing through Eve.

All paths were divided into two equivalent paths: one path passing and the other not passing through Eve. We assumed that power allocation factor μ correspond to the path passing through Eve, whereas the other path has a power allocation factor of $1 - \mu$. The variation of the SKGR according to μ under an SNR for legitimate nodes 10 dB is shown in Figure 2. The results show that the SKGR based on the public pilot (PBP) decreases as μ increasing, reaching an SKGR of 0 when $\mu = 1$. This results is

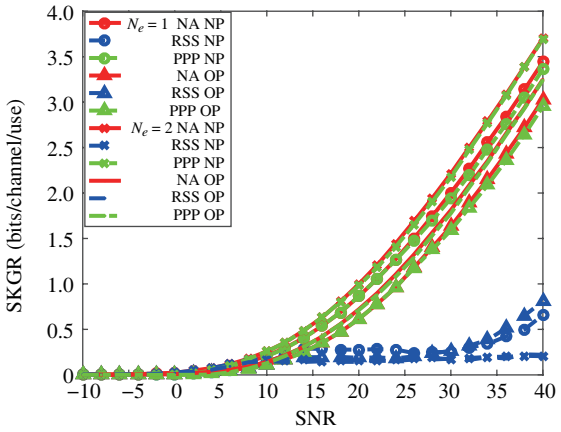

Figure 3 SKGR according to SNR under MITM attack with $N_A = N_B = 1$, $N_E = 1$ or 2.

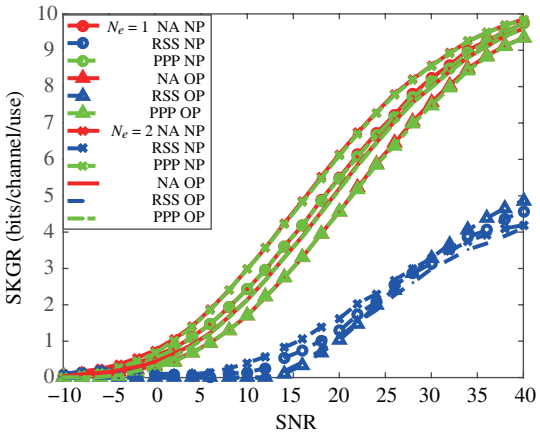

Figure 4 SKGR according to SNR under MITM attack with $N_A = 4$, $N_B = 1$, $N_E = 1$ or 2.

consistent with Theorem 1 that claims a zero SKGR when all paths pass through Eve. When $\mu = 0$, the SKGR is equal to that with no attacker or no path passing through an attacker (NA), which agrees with Theorem 2. The SKGR based on the private pilot (PP) first rises and then decreases, reaching its peak at $\mu = 0.5$. This result indicates that a maximum SKGR can be obtained when the transmitter equally distributes the transmitting power between the paths.

Theorem 1 claims that Eve can intercept all the CSI passing through it. In contrast, if the channel gains of the other paths are above those of the paths passing through Eve, it will be more difficult for Eve to obtain all CSI. Thus, for subsequent simulations, we only considered the users whose available paths were very weak or had no other path available. We assumed that the rate of the channel gains between the available paths and those passing through Eve is 0.2. In addition, the power allocation factor of the transmitter was $\mu = 0.5$ and the power amplification factor of Eve was $\alpha = \beta$.

We explored the SKGR for three cases: 1) no attacker or eavesdropper (NA); 2) the SKG method based on MIMO received signal spaces (RSS) under an MITM attack [7]; and 3) our proposed SKG method based on private pilot (PPP) under an MITM attack.

In the first scenario, when $N_A = N_B = 1$, $N_E = 1$ or 2, we examined the changes of SKGR according to the SNR of the legitimate nodes, which are shown in Figure 3. First, all the SKGR curves in the two scenarios rise as the SNR increase. Second, the SKGR values when all paths pass through Eve are larger than those when there are other available paths, because of the allocated power for legitimate nodes and the amplifying-and-forwarding of Eve. Next, the SKGR of the proposed scheme is close to that with NA, which demonstrates that the proposed SKG scheme can ensure the security of the SKG in the two scenarios. Moreover, the SKGR of our proposed approach is much larger than that of the RSS method, which shows that our approach outperforms the RSS method described in [7]. Furthermore, the SKGR obtained from the RSS method is small in the two scenarios because Eve can estimate the channels between it and Bob, and receive the signals transmitted by Alice, hence receiving the signals from Bob. Thus, Eve can intercept most of the generated secret keys according to the SKG agreements, which proves that the MITM attack is a serious threat to the SKG based on wireless channels. Finally, the SKGR obtained from the RSS method is smaller when $N_E = 2$ than when $N_E = 1$, which suggests that Eve can improve its eavesdropping performance with the increase on the number of its antennas.

In the second scenario, when $N_A = 4$, $N_B = 1$, and $N_E = 1$ or 2, Figure 4 shows that the SKGR rises with increasing of SNR. Compared with Figure 3, we find that the SKGR rises when the number of antennas increasing for the legitimate nodes increasing, which shows that a multi-antenna configuration can enhance the wireless channels and increase the channel gains. In addition, the SKGR from the RSS method shows a substantial improvement compared to that shown in Figure 3, which suggests that the multi-antenna configuration provides anti-eavesdropping features.

The SKGR results for the third scenario, when $N_A = 4$, and $N_B = N_E = 2$, are shown in Figure 5.

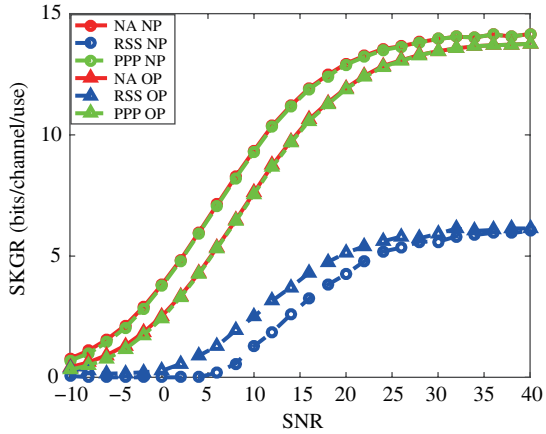

Figure 5 (Color online) SKGR according to SNR under MITM attack with $N_A = 4$, $N_B = N_E = 2$.

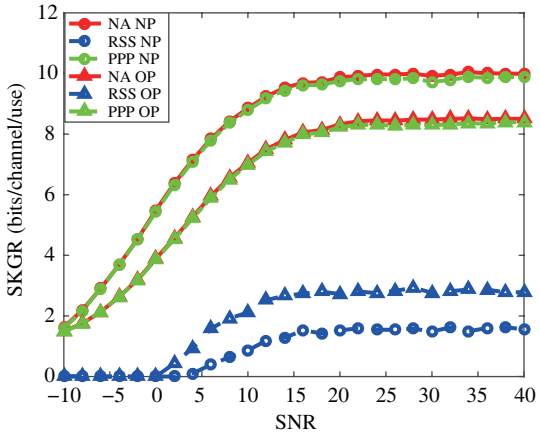

Figure 6 (Color online) SKGR according to SNR under MITM attack with legitimate node SNR of 10 dB, $N_A = 4$, $N_B = N_E = 1$.

Compared with Figures 3 and 4, we observe that although the increase in the number of antennas for Eve can enhance the interception of information for the RSS method, the same increase has negligible effect on the SKGR obtained from our proposed scheme, which suggests its robustness.

In the fourth scenario, when $N_A = 4$ and $N_B = N_E = 1$, the SKGR according to the Eve SNR is shown in Figure 6. From the figure, we observe that the SKGR increases up to steady values with the increasing Eve SNR, which indicates that Eve can gain control over the wireless channels up to a certain limit.

To verify the performance and suitability of the proposed private pilot, we compared our scheme with both SKG method described in [7] under passive eavesdropping and the method in [31] under an impersonation attack. The results are reported in the following.

5.2 Passive eavesdropping

In this scenario, passive eavesdropper Eve only monitors the information transmitted over the wireless channels during the SKG, and tries to acquire the channel information of the legitimate nodes from any intercepted information. Once Eve obtains enough information about channels $\tilde{\boldsymbol{H}}_{AB}$ and $\tilde{\boldsymbol{H}}_{BA}$, it can infer a part of the generated keys. However, provided Eve is silent during the SKG, it is still considered as a passive eavesdropper.

Given the private pilot security, Eve cannot estimate channels $\tilde{\boldsymbol{H}}_{AB}$ and $\tilde{\boldsymbol{H}}_{BA}$ and can only acquire signals $\boldsymbol{Y}_{AE}$ and $\boldsymbol{Y}_{BE}$, even if Eve is close to Alice or Bob. Thus, the proposed SKG scheme with private pilot is secure against passive eavesdroppers.

We explored the scenario where $N_A = 4$, $N_B = 1$, which applies to MISO scenario in the MTC in the presence of a single antenna or multi-antenna (false base station) eavesdropper. For passive eavesdropping, we assumed the worst case where either the eavesdropper is very close to a legitimate node or the wireless channels slowly change such that Eve can determine the channel of the legitimate nodes. Thus, we assumed that $\boldsymbol{H}_{AB} \approx \boldsymbol{H}_{AE}$ and compared the SKGR obtained from our approach with that obtained from the RSS method described in [7].

Figure 7 shows that the SKGR rises with the increasing SNR of the legitimate nodes, and there are some other characteristics: first, the SKGR obtained from the proposed scheme is almost equal to that with NA, which verifies the security of our scheme; second, the SKGR obtained from the RSS method is zero because of $\boldsymbol{H}_{AB} \approx \boldsymbol{H}_{AE}$, which makes the received signals of Bob and Eve almost the same, and thus, Eve can intercept all the secret keys according to the SKG agreements; third, the SNR of the legitimate nodes have a neglible effect on the SKGR obtained from the RSS method; finally, the increase in the number of eavesdropper antennas has no effect on the SKGR obtained from the proposed SKG scheme.

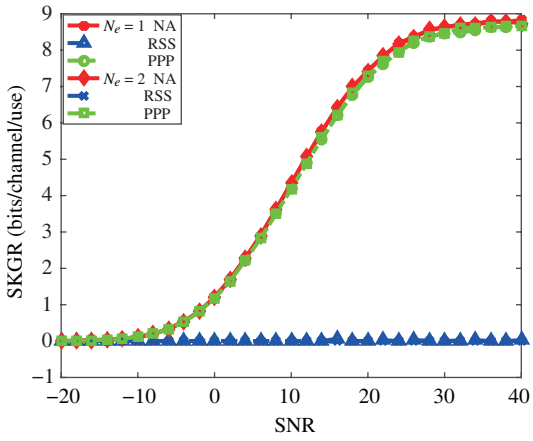

Figure 7 (Color online) SKGR according to SNR under passive eavesdropping with $N_A = N_B = 1$, $N_E = 1$ or 2.

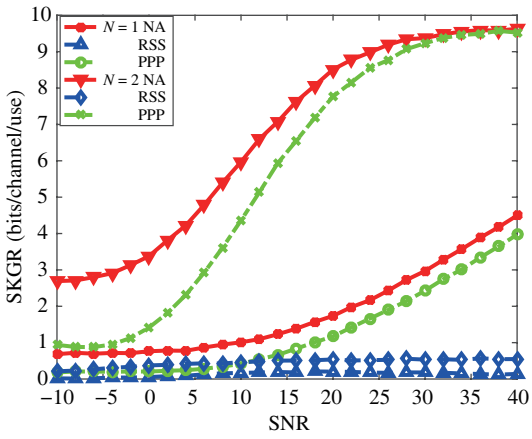

Figure 8 (Color online) SKGR according to the SNR of legitimate nodes under replay attack with $N_A = N_B = N_E = 1$ or $N_A = N_B = N_E = 2$.

5.3 Impersonation attack

The impersonation attack is another active attack model, where the attacker stores the information transmitted by a legitimate node, and then imitates that node to communicate with other legitimate nodes [31–33]. The impersonation attack mainly includes signal and feature replay attacks. We only discuss the signal replay attack here, because it is similar to the feature replay attack, which reproduces the features of the legitimate users instead of its signals.

Signal replay attack [33] is an attack where, for instance, attacker Eve stores the signals (including the pilot) transmitted by Alice, and then impersonates Alice to communicate with Bob by transmitting the stored signals and acquire private information from Bob. However, Eve cannot execute a replay attack on legitimate users who communicate and generate secret keys using the dynamic private pilot. The reason is that Alice and Bob use different private pilots to estimate the wireless channels each time, and thus Eve cannot demodulate the channels and intercept their information due to the private pilot security.

We only present two scenarios in this subsection, namely $N_A = N_B = N_E = 1$ and $N_A = N_B = N_E = 2$. These scenarios correspond to independent pairwise models to generate secret keys in SISO or MIMO cases. The SKGR obtained from our approach is compared with that obtained from the impersonation attack described in [31]. We examined the change of SKGR according to SNR under a signal replay attack and assumed that the SNR of the attacker is same as that of the legitimate nodes.

Figure 8 shows that the SKGR rises for increasing SNR values. The results can be summarized as follows: first, the SKGR for the method described in [31] remains low given the private pilot security, which prevents the attacker to cancel the pilot, and the illegitimate pilot signals are considered as noise by the legitimate nodes; second, the SKGR difference between our proposed approach and the NA case decreases for high SNR values because the effect of the interference signals transmitted by the attacker decreases as the SNR increase; finally, the SKGR for the method described in [31] is notably lower than that for our proposed scheme, which indicates that the proposed scheme based on private pilot can effectively prevent impersonation attacks; furthermore, the legitimate nodes can detect the attacker according to the SKGR.

6 Conclusion

In this paper, we explore the MITM transparent forwarding attack, where the attacker can intercept all the CSI passing through it when the physical-layer secret key is generated with a public pilot by using reverse-link channel training. To prevent this type of attacks, we designed an SKG scheme based on a dynamic private pilot, where both the secret key and seed information of the private pilot are

generated from the wireless channels, which are estimated with the private pilot instead of a public pilot. The proposed private pilot not only encrypts the wireless channels but also authenticates the channels. Simulation results show that the proposed scheme withstands MITM attacks, passive eavesdropping and impersonation attacks. Given the security and privacy of the proposed private pilot, the attackers cannot demodulate the received signals without the private pilot in the wiretap model. However, how the legitimate nodes can discover the wiretapped channels and attacker locations need to be further explored. Thus, the use of the proposed private pilot in a general wiretap channel model is part of our future investigation.

Acknowledgements The authors would like to thank the anonymous reviewers for their detailed evaluation and constructive comments. This work was partially supported by National High-Tech R&D Program of China (863) (Grant No. SS2015AA011306), National Natural Science Foundation of China (Grant Nos. 61601514, 61379006, 61401510, 61521003, 61501516), and China Postdoctoral Science Foundation (Grant No. 2016M592990).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Li N, Tao X F, Wu H C, et al. Large system analysis of artificial noise assisted communication in the multiuser downlink: ergodic secrecy sum-rate and optimal power allocation. IEEE Trans Veh Technol, 2016, 65: 7036–7050
2 Qi X H, Huang K Z, Zhong Z H, et al. Physical layer security of multi-hop aided downlink MIMO heterogeneous cellular networks. China Commun, 2016, 13: 120–130
3 Ji X S, Kang X L, Huang K Z, et al. The full-duplex artificial noise scheme for security of a cellular system. China Commun, 2015, 12: 150–156
4 Li M L, Guo Y F, Huang K Z. Secure power and subcarrier auction in uplink full-duplex cellular networks. China Commun, 2015, 12: 157–165
5 Zhang L J, Jin L, Luo W Y, et al. Robust secure transmission for multiuser MIMO systems with probabilistic QoS constraints. Sci China Inf Sci, 2016, 59: 022309
6 Li X Y, Jin L, Huang K Z, et al. Transmission frequency-band hidden technology in physical layer security. Sci China Inf Sci, 2016, 59: 019301
7 Lou Y M, Jin L, Zhong Z, et al. Secret key generation scheme based on MIMO received signals spaces (in Chinese). Sci Sin Inform, 2016, 47: 362–373
8 Khisti A. Interactive secret key generation over reciprocal fading channels. In: Proceedings of 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, 2012. 1–8
9 Kapetanovic D, Zheng G, Rusek F. Physical layer security for massive MIMO: an overview on passive eavesdropping and active attacks. IEEE Commun Mag, 2015, 53: 21–27
10 Zhou X, Maham B, Hjrungnes A. Pilot contamination for active eavesdropping. IEEE Trans Wirel Commun, 2012, 11: 903–907
11 Zhou H, Lauren M H. Secret key generation in the two-way relay channel with active attackers. IEEE Trans Inf Forens Secur, 2014, 9: 476–489
12 Zafer M, Agrawal D, Srivatsa M. Limitations of generating a secret key using wireless fading under active adversary networking. IEEE/ACM Trans Netw, 2012, 20: 1440–1451
13 Bellovin S M, Merritt M. Encrypted key exchange: passwordbased protocols secure against dictionary attacks. In: Proceedings of IEEE Computer Society Symposium on Research in Security and Privacy, Oakland, 1992. 72–84
14 Demillo R, Merritt M. Protocols for data security. Computer, 1983, 2: 39–51
15 Baker W, Goudie M, Hutton A, et al. Data breach investigations report. Methodology, 2011, 36: 1–63
16 CAPEC. Capec-94: Man in the middle attack. 2014. http://capec.mitre.org/data/definitions/94.html
17 Frankel S, Eydt B, Owens L, et al. Establishing wireless robust security networks: a guide to IEEE 802.11i. National Institute of Standards and Technology, Gaithersburg. Report No. NIST SP 800-97. 2007
18 Mayank A, Santosh B, Sukumar N. Advanced stealth Man-in-The-Middle attack in WPA2 encrypted Wi-Fi networks. IEEE Commun Lett, 2015, 19: 581–584
19 Song I-A, Lee Y-S. Improvement of key exchange protocol to prevent Man-in-The-Middle attack in the satellite environment. In: Proceedings of 8th International Conference on Ubiquitous and Future Networks (ICUFN), Vienna, 2016. 408–414
20 Conti M, Dragoni N, Lesyk V. A survey of Man-in-The-Middle attacks. IEEE Commun Surv Tutor, 2016, 18: 2027–2051
21 Ye C, Mathur S, Reznik A, et al. Information-theoretically secret key generation for fading wireless channels. IEEE Trans Inf Forens Secur, 2010, 5: 240–254
22 Thomas M, Joy A T. Elements of Information Theory. New York: Wiley-Interscience, 1991
23 Bjornson E, Ottersten B. A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance. IEEE Trans Signal Process, 2010, 58: 1807–1820

24 Shariati N, Wang J, Bengtsson M. Robust training sequence design for correlated MIMO channel estimation. IEEE Trans Signal Process, 2014, 62: 107–120
25 Soltanalian M, Naghsh M M, Shariati N, et al. Training signal design for correlated massive MIMO channel estimation. IEEE Trans Wirel Commun, 2017, 16: 1135–1144
26 Chae S H, Choi W, Lee J H, et al. Enhanced secrecy in stochastic wireless networks: artificial noise with secrecy protected zone. IEEE Trans Inf Forens Secur, 2014, 9: 1617–1628
27 Ren K, Su H, Wang Q. Secret key generation exploiting channel characteristics in wireless communications. IEEE Wirel Commun, 2011, 18: 6–12
28 Ye C, Mathur S, Reznik A, et al. Information-theoretically secret key generation for fading wireless channels. IEEE Trans Inf Forens Secur, 2010, 5: 240–254
29 Yang B, Wang W J, Yin Q Y. Secret key generation from multiple cooperative helpers by rate unlimited public communication. In: Proceedings of IEEE Internation Conference on Acoustics, Speech Signal Process (ICASSP), Florence, 2014. 8183–8187
30 Szabo Z. Information theoretical estimators toolbox. J Mach Learn Res, 2014, 15: 283–287
31 Tayebi A, Berber S, Swain A. Syncim: a new impersonation attack against chip synchronization in WSN. In: Proceedings of 9th International Conference on Sensing Technology, Auckland, 2015. 128–132
32 AlQahtani S, Gamble R. Mitigating service impersonation attacks in clouds. In: Proceedings of Future Technologies Conference (FTC), San Francisco, 2016. 998–1007
33 Kashima K, Inoue D. Replay attack detection in control systems with quantized signals. In: Proceedings of European Control Conference (ECC), Linz, 2015. 782–787