Secure TDD MIMO Networks Against Training Sequence Based Eavesdropping Attack

Yunlong Mao, Ying He, Yuan Zhang, Jingyu Hua, and Sheng Zhong

Abstract—Multi-User MIMO (MU-MIMO) has attracted much attention due to its significant advantage of increasing the utilization ratio of wireless channels. However, Frequency-Division Duplex (FDD) systems are vulnerable to eavesdropping, since the explicit CSI feedback can be manipulated. In this paper, we show that Time-Division Duplex (TDD) systems are insecure as well. In particular, we show that it is possible to eavesdrop on other users' downloads by tuning training sequences. In order to defend MU-MIMO against such threats, we propose a secure CSI estimation scheme, which can provide correct estimates of CSI when adversarial users are in presence. We prove that our scheme is secure against training sequence based eavesdropping attack. We have implemented our scheme for TDD MU-MIMO systems and performed a series of experiments. Results demonstrate that our secure CSI estimation scheme is highly effective in protecting TDD MIMO networks against eavesdropping attack. Furthermore, we extend our scheme to support massive MU-MIMO networks, with a carefully redesigned uplink protocol and optimized power allocation to achieve higher spectral efficiency. To be more practical, we also take mismatch channel issue into our consideration. An enhancement scheme is proposed and we show that our scheme with enhancement is secure and correct under mismatch channel.

Index Terms—MU-MIMO, time-division duplex, channel state information, eavesdropping attack, security

1 Introduction

Playing a key role in 5G networks, MIMO systems have received much attention from both academia and industry recently [2]. In MIMO systems, the Base Station (BS) is equipped with multiple antennas. With properly designed precoding symbols for each antenna, it can serve multiple Mobile Stations (MSs) simultaneously, and thus enhance the efficiency of wireless channels. To make such MU-MIMO applicable, two important problems need to be solved first: the multi-user detection in uplink and the multi-user interference cancellation in downlink. The prerequisite for solving both problems, especially the latter, is the Channel State Information (CSI) which characterizes coefficients of channels and is used in the precoding design [3].

The acquisition of CSI has been extensively studied in benign settings with no malicious attackers. Based on how CSI is learnt by the BS, it can be categorized into two types: explicit CSI estimation [4] and implicit CSI estimation [5]. In explicit CSI estimation, training sequences or so called "pilots", which are publicly known to all MSs and the BS, are transmitted from the BS to MSs. Based on received signals of these sequences, each MS estimates its CSI and sends it back to the BS. In implicit CSI estimation, the publicly known training sequences are sent from MSs to the BS on the contrary. Next, the BS estimates each MS's CSI with received signals.

Recently, a serious eavesdropping attack on downlink of MU-MIMO with explicit CSI estimation has been identified [6]. The attacker feeds back forged CSI to the BS to make the multi-user downlink signal cancellation fails, and correspondingly the attacker's received signal will be a mixture of its own download and the victim's download. With a careful selection of forged CSI, the attacker is able to extract the victim's downlink content.

Note that the above eavesdropping attack is based on manipulating the explicit CSI feedback from the MS which does not exist in implicit CSI estimations, it is feasible only in MU-MIMO with explicit CSI estimation. Noticing this, we are interested in whether similar attacks can be launched when the implicit CSI estimation is adopted. In fact, implicit CSI estimation is often used in TDD systems. Due to the reciprocity of TDD channels, estimation of CSI at transmitting end can be done by implicit CSI estimation with only one single pass of training sequences.

In this paper, we first propose a new eavesdropping attack which can be launched in TDD MU-MIMO networks with implicit CSI estimation. Our attack can result in the downlink leakage in TDD MU-MIMO networks. The main idea is as follows. The attacker keeps eavesdropping on the BS. Because training sequences are publicly known, the attacker is able to estimate the victim's CSI with the signal of victim's training sequence at the BS. With this knowledge, we show that the attacker is able to compute a "poisoned" training sequence, and use it to mislead the BS to adopt a seriously ill precoding (Please see details in Section.4). Under this precoding, the attacker is able to extract the victim's download content.

We point out that the downlink leakage and the above attack are highly challenging to be prevented. First, this eavesdropping attack in implicit CSI estimation is difficult

- *The authors are with the State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, Jiangsu 210008, China, and also with the Computer Science and Technology Department, Nanjing University, Nanjing, Jiangsu 210008, China. E-mail: {maoyl, zhangyuan, zhongsheng}@nju.edu.cn, 151220038@smail.nju.edu.cn, huajingyu2012@gmail.com.*

Manuscript received 31 Mar. 2018; revised 29 Apr. 2019; accepted 12 Aug. 2019. Date of publication 26 Aug. 2019; date of current version 3 Nov. 2020.
(Corresponding author: Yunlong Mao.)
Digital Object Identifier no. 10.1109/TMC.2019.2937081

to be identified, since there is no evidence of the transforming of training sequences for the BS to tell whether its received training sequences are original ones. Second, there is a conflict in protecting CSI estimation, where CSI and training sequences should be available to the BS but should not be available to other MSs even in the situation in which the BS may be eavesdropped on. Third, when users' CSI has not been estimated in TDD systems, we cannot assume that the BS can decompose users' upload concurrently, which means the availability of users' uplink will be very limited. Fourth, complex cryptographic tools cannot be adapted. Construction of secure communication or secure key exchanging in unstable transmitting is too inefficient for estimation phase. Masking or confusing techniques cannot be adapted because training sequences and CSI need to be as accurate as possible.

Despite these challenges, we design a novel secure CSI estimation scheme that successfully defends TDD MU-MIMO networks against the above eavesdropping attack. Our secure CSI estimation is designed to be a two-phase scheme. The first phase is for the BS to collect commitments of training sequences from MSs, which should happen in the previous coherence interval before expected downlink interval. A fuzzy commitment scheme [7] is used to generate commitments of training sequences for MSs, which can help the BS ensure that when the BS does CSI estimation, MSs will use exactly the same sequences as they have committed. The first phase mainly involves uplink of MSs. This phase can be regarded as a process of multi-user detection in conventional uplink schemes. The second phase should happen in expected coherence interval. In the beginning of the second phase, MSs should reveal their training sequences to the BS. Then the BS verifies whether these training sequences can match those that have been committed before. If the training sequence of any MS is legal, the BS will recover the original training sequence from the MS's commitments and estimates this MS's CSI with the original training sequence and received training sequence.

The above countermeasure is based on the observation that instantaneous CSI is changing rapidly. Therefore we can make use of the timeliness of CSI to ensure that current CSI is secure against attackers. But there may be relatively statistical CSI cases which should be also taken into consideration. To this end, we extend our secure CSI estimation scheme to achieve adaptive security in relatively statistical CSI case. Adaptive security here means that varying degree of security can be achieved according to the selection of security parameter. If a security threshold is determined, then varying degree of resources will be expended according to the statistical degree of CSI.

To be more practical, we have also designed a new uplink scheme to make our secure CSI estimation scheme feasible for massive MIMO networks. This is important because when the amount of users is massive, the overhead of our original scheme will increase significantly. Especially the uplink transmission will congest. To address this problem in massive MIMO networks, we introduce an improved uplink design for our secure CSI estimation in Section.8. This uplink design can guarantee the robustness of uplink protocol against eavesdropping attacker in massive MIMO systems, which has not been considered in previous work [1]. We analyse the spectral efficiency and power allocation of our new uplink design theoretically and experimentally. Our results show that, with our secure CSI estimation applied, we can still optimize power allocation for uplink under the constraint of mobile devices' energy budget.

Furthermore, we also take mismatch channel into consideration, which is an important issue in TDD MIMO systems [8], [9], [10], [11], [12], [13], [14]. When mismatch channel is considered in TDD systems, the reciprocity of channel will be broken. Since the reciprocity of channel is the foundation of implicit CSI estimation, some enhancement should be made to keep CSI estimation correct and secure under the mismatch channel. We enhance our secure CSI estimation scheme in Section.9 to deal with mismatch channel and we show that the enhancement can defend against the adversary efficiently through analysis.

Our contributions can be summarized as follows:

- We propose a feasible approach to make eavesdropping attack feasible in TDD MU-MIMO networks with implicit CSI estimation. We have verified this attack in a prototype TDD MU-MIMO network.
- We propose a secure CSI estimation scheme which takes advantage of the timeliness of instantaneous CSI. Solutions for both uplink and downlink have been given to implement our scheme. In case that CSI is relatively statistical, we also propose an adaptive security approach. This approach can be integrated to ensure security with varying cost according to the statistical degree of CSI.
- To support massive MU-MIMO networks, we propose a new uplink design for our secure CSI estimation. To achieve higher spectral efficiency, power allocation for uplink is also optimized to meet mobile devices' energy budget.
- We propose an enhancement for our secure CSI estimation scheme under mismatch channel situation. This enhancement can guarantee that our scheme is correct and secure against the adversary in both instantaneous CSI case and statistical CSI case.

2 Related Work

MU-MIMO systems have attracted more and more attention since it emerged. Many essential techniques are growing mature. For uplink, minimum mean square error method and maximum likelihood method have been proved efficient in multi-user detection [15], [16]. Some uplink protocols are designed to contend for channels without coordination [17], [18], [19]. This kind of methods is easy to establish, and flexible to use. Some other protocols are designed to use uplink with coordinated access [20], [21], [22]. This kind of methods usually has higher utility of channels, but the prerequisite is that full CSI of users must be available.

As for downlink of MU-MIMO, CSI is essential for highly reliable transmission. To obtain CSI within small delay for transmitting, lots of subtle estimation schemes have been proposed. For explicit CSI estimation, many efficient feedback schemes and training techniques have been proposed, such as [4], [23], [24]. Meanwhile, implicit CSI estimation has an advantage of single-pass training form. This can be used with reciprocity characteristic of time-duplex division systems to shape a complete CSI estimation scheme [5], [25] for both transmitters and receivers.

Although much work about CSI estimation has been done, only a small part of work focuses on the security of CSI estimation. It is commonly agreed that CSI should be fed or learnt in plain text. To achieve physical layer security, artificial noise is usually used. In [26], secrece transmission rate with artificial noise in massive MIMO networks has been thoroughly studied. The authors of [27] report an attack that can exploit the spatial uncorrelation property to confuse the location distinction by generating a virtual multipath channel. An adversary can simulate a real multipath by controlling the time of sending signals and attenuation factors. To eliminate the effect of obstacles in real world, the attackers can utilize reverse-engineering existing wireless channel estimation algorithms and perform linear transformations to their signals. The authors propose a detection method to defend this attack [27]. By placing another receiver at a different position as an auxiliary, honest users can compare the channel characteristics of the crafted signal in the auxiliary receiver with the target receiver to find the inconsistency.

Recently, a serious eavesdropping attack has been proved to be practicable in MU-MIMO system even with protection of artificial noise [6]. This attack can threaten users' download if CSI is available to the attacker. The attack will be able to extract the victim's downloading message if forged CSI is reported to the BS. This attack is proposed and proved practical in explicit CSI estimation with FDD mode. But in implicit CSI estimation, this attack will be inapplicable because CSI is not explicitly reported. In [1], the authors have shown a feasible approach to launch the eavesdropping attack in implicit CSI estimation. Different with [6], this attack is launched by misleading the BS with the fraudulent training sequence instead of forging a CSI report. The authors of [28] have given a nice theoretical analysis of achievable secrecy rate under a pilot contamination attack which is a similar to the attack studied in [1], [6]. The main difference between [28] and previous work is that [28] has theoretically proved the vulnerabilities of massive MIMO systems to fake pilot based attacks and achievable individual secrecy rate is given. Unlike [28], we have investigated the eavesdropping attack based on fine-tuned training sequence in a more practical way. Instead of focusing on theoretical analysis only, we have also advanced the eavesdropping attack and the proposed secure transmission scheme in real world implementation. Furthermore, an extension of our scheme will be given in this paper to deal with mismatch channel situation.

Although a secure CSI estimation scheme is proposed in [1], there are still some practical issues to be taken into account. So in this paper, we will consider the extension for two important application scenarios, i.e., massive MU-MIMO and mismatch channel.

3 Preliminaries

We will focus on the leakage of downlink in a single-cell TDD MU-MIMO network in this paper. We note that the biggest difference between the TDD system and FDD system is that the TDD system uses the reciprocity of channels for transmitters and receivers. This is also the main factor that makes secure CSI estimation in TDD systems difficult to be achieved. In the favor of TDD systems, we use implicit CSI estimation method as the default setting. We assume the channels are stochastic block fading [29]. Before the introduction of our attack model and the proposed scheme, we will introduce TDD MU-MIMO system and implicit CSI estimation first.

3.1 Downlink in MU-MIMO

In this MU-MIMO network, we assume that the BS uses M antennas to communicate with K single-antenna MSs. As mentioned before, there are two elementary requirements to design MU-MIMO MAC protocols: multi-user detection scheme in uplink and multi-user interference cancellation scheme in downlink [16]. Although dirty paper coding [30] has been proved to be the most effective interference cancellation scheme theoretically, zero-forcing (ZF) is now the most popular scheme in practical use. So we use ZF scheme as the foundation of downlink. The stochastic block-fading channel of downlink between the BS and MSs can be represented by a $K \times M$ matrix:

$$\boldsymbol{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1M} \\ h_{21} & h_{22} & \cdots & h_{2M} \\ \vdots & & \ddots & \vdots \\ h_{K1} & h_{K2} & \cdots & h_{KM} \end{bmatrix}, \tag{1}$$

where $h_{ij}, i \in [1:K], j \in [1:M]$ is the complex coefficient from the jth antenna of the BS to the ith MS. According to IEEE 802.11ac standard, download data should be modulated into multiple streams for N subcarriers based on Orthogonal Frequency-Division Multiplexing (OFDM). We should use $\boldsymbol{H}_k$ to denote the channel coefficient on the kth subcarrier. But to be succinct, we ignore subscript k for subcarriers unless when the statement is in need of specific subcarriers. We use $\boldsymbol{h}_i$ to denote the ith row of $\boldsymbol{H}$, which characterizes full CSI from the all M antennas of the BS to the ith MS. In this way, the received signal of MSs in downlink can be represented as:

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{X} + \boldsymbol{Z}, \tag{2}$$

where $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_M]^T$ is modulated symbols from BS's M antennas, $\boldsymbol{Z} = [z_1, z_2, \ldots, z_M]^T$ is the additive white Gaussian noise (AWGN), satisfying $E\{z_i\} = 0, Var\{z_i\} = \sigma^2, i \in [1, M]$. We use $\boldsymbol{H}^T$ to denote transpose of $\boldsymbol{H}$, and $\boldsymbol{H}^\dagger$ to denote Hermitian transpose of $\boldsymbol{H}$.

3.2 Channel Estimation

According to ZF scheme, symbols $\boldsymbol{X}$ modulated by the BS should be precoded with inverse matrix of CSI. To obtain full CSI at the transmitter end, MSs should transmit commonly known training sequences on N subcarriers to the BS in the training phase. Assume that all these N subcarriers are orthogonal perfectly. We can denote training symbols of these N subcarriers as $\boldsymbol{s} = [s_1, s_2, \ldots, s_N]^T$, where $\boldsymbol{s}_i$ is training symbol of the ith subcarrier. Note that the length of training symbols should be larger than M. The BS will estimate MS's uplink CSI with known training sequences $\boldsymbol{s}$ and received training sequences $\boldsymbol{Y}_s$. In order to calculate the approximation of $\boldsymbol{H}$, minimum mean square error (MMSE) is commonly used. The linear MMSE estimator [31] of $\boldsymbol{H}$ is:

$$\hat{H}_{MMSE}(\boldsymbol{Y}_s, \boldsymbol{s}) = \boldsymbol{Y}_s(\boldsymbol{S}^H \boldsymbol{R_H} \boldsymbol{S} + \sigma_0^2 K I)^{-1} \boldsymbol{S}^H \boldsymbol{R_H}, \tag{3}$$

where $R_H = E\{\boldsymbol{H}^H \boldsymbol{H}\}$. Then according to the reciprocity of channels, we can have downlink CSI as $\hat{\boldsymbol{H}}^T$.

3.3 Fuzzy Commitment Scheme

Fuzzy commitment is usually applied to biometric templates, such as fingerprint authentication system. Since the readings of the same fingerprint are not always identical, biometric templates, as important as passwords, require resilience to small corruptions. Fuzzy commitment was proposed to meet this need. Recently, fuzzy commitment scheme is also used to construct some components of secure communication, such as [32], [33], [34]. To the best of our knowledge, the scheme proposed in [1] is the first solution to use fuzzy commitment in MIMO systems. In this paper, we share the same use of fuzzy commitment, which means that training sequence will be protected by fuzzy commitment scheme. Fuzzy commitment r,ather than other cryptography tool is used because we need to recover the committed training sequence from unreliable channel. A fuzzy commitment scheme can allow a blob $y = F(b, x)$, where $F()$ represents fuzzy commitment function, x is confidential message, b is ,blurring parameter, to be opene,d using any witness x' that is close to ,x in some appropriate metric, but not necessarily identical to x. The,re are two main tasks for the fuzzy commitment employed in our scheme: one is to protect MSs,' training sequences against malicious users, and the other one is to deal with unreliable channel when MSs reveal commitments. For more detailed description and applications of fuzzy scheme, please refer to [35].

4 Eavesdropping Attack Based on Training Sequence

To do channel estimation normally, every MS should transmit commonly known training sequence. However, some malicious MS is capable of misleading the BS by transmitting elaborately forged training sequence instead of the benign one. More specifically, the BS will give estimation $\hat{\boldsymbol{H}} \approx \boldsymbol{H}$ when the benign training sequence $\boldsymbol{s}$ is transmitted. If the MS transmits $\Delta\boldsymbol{s}$ instead, the BS will give estimation $\hat{\boldsymbol{H}} \approx \boldsymbol{H}\Delta\boldsymbol{H}$, since the BS still uses $\boldsymbol{s}$ as expected training sequence. This kind of eavesdropping attack is hard to be identified because the attacker can forge any channel state by transforming its training sequence and the BS has no ability to find out whether CSI is forged or not. The main reason is that conventional CSI estimation is based on the assumption that both the BS and MSs are trusted.

4.1 Threat Model

The BS must be trusted in any scenario, because every effort for physical security will be in vain if the BS is compromised. We do not consider attackers from outside of this MU-MIMO network because the attack that we study here can only be launched inside, because attackers from outside cannot have interactive behavior such as estimating and downloading with the BS. This means our attacker is some MS in the MU-MIMO network. Other kinds of attacks from outside or higher layers are out of our concern. We will focus on physical security of MSs' CSI and downlink.

The attacker in our work is aggressive, who is able to eavesdrop on the BS to grab all signals which are supposed to be received by the BS. We find this assumption of eavesdropping on multiple antennas is reasonable because this kind of eavesdropping has been widely studied in recent literatures, such as [6], [36], [37], [38]. Please also note that there is no need for eavesdropping equipment to be at the exact location as the BS. If the victim keeps a relatively long distance from the BS, then the distance between eavesdropping equipment and the BS can be acceptable in a short range. Since the attacker can choose its victim, this prerequisite can be simply achieved. More details about the limitation of this eavesdropping attack will be discussed in an implementation example in evaluation.

Since the training sequences are transmitted in plain text, the attacker can get BS's received training sequences easily. Generally, we assume that the victim is some MS, say MS$_1$, and the attacker is another MS, say MS$_2$. MS$_1$'s training sequence received at the BS is also known to MS$_2$. Then MS$_2$ can do the same estimation as the BS and obtain CSI of MS$_1$ in the coherence interval. It is assumed that the attacker knows its own download content before the BS begins transmission. This can be easily achieved by repeating some previous download requests [6]. The BS and MSs are assumed to use traditional communication protocols as default setting if we do not make explicit redefinition for some particular component. We assume that all transmitters of the BS are protected with artificial noise against illegitimate users. This can be achieved by using spare antennas of the BS to send artificial noise on null space [39]. Our attack will be considered under this simple security assumption. Other more complex artificial noise based physical security schemes (e.g., cooperated jamming scheme [40]) need separate study and are out of our discussion here.

4.2 Transforming of Training Sequence

The attack we investigated is based on the observation that training sequences can be transformed without BS's awareness. For brevity, we will illustrate training sequence based attack in the context of a 2 × 2 MU-MIMO TDD system. The BS uses two antennas to transmit messages to two MSs with single antenna respectively. The setting of this system can be extended to more complicated systems with more antennas. In this 2 × 2 MU-MIMO system, the received signals of downlink at two MSs can be written as:

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_1 \\ \boldsymbol{h}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \end{bmatrix} \begin{bmatrix} \sqrt{p_1}\boldsymbol{x}_1 \\ \sqrt{p_2}\boldsymbol{x}_2 \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \tag{4}$$

where $[\boldsymbol{c}_1, \boldsymbol{c}_2] = [\boldsymbol{h}_1, \boldsymbol{h}_2]^{-1}$ is the precoding matrix that the BS precodes downlink streams according to ZF, $\sqrt{p_1}, \sqrt{p_2}$ are transmitting power of signal, and $\boldsymbol{x}_1, \boldsymbol{x}_2$ are signals transmitted to MS$_1$ and MS$_2$.

Generally, we assume that MS$_2$ is coveting downlink message $\boldsymbol{x}_1$ that MS$_1$ are downloading from the BS. To achieve its attempt, MS$_2$ first eavesdrops on training sequences that the BS receives from other MSs including MS$_1$. Then MS$_2$ estimates MS$_1$'s channel coefficient $\hat{\boldsymbol{h}}_1$. Attacker MS$_2$ now knows both MS$_1$'s CSI $\hat{\boldsymbol{h}}_1$ and its own CSI $\hat{\boldsymbol{h}}_2$. Then, MS$_2$ can calculate the difference $\Delta \boldsymbol{H}$ between its channel coefficient and MS$_1$'s channel coefficient:

$$\hat{\boldsymbol{h}}_1 = \hat{\boldsymbol{h}}_2 \Delta \boldsymbol{H}. \tag{5}$$

Since the BS can only estimate MSs' CSI by commonly known training sequences, it is possible for attacker MS$_2$ to mislead the BS to learn MS$_2$'s CSI by transforming benign training sequence. With CSI difference $\Delta \boldsymbol{H}$ learnt, if MS$_2$ wants its CSI to be estimated by the BS to be the same as MS$_1$, MS$_2$ should transform its training sequence into:

$$\boldsymbol{s}_2^f = \Delta \boldsymbol{s} \boldsymbol{s}_2, \tag{6}$$

where $\Delta \boldsymbol{s}$ is a solution to the following equation,

$$\hat{\boldsymbol{h}}_{2MMSE}((\Delta \boldsymbol{s} \boldsymbol{s}_2)^{MS_2}, \boldsymbol{s}_2^f) = \hat{\boldsymbol{h}}_2 \Delta \boldsymbol{H}. \tag{7}$$

When MS$_2$ uses $\boldsymbol{s}_2^f$ instead of $\boldsymbol{s}_2$. The training sequence of MS$_2$ received by the BS will be:

$$\boldsymbol{y}_2^f = \boldsymbol{h}_2 \Delta \boldsymbol{s} \boldsymbol{s}_2 + z_2. \tag{8}$$

The BS will estimate MS$_2$'s CSI as $\widehat{\boldsymbol{h}_2}^f$, which will seem like $\hat{\boldsymbol{h}}_1$. Forged CSI $\widehat{\boldsymbol{h}_2}^f$ surely is different to MS$_1$'s actual CSI $\boldsymbol{h}_1$. But we should not worry about this, because what we want is to let $\hat{\boldsymbol{h}}_2^f$ looks like $\hat{\boldsymbol{h}}_1$, not $\boldsymbol{h}_1$, in BS's observation. A bound of difference between $\hat{\boldsymbol{h}}_2^f$ and $\hat{\boldsymbol{h}}_1$ is given in Theorem 1, which directly follows a property of MMSE [41].

Theorem 1 (Difference of CSI). *The difference between forged CSI $\hat{\boldsymbol{h}}_2^f$ and estimated CSI $\hat{\boldsymbol{h}}_1$ can be bounded by $(\frac{2}{s_2 \Delta H})^n \sqrt{n!}$.*

Proof. We regard channels' coefficient of both MS$_2$ and MS$_1$ as variable independently. Estimated CSI $\hat{\boldsymbol{h}}_2$ and $\hat{\boldsymbol{h}}_1$ are known to the BS. Transformed training sequence is only corresponding to $\hat{\boldsymbol{h}}_2$ and $\hat{\boldsymbol{h}}_1$. MMSE estimator of this sequence will be:

$$\begin{aligned} mmse(\boldsymbol{h}_2^f, (s_2 \Delta \boldsymbol{s})^2) &= E(\boldsymbol{h}_2^f - E\boldsymbol{h}_2^f | (s_2 \Delta \boldsymbol{s}) \boldsymbol{h}_2^f + z_2)^n \\ &\leq (\frac{2}{s_2 \Delta \boldsymbol{s}})^n \sqrt{n!}. \end{aligned}$$

□

4.3 Eavesdropping Attack

In the honest MU-MIMO scenario, after receiving training sequences from MSs, the BS will estimate MSs' CSI and prepare download for MSs. Having download content precoded with inverse matrix of full CSI by the BS, MSs are supposed to receive their own downlink content. But when there is an adversarial user who covets another MS's downlink content, received content of the attacker will be a mixture of the attacker's download and the victim's download. The attacker must use its own CSI and the victim's CSI to eliminate channel coefficient and precoding matrix. Hence, a big problem for eavesdropping attack is the elimination of interference. To solve this problem, the prerequisite for eavesdropping is acquisition of the victim's CSI, assuming that CSI is in plain text.

Although MS$_2$ can transform its training sequence to $\boldsymbol{s}_2^f = \Delta \boldsymbol{s} \boldsymbol{s}_2$ to imitate CSI of MS$_1$, this is not the best choice to eavesdrop on MS$_1$. When MS$_2$ fakes its CSI as $\boldsymbol{f}_2 = [f_{21}, f_{22}]$ by the transforming of training sequence $\boldsymbol{s}_2^f$, the download messages of MS$_1$ and MS$_2$ from the BS with ZF precoding will be:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_1 \\ \boldsymbol{h}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{f}_2 \end{bmatrix} \begin{bmatrix} \sqrt{p_1} \boldsymbol{x}_1 \\ \sqrt{p_2} \boldsymbol{x}_2 \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}. \tag{9}$$

When MS$_2$ does eavesdropping and downloading at the same time, the downlink content of MS$_1$ in MS$_2$'s observation will be:

$$\begin{aligned} \boldsymbol{x}_1^{MS_2} &= \frac{(h_{11} f_{22} - f_{21} h_{12})}{\sqrt{p_1}(h_{21} f_{22} - h_{22} f_{21})} (\boldsymbol{y}_2 - \frac{(h_{11} h_{22} - h_{12} h_{21})}{(h_{11} f_{22} - f_{21} h_{12})} \sqrt{p_2} \boldsymbol{x}_2) \\ &= \boldsymbol{x}_1 + \frac{(h_{11} f_{22} - f_{21} h_{12})}{\sqrt{p_1}(h_{21} f_{22} - h_{22} f_{21})} z_2. \end{aligned} \tag{10}$$

As shown in expression of $\boldsymbol{x}_1^{MS_2}$, in order to maximize the attacker's eavesdropping and minimize the interference of its own downloading message, the key is to forge the attacker's CSI. A weighted sum of genuine CSI has been proved to be the best choice [6]. $\boldsymbol{f}_2 = [wh_{11} - h_{12}, wh_{21} - h_{22}]$, where w is a adjustable coefficient. This is the approach proposed by prior work [6] in a FDD system with explicit CSI estimation. We in this paper, make this eavesdropping approach also feasible in a TDD system with implicit CSI estimation by transforming training sequences. To ensure that the BS can learn $\boldsymbol{f}_2$ as MS$_2$'s CSI, the training sequence of MS$_2$ should be transformed in the way shown in Equation 6. The received signal at attacker MS$_2$ contains a mixture of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. In order to decode $\boldsymbol{x}_1$ of MS$_1$, MS$_2$ should download known message $\boldsymbol{x}_2$ from a colluded or spurious server [6]. Thus, MS$_2$ can remove $\boldsymbol{x}_2$ and its own interference from the received signal. In this way, we can launch this kind of eavesdropping attack in TDD systems, but no existing scheme can prevent it effectively.

5 Secure Estimation of Instantaneous CSI

There are generally two types of CSI: statistical CSI and instantaneous CSI. Statistical CSI usually has strong correlation in space, time and frequency, and can be described by statistical characteristics of the channel. This type of CSI usually has no need to be estimated over time. Thus, we will focus on instantaneous CSI first which needs to be estimated continually. Our secure CSI estimation procedure of MSs is designed to have two phases. The first phase is MSs' generating commitments, which should happen in the previous coherence interval before MS's expected downlink interval. The second phase is revealing commitments, which should happen in the same coherence interval with expected download. As for the BS, commitments of training sequences should be collected in the first phase, and CSI estimation will be done in the second phase.

In order to show the changing of CSI in different intervals, we have performed some experiments to measure

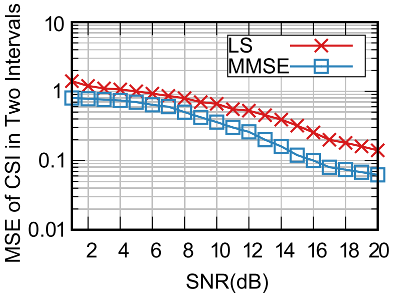


Fig. 1. Difference of CSI in two intervals. The difference is measured in MSE.

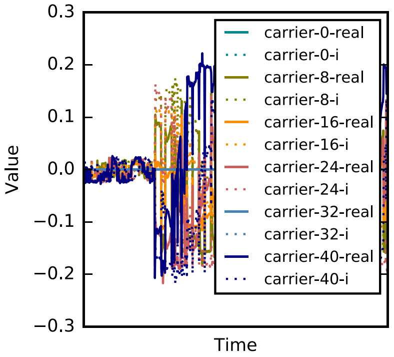


Fig. 2. State of OFDM MIMO channels in experiments.

how much difference of CSI between adjacent coherence intervals will be. The result is shown in Fig. 1. The vertical axis indicates the difference of CSI between adjacent coherence intervals. Mean squared error is used to measure the difference. CSI estimated by least squares (LS) method causes larger difference than the CSI estimated by MMSE method. The horizontal axis indicates SNR increasing. It can be seen that difference of CSI gets smaller when SNR gets higher. However, even if SNR is high enough, this difference of CSI is still significant. Fig. 2 shows practical state of OFDM MIMO channels in our experiments. Complex components of 8 subcarriers are plotted.

Specifically, each MS generates a random training sequence in the previous coherence interval just ahead of its expected download. Then each MS makes a commitment of this random training sequence to the BS. The BS will put these MSs, whose commitments have been collected in the first phase, into the scheduling of next interval. When the second phase begins, each MS reveals its commitment of training sequence to the BS. The BS then can do correct estimation of each MS's CSI. An illustration of the two phases is given in Fig. 4. Some important notations used in this paper can be referred to Table 1.

5.1 First Phase: Generating Commitments

Different from FDD systems, when CSI of downlink is unknown in TDD systems, a reliable uplink cannot be assumed because of the reciprocity of TDD channels. Thus, any secure CSI estimation scheme for TDD systems must be constructed from the very beginning of communication, including uplink and downlink. Because of lacking CSI, coordinated uplink access or synchronous uplink transmissions [16] cannot be achieved directly. To be realistic and efficient in establishing communication, we use a spatial multiple access scheme similar with SAM [18]. Then, we will construct our secure downlink scheme based on this uplink scheme to achieve an effective and secure TDD MU-MIMO system.

TABLE 1
Notation Table for Quick Reference

Notation	Definition
H	Channel coefficients between the transmitter and the receiver.
ΔH	The difference between channel coefficients of two MSs.
s	Training sequence used by MS.
Δs	The difference between two training sequences.
r_i	Randomly generated training sequence for MS$_i$.
r_i'	MS$_i$'s random training sequence received at the BS.
E_s	Symbol transmitting energy.

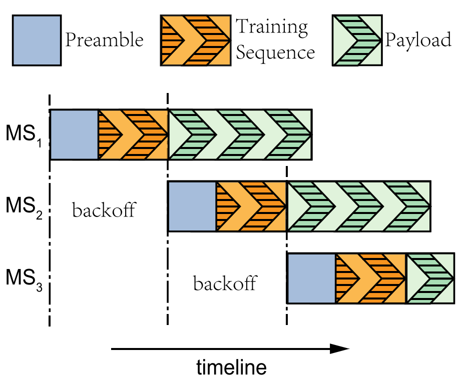


Fig. 3. A sketch of uplink design in our proposed scheme.

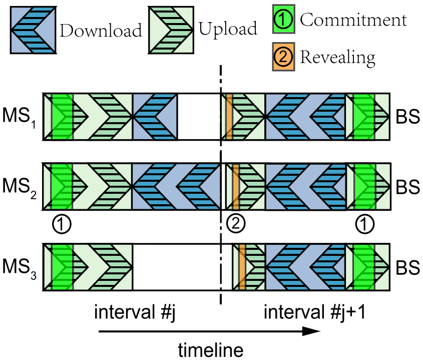


Fig. 4. Examples of the relation between two phases. ① represents locations of commitments. ❷ represents locations of revealing messages.

SAM [18] uses a chain decoding scheme to decompose the overlapped parts of frames. In this paper, we use a very similar design to implement our uplink protocol. When CSI of MS's channel to the BS is unknown, each MS needs a clear channel to ensure the transmission of its preamble and training sequence. Like SAM, we use a chain decoding method to implement an efficient spatial multiple access. An uplink example of our design is illustrated in Fig. 3. Say there are three users served by the same BS. Since their CSI is unknown to the BS, they need to detect collisions and then contend for the access to the upload channel. If MS$_1$ gets the first opportunity to send its frames, MS$_2$ and MS$_3$ have to wait a backoff window. After that, MS$_2$ and MS$_3$ will begin another contending. Generally, our uplink protocol frames are composed of three parts, preamble, training sequence and payload. When the BS receives the transmission of MS$_1$, the BS will estimate MS$_1$'s CSI and recover its payload. When the BS wants to decompose the payload of MS$_2$, the BS needs to remove all frames of MS$_1$ first. Then the BS can estimate MS$_2$'s CSI and recover its payload just like MS$_1$. Other MS's upload can be processed in the same way successively. The authors also put the above explanation into the revision. We hope this can make our proposed scheme more complete.

It is commonly agreed that training sequences should be inserted in data stream of uplink to avoid frequent interruption of continuous download in TDD systems [42], [43]. We follow this rule in our scheme. But what's different is that we insert the commitment of training sequence for the next coherence interval in current uplink payload. This means MSs always commit their training sequences that will be used in the next coherence interval beforehand. MSs will reveal their training sequences in the second phase which is in the next coherence interval. MSs who newly join this MU-MIMO network should wait for an guard interval. Then these new MSs will be able to use their training sequences for CSI estimation at the BS, and start their download process.

Take the scenario shown in Fig. 4 as an example. In the middle of coherence interval j, MS$_1$, MS$_2$ and MS$_3$ happen to be in upload slot. Each of them will put valid payload in its uplink since their uplink CSI has been estimated by the BS in the beginning of the interval. Part of the payload is highlightened in light green and indicated by ①. This is where each MS's commitment is placed. When upload slot ends, there may be some download slots until the end of the coherence interval. So the first phase of our scheme is completely included in MS's upload. If one MS wants to join the BS in the next interval, then this MS must place its commitment of generated training sequence in one of its uplink payload before current interval ends.

All MSs that want to transmit to the BS should follow our spatial multiple uplink access scheme. Different from SAM [18], we use a fixed-format package to contain commitments of MSs' training sequences in uplink payload. And we employ a relatively loose time window of uplink channels for MSs. This is due to two reasons: ensuring interval to be larger than safe threshold and guaranteeing the demultiplex of potential transmitters. Without loss of generality, for any user MS$_i, i \in [1, K]$, a random training sequence r_i is generated. Then a linear error correcting code $ECC(n_e, t_e, Enc(), Dec())$ is used to encode r_i, where n_e is the length of codeword, t_e is the error-correcting capability, and $Enc()$ and $Dec()$ are encoding operation and decoding operation respectively. $Enc(r_i)$ will yield corresponding codeword $c_i \in 0, 1^{n_e}$. Without loss of generality, we use Hamming distance to depict difference between r_i and c_i, which will be $\delta_i = r_i \oplus c_i$, where $\oplus$ is XOR operation. But in order to keep readable, we still use numeric plus and minus hereafter. Then the construction of commitment function F is:

$$F(c_i, r_i) = (h(c_i), \delta_i), \tag{11}$$

where $h : \{0,1\}^{n_e} \rightarrow \{0,1\}^l$ is a secure hash function. This construction follows [7], so our commitment function F meets both binding condition and concealing condition. According to carrier sense multiple access with collision avoidance (CSMA/CA), MS$_i$ will send commitment $(h(c_i), \delta_i)$ to the BS when the backoff timer of MS$_i$ is up. The BS will keep collecting commitments generated by different MSs during time window of the first phase. Algorithm 1 shows summarized construction of commitments in the first phase.

Algorithm 1. Commitment of Training Sequence

```
for all $MS_i, i \in [1, K]$ do
    $MS_i$ generates random training sequence $r_i$,
    $MS_i$ chooses codeword $c_i \in C$,
    $MS_i$ computes $F(c_i, r_i) = (\alpha_i, \delta_i) = (h(c_i), r_i - c_i)$.
    $MS_i$ sends $F(c_i, r_i)$ to the BS.
end for
```

In the end of the first phase of our scheme, each MS that wants to join the BS in the next interval will have sent its commitment $F(c_i, r_i)$ to the BS. Since this happens in the middle of the interval, the BS can correctly receive all commitments. Because each MS's CSI has been estimated in the beginning of the interval.

5.2 Second Phase: Revealing Commitments

When current coherence interval ends, those MSs whose commitments of training sequences have not been collected will not be available candidates in the next coherence interval. These MSs whose commitments have been collected will enter the second phase as soon as the BS broadcasts an *end of first phase* message which is similar to CTS package in RTS/CTS protocol. To leverage fresh CSI as soon as possible, we set time window of the second phase to be compact. So MSs are required to send the revealing message simultaneously once the BS's broadcasting is detected.

We still take the scenario shown in Fig. 4 as an example. When interval j ends, the BS will broadcast the start of interval $j + 1$. Since this is a new coherence interval, no MS CSI is available at the BS. So all MSs have to follow our uplink protocol to ensure concurrent upload can be correctly decomposed. Please recall the design of uplink protocol which has been shown in Tigure.3. Without our secure CSI estimation scheme, the second part of upload frame should be a commonly known training sequence. Now in our second phase, the second part of the upload should be the revealing message, which is randomly generated training sequence r_i in the previous interval by MS$_i$. Following chain decoding rule, the BS can receive MS$_i$'s generated training sequence as r'_i.

To prevent attackers who replay commitment and revealing message of a target MS to disturb MU-MIMO system, we will do cheating detection after revealing commitments. If replicate training sequences are detected, only the first MS using this sequence will be allowed to the next step according to the timestamps when the BS received these sequences in the second phase. When the BS obtains a revealing message r'_i for any MS MS$_i$, the BS should check whether the commitment (α_i, δ_i) can be revealed by r'_i. If the revealing message transmitted by MS$_i$ is the same as the training sequence which has been committed, then the following equation will hold:

$$\alpha_i = h(f(r'_i - \delta_i)), \tag{12}$$

where $\alpha_i = h(c_i)$, $\delta_i = r_i - c_i$, f is decoding function, mapping arbitrary n_e-bits strings to nearest codewords. If MS$_i$'s commitment can be revealed correctly, then the BS will estimate the CSI using r_i and (r'_i) which are known training sequence and actually received training sequence respectively. Now, the BS can run the same MMSE as usual to calculate approximate CSI. In this way, the BS can get MS$_i$'s' CSI estimation, $\hat{\boldsymbol{h}}_{iMMSE}(r'_i, r_i)$. The whole sketch of revealing procedure is shown in Algorithm 2.

Algorithm 2. Revealing Commitment

```
for all $MS_i, i \in [1, K]$ do
    if $\alpha_i = h(f(r'_i - \delta_i))$ then
        The BS recovers $r_i = f(r'_i - \delta_i) + \delta_i$,
    end if
end for
for all $r_i, i \in [1, K]$ do
    The BS runs replicate cheating detection,
    the BS estimates $\hat{\boldsymbol{h}}_i \leftarrow \hat{\boldsymbol{h}}_{iMMSE}(r'_i, r_i)$.
end for
```

5.3 Security Analysis

This commitment scheme can protect instantaneous CSI of all MSs perfectly. But CSI will still be available to the attacker in the end of each interval, because this attacker keeps eavesdropping on all information that the BS has. Although the attacker cannot use this CSI immediately, he can use this as an outdated CSI in the next coherence interval. To guarantee that any message of any MS will be safe, we need a bound of changing of CSI so that attackers can reveal nothing by outdated CSI. To this end, the SNR of received signal should be kept under threshold δ_{snr} when the attacker tries to reveal messages with outdated CSI. Assume that both the BS and attackers can learn perfect CSI between the BS and MSs so that maximal ratio transmission (MRT) can be achieved. Then the output SNR of MRT with perfect CSI can be given by [44]:

$$\gamma_{MRT} = \lambda_{max} \frac{E_s}{\sigma^2}, \tag{13}$$

where E_s is transmitting energy for symbols, λ_{max} is the largest eigenvalue of $\boldsymbol{H}^\dagger \boldsymbol{H}$. If outdated CSI is denoted by $\tilde{\boldsymbol{H}}$ then the changing of CSI is $\delta_{\boldsymbol{H}} = \boldsymbol{H} - \tilde{\boldsymbol{H}}$. Since CSI of every moment can be regarded to follow independent Gaussian distribution, $\delta_{\boldsymbol{H}}$ can be seen as an independent difference following Gaussian distribution with zero-mean and variance σ_d^2. If the attacker uses outdated CSI to reveal other MSs' messages from received signal, the SNR will be associated with $\delta_{\boldsymbol{H}}$ by Theorem 2 which is derived from an existing theorem [45].

Theorem 2 (SNR with outdated CSI). *The SNR of received signal with outdated CSI can be given by:*

$$\gamma_{MRT} = \frac{\tilde{\lambda}_{max}}{(1 + \sigma_d^2)(\sigma_d^2 + \frac{(1+\sigma_d^2)\sigma^2}{E_s})}, \tag{14}$$

where $\tilde{\lambda}_{max}$ is the largest eigenvalue of $\tilde{\boldsymbol{H}}^\dagger \tilde{\boldsymbol{H}}$, E_s is transmitting energy of data symbol.

If a fixed SNR threshold δ_{snr} is given, then output message with SNR $\gamma_{MRT} \leq \delta_{snr}$ will be regarded as useless, i.e., no information leakage. Except for outdated CSI, preambles which are inserted to cancel frequency offset and to demultiplex MSs' signals may be leveraged by the attacker too. But everything that the attacker can get in the first phase will be outdated in the next coherence interval except for commitments. And the attacker can get nothing from these commitments. As for the second phase, it is too late for the attacker to leverage CSI which is obtained from revealing messages or short orthogonal preambles, because the BS will do zero-forcing with the CSI that the attacker has committed in the first phase.

We have taken several possible attacks against our proposed scheme into consideration. None of them is efficient enough to be practically applied except for faking training sequence. Theoretically, the attacker is still capable of generating fake training sequences. Without any knowledge of its victim's training sequence, the attacker can wildly guess or generate totally random training sequences to match victim's training sequence. This kind of attack could be successful only if a very small probability event happens. The probability of this event will be no larger than $\aleph_0^n$, where n is the length of training sequences. This attack probability is small enough to keep our scheme secure. Furthermore, for an aggressive attacker, victim's outdated CSI and training sequences may be accessible if the attacker keeps eavesdropping on the BS and victim. Since training sequences are randomly generated, previous sequences are useless for attacker's guessing newly generated sequences. However, outdated CSI can be useful when channels between the BS and MSs are relatively statistical, which means CSI do not change significantly over time. Under this circumstance, we propose an adaptive security scheme which will be introduced in details next.

6 Adaptive Security with Statistical CSI

Generally, CSI estimation is used for instantaneous CSI because this kind of CSI keeps varying rapidly. As for statistical CSI scenario, fixed CSI is usually used instead of periodic estimation. But there is always a blurred area where instantaneous CSI may vary not that fast sometimes. Or sometimes instantaneous CSI may also change slowly. Hence, we have also considered the situation where CSI is relatively statistical. Strictly speaking, we will achieve adaptive security for slow-varying CSI when SNR with changing CSI is higher than δ_{snr}. By adaptive security, we mean that varying degree of security can be achieved according to the selection of security parameter. If a security threshold is determined, then varying degree of resources will be expended according to statistical degree of CSI.

Since $\sigma_d^2 \propto (\frac{E_p}{N_0})^{-1}$ [46], where E_p is the pilot symbol energy, according to Theorem 2, it is possible to weaken the energy of training sequences to achieve a lower SNR. But the quality of downlink of MU-MIMO will be disappointing in this way because of inaccurate CSI. The hint of lower SNR can lead to another possible solution: higher Bit Error Rate (BER). This solution sounds like irrational, but we do find out that the attacker's eavesdropping can be thwarted effectively with higher BER of download content. And this BER can be bounded in a reasonable range. In other words, a little higher BER can be traded for more secure downlink. To avoid unnecessary loss of bandwidth, we propose an adaptive security scheme to control BER of downlink to prevent eavesdropping.

Assume that the SNR of current downlink is t_{snr} with CSI $\boldsymbol{H}_t$ of current coherence interval. The attacker's eavesdropping is based on previous CSI $\boldsymbol{H}_p$, so the BS can always use current CSI to calculate how much BER should be added for the next coherence interval. The BER function of QPSK modulation with AWGN can be given by:

$$BER = 1/2er\ fc\left(\sqrt{\frac{E_b}{N_0}}\right), \tag{15}$$

where $\frac{E_b}{N_0}$ is energy per bit to noise power spectral density ratio, $erfc()$ is complementary error function:

$$er\ fc(x) = 1 - \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt. \tag{16}$$

As for our MU-MIMO system,

$$\begin{aligned} \frac{E_s}{N_0}(dB) &= \frac{E_b}{N_0}(dB) + \log_2(2\log_2(M)), \\ \frac{E_s}{N_0}(dB) &= 10\log_{10}\left(\frac{T_{symbol}}{T_{sampling}}\right) + SNR(dB). \end{aligned} \tag{17}$$

Then we can calculate the compensation of SNR that is needed to achieve δ_{snr}:

$$10\log_{10}\frac{\delta_{snr}}{t_{snr}} = \left(\frac{E_b}{N_0}\right)_\delta(dB) - \left(\frac{E_b}{N_0}\right)_t(dB), \tag{18}$$

where $(\frac{E_b}{N_0})_\delta$ is corresponding to δ_{snr}, so $(\frac{E_b}{N_0})_t$ can be calculated in Equation (18). Then Equations (15) and (16) are used to calculate necessary BER δ_{ber} which should be added to downlink of each MS.

A more intuitive explanation for our adaptive security scheme is here. Recall that a mixture of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ is received when attacker MS$_2$ is eavesdropping on MS$_1$. The extraction of $\boldsymbol{x}_1$ is based on not only full CSI but also the known content of $\boldsymbol{x}_2$. When additional BER is introduced to downlink of each MS, both $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ will be less precise. When MS$_2$ tries to extract $\boldsymbol{x}_1$ as shown in Equation (10), introduced BER will be cumulated. In this way, a low SNR will be achieved. If the BS calculates the compensation of SNR in each interval, then only necessary BER should be added to downlink of each MS. We can integrate this adaptive security scheme into our secure estimation scheme easily, because when SNR is higher than δ_{snr}, no additional BER is needed, which means no loss of bandwidth in instantaneous CSI situation. Algorithm 3 shows the detailed procedure after the integration of two schemes.

Algorithm 3. Secure Estimation for Arbitrary CSI Case

```
Commitment Phase:
for all $MS_i, i \in [1, K]$ do
    $MS_i$ generates random training sequence $r_i$,
    $MS_i$ chooses codeword $c_i \in C$,
    $MS_i$ computes $F(c_i, r_i) = (\alpha_i, \delta_i) = (h(c_i), r_i - c_i)$.
    $MS_i$ sends $F(c_i, r_i)$ to the BS.
end for
Revealing Phase:
for all $MS_i, i \in [1, K]$ do
    if $\alpha_i = h(f(r'_i - \delta_i))$ then
        The BS recover $r_i = f(r'_i - \delta_i) + \delta_i$,
    end if
end for
for all $r_i, i \in [1, K]$ do
    The BS does replicate cheating detection,
    The BS estimates $\hat{H}_i \leftarrow \hat{\boldsymbol{H}}_i(r_i, r'_i)$.
end for
The BS calculates the compensation of SNR with previous
    $t_{snr}$.
if $\delta_{snr} < t_{snr}$ then
    The BS calculates necessary BER $\delta_{ber}$ which should be
    introduced to downlink of each MS.
    The BS introduces $\delta_{ber}$ by flipping bits or decreasing $E_b$.
end if
```

7 Implementation and Experiments

We have introduced our secure CSI estimation scheme theoretically. In this part, we will give one possible way to implement the proposed scheme with state-of-the-art techniques. Some parts which have been introduced in previous sections will also be reviewed in an engineering way. We will introduce our experiments implemented with GNU Radio and Universal Software Radio Peripheral (USRP). In the second subsection, numerical results of our experiments will be discussed.

When we construct uplink of MU-MIMO, we use a spatial multiple access scheme similar with SAM [18]. The core idea is to monitor any other MSs' upload frame and ensure that preambles of any two MSs are not overlapping. In the payload of uplink, a customized packet is employed to convey commitment of training sequences. This customized packet is easy to parse. There are two elements within this packet, one is the hash of codeword, and the other one is distance between codeword and training sequence. Both elements are within fixed length. Except for extra payload length, these two elements will cause no further influence to original protocol structure. We choose SHA-256 as secure hash function $h()$, and a low-density parity-check code [47], [48] is used as $ECC()$ for commitment scheme. Downlink construction is straightforward. After MSs' commitments are revealed, CSI will be estimated by MMSE method. Then the BS can do ZF precoding with full CSI in the same way as other normal MU-MIMO downlink [49].

7.1 Evaluation of Eavesdropping Attack

To demonstrate the practical threat of training sequence based eavesdropping attack, we have performed a series of indoor experiments. The BS equipped with 4 antennas is

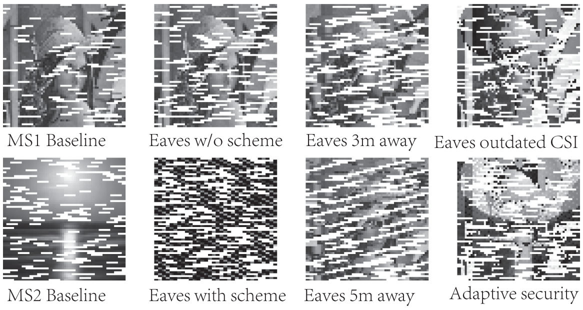


Fig. 5. Eavesdropping attack in different conditions.

TABLE 2
Transmission BER of Different Transmission Scenarios

Scenario	MS_1 baseline	MS_2 baseline	Eaves w/o scheme	Eaves with scheme
BER	0.161	0.222	0.260	0.589
Scenario	Eaves 3m away	Eaves 5m away	Eaves outdated CSI	Adaptive security
BER	0.455	0.471	0.298	0.530

serving two MSs equipped with single antenna respectively. One MS is the attacker while the other one is the victim. The attacker sets up additional radio equipments nearby the BS to keep eavesdropping on messages transmitted to the BS. We first place the eavesdropping equipment nearby the BS and try to recover victim's download. Then we will investigate the effect of relative distance between the BS and attacker's additional eavesdropping equipment. At last, we will apply our adaptive security enhancement.

To be more practical, we have implemented a simple UDP as application layer protocol based on our proposed scheme. Since there are two users in our toy example, the BS will use two antennas serve MS_1 (i.e., the victim) and MS_2 (i.e., the attacker). The other two spare antennas are used to transmit noise to protect transmission of the BS. The BS and MSs are placed in a 6 m $\times$ 6 m area. The victim MS and the attacker mobile station are 3 meters away from the base station respectively but there is no requirement about how far the attacker MS should be away from the victim MS. Since the only requirement is that two MSs should be within the signal coverage of the BS, we place the attacker MS 3 meters away from the victim MS.

First, we use conventional MU-MIMO implementation as a baseline with only artificial noise as additional protection. MSs are downloading two 64 $\times$ 64 pixels images respectively. Since we do not use any error correction, some invalid packets will be dropped directly. The results of our UDP baseline are shown in the first column of Fig. 5. To show the effect of relative distance between the BS and the eavesdropping equipment, we test attacks in three different relative distances. The eavesdropping equipment is placed nearby, 3 meters away from, 5 meters away from the BS. Then eavesdropping results of three conditions are shown at row 1 column 2, row 1 column 3 and row 2 column 3 in Fig. 5 respectively. Eavesdrop result will be damaged while relative distance increases. However, please also note that we are using a toy example in a small area. The relative distance can increase when antenna array scales up. Corresponding transmission BERs of scenarios in Fig. 5 are calculated in Table 2. It should be noted that the BERs calculated in Table 2 have taken loss errors during the transmission into consideration as bit errors after binarization.

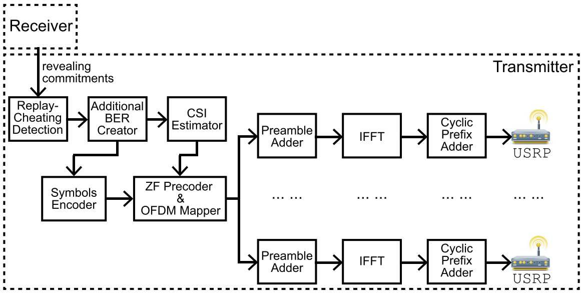


Fig. 6. Primary modules for BS transmitters.

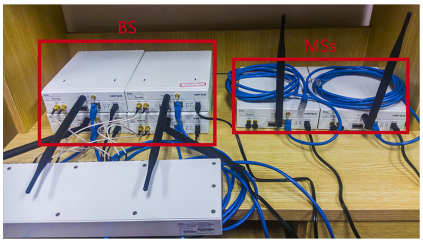


Fig. 7. Demo of the BS and MSs. This BS is constructed with 4 USRP N210. Two MSs are ready to be deployed in other place with laptop connected respectively.

To verify that our secure CSI estimation can protect user's download effectively, we test the eavesdropping attack again after applying the proposed CSI estimation. From the result shown at row 2 column 2 in Fig. 5, we can conclude that the attacker cannot obtain other user's download information correctly. When the attacker uses victim's historical CSI in a short time, partial download information can be decomposed under a relatively stable channel. This situation can be stemmed by our adaptive security scheme. From comparison between two attacks with outdated CSI as shown in column 4 in Fig. 5, we can tell that our adaptive security can prevent the eavesdropping under a relatively stable channel effectively.

7.2 Evaluation of Proposed Scheme

We have implemented secure CSI estimation in the MU-MIMO TDD system with the help of Gnu Radio [50] and USRP [51]. We use 4 USRP N210 to construct the BS with 4 omnidirectional antennas. And each MS is functioned with single USRP N210. An OctoClock-G is used to feed synchronized 10MHz clock and 1 PPS reference to the BS. Data of each USRP N210 is transmitted through Gigabyte cable to a computer for analysis. Fig. 7 shows a demo of the BS and MSs.

Gnu Radio v-3.7.9 is used to build up our MU-MIMO scheme. Original OFDM modules of Gnu Radio will be modified. Some new modules are added to run proposed secure CSI estimation. Primary modules which are needed for BS's transmitter are shown in Fig. 6, including modules that have been modified or added.

In order to verify the feasibility of our secure CSI estimation, we compare the information leakage of downlink with secure estimation and downlink without it. We use the BER of eavesdropping download to measure the information leakage of the victim. Our eavesdropping attack uses a heuristic selection of forged CSI [6] as reference, where attacker

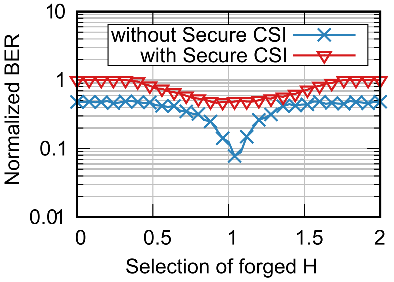


Fig. 8. Eavesdropping results of downlink with secure CSI estimation and downlink without secure CSI estimation.

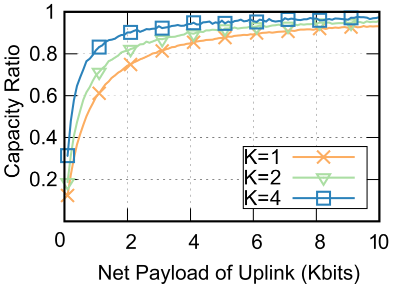


Fig. 10. Capacity ratio is calculated by 1 - (capacity of commitments)/(total capacity) for all MSs.

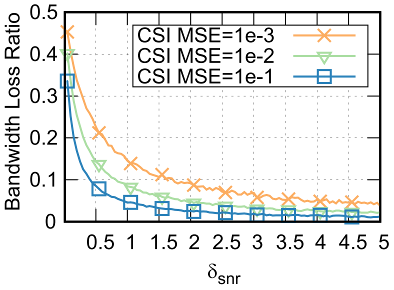


Fig. 9. Loss ratio of bandwidth, which is calculated by (loss of bandwidth)/(total bandwidth). Results are calculated for single MS.

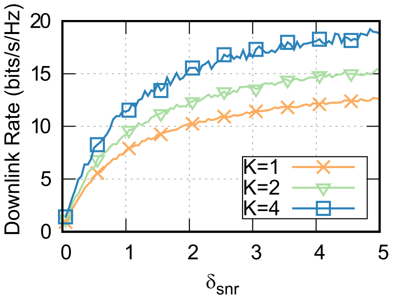


Fig. 11. Download rates of different SNR threshold δ_{snr} are measured for total download links between all MSs and the BS.

MS$_2$'s CSI should be forged to be $w\boldsymbol{H}_1 - \boldsymbol{H}_2$. As for our attack model, we use this forged CSI to calculate corresponding $\Delta\boldsymbol{s}$. Hence, the result of eavesdropping will depend on the selection of weighting parameter w. Results of eavesdropping are shown in Fig. 8. An appreciable eavesdropping can be achieved in unprotected downlink when w is about 1.1. Downlink with secure CSI estimation can avoid this low BER eavesdropping by disturbing the attacker's selection of forged training sequence. In secure CSI estimation scenario, we use previous CSI to extract the victim's download. The BER is so high that the attacker cannot extract precise downloading content of the victim.

When the changing of instantaneous CSI is relatively slow, our adaptive security scheme will do its job. When the adaptive security is working, additional bit errors will be introduced. Definitely, additional bit errors will effect valid bandwidth, but results in Fig. 9 show that loss ratio of bandwidth will decrease significantly when SNR threshold δ_{snr} is low. If we choose $\delta_{snr} = 2$ which is small enough for protection, loss of bandwidth will be lower than 10 percent even in the condition where MSE of CSI between two adjacent intervals (referred as CSI MSE in Fig. 9) equals to 10^{-3}. When MSE of CSI between two adjacent intervals keeps lower than 10^{-3} with MMSE method, we think this situation can be regarded as a statistical CSI case where statistical characterization should be used instead of CSI estimation.

Recall that we have modified the uplink in the first phase. Transmission of commitments of training sequences cannot be negligible. Since commitments are always transmitted with users' payload of uplink, capacity ratio of uplink can be effected by the occupancy of commitments. In our implementation, we use short training sequence which is randomly generated and the commitment of training sequence is always put in the first symbol of of frames for every subcarrier. We have measured the capacity ratio of uplinks between all MSs and the BS. As shown in Fig. 10, the capacity ratio of net payload will reach a nearly saturation if net payload of each MS keeps increasing. It is reasonable that the more MSs there are, the more quickly the capacity ratio will approach to 1.0. But even there is only one MS in the network, the capacity ratio can still approach to 60 percent with 1Kbits net payload.

With secure CSI estimation and adaptive security employed, results of the experiments can prove the feasibility of our scheme. Besides, downlink rate of MU-MIMO is an important evaluation. To find out how much downlink rate can be achieved with our secure CSI estimation and adaptive security employed, we have measured downlink rates with conditions of different SNR threshold δ_{snr}. We will consider power allocation problem in the following section. Before that, we will assume that the BS keeps the same per-client transmitting power. Results are shown in Fig. 11. SNR threshold δ_{snr} has significant effect on downlink rate, because of the loss of bandwidth which is introduced by adaptive security. Results are following a similar pattern in scenarios of different amounts of MSs. If δ_{snr} is set to be 2 as recommended in the discussion of bandwidth loss, a downlink rate of 10bits per symbol per Hz can be achieved even there is only one MS in the network.

Last but not least, we have measured the overhead of computation in CPU cycles. In the first phase, almost all work is done by MSs in the previous interval, which means MSs can do commitments' computation while downloading or waiting. Thus we will focus on the overhead of BS's computation. Fig. 12 shows results of the overhead of computation for two main procedures of second phase: revealing commitments and applying adaptive security. Revealing commitments of training sequences takes most of the overhead. Although the overhead of revealing commitments increases in a nearly linear tendency, the order of

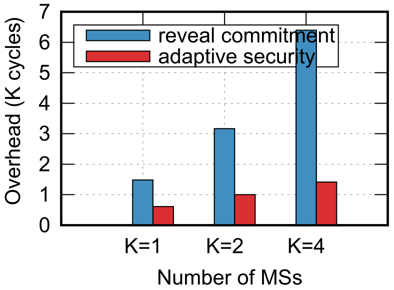


Fig. 12. Overhead of computation for two main procedures of second phase are measured for different K.

magnitude of CPU cycles will keep in 10^3 unless there are about 10^3 MSs, computation can be done in microseconds by a CPU clocked in GHz. This means the overhead introduced by our scheme is acceptable.

8 Optimization of Uplink Design and Power Allocation

The main goal to refine our previous scheme is to extend eavesdropping detection scheme to support massive MIMO scenario. Massive MIMO system has many attractive properties, such as massive antennas, massive users, massive traffics [52]. Those interesting properties also bring potential challenges. Among all of those properties and challenges, we try to deal with massive users in this section. A reported in previous research, the user amount will exceed 1000 in massive MIMO systems [53]. Current researches are aggressively trying to serve 100 times more devices delivering 1000 times [54] traffic than Long Term Evolution (LTE).

In previous work, we use spatial multiple access scheme to implement uplinks. When users amount expands into masses, spatial multiple access will be a real obstacle to quality transmission because of very long queuing upload. To meet the demand of serving massive users, we need improve uplink transmission to achieve higher transmitting rate. To this end, we will refine revealing commitments phase with simultaneous transmission based on imperfect CSI. Then we can bound revealing phase in a short time window where users can upload revealing messages simultaneously so that channels utility can be improved significantly. After that, we will take energy efficiency into consideration. An optimal power allocation scheme for uplink with secure training sequences will be given.

Moving revealing messages into the same coherence interval as commitment means that more data transmission time in the interval will be occupied. Thus how to design real-time revealing phase while ensuring efficient data transmission will be a challenge. Generally, CSI estimation should capture channels changing as soon as possible. However, due to the requirement of transmitting rate and the limitation of energy, estimation cannot be performed all the time. But if estimation frequency lags behind CSI fluctuations, it will be easier to trigger false positives in our eavesdropping detection scheme. Specifically, in interval j, MS$_i$ has commitment $F(c_i, r_i)$, estimated CSI between MS$_i$ and the BS is $\hat{H}$. After data transmission or the guard time, the BS requires all the MSs to send revealing messages within time window T_r. If T_r stays within current coherence interval, the BS can decode all messages correctly. But if the sum of data transmission window and T_r exceeds coherence time, revealing messages or partial revealing messages will be decoded with errors. Our original scheme is error-tolerant, but error correcting capability of $ECC()$ is limited. Thus whether revealing messages can be correctly decoded or not still depends on promoted uplink design.

Given error-correcting capability t_e of $ECC()$, training sequence string length n, the fuzzy commitment scheme should be $\frac{100t_e}{n}\%$ error resilience [7]. For QPSK modulation we used, BER is given by $1/2erfc(\sqrt{\frac{E_b}{N_0}})$. Since Theorem 2 has given the maximum SNR with outdated CSI and BER is inversely proportional to SNR, we can have minimum BER with outdated CSI as P_b. In the worst case where entire revealing window falls behind coherence interval, $\frac{100t_e}{n}\%$ should be no less than P_b if we want to decode all revealing messages correctly. But this requires very strong error-correcting capability. Large t_e may lead to impractical encoded string which is too long to be transmitted in short time window. Our goal in promoted uplink design is to achieve higher data transmission rate, which means we want to minimize revealing window (i.e., revealing message length). The paradox here is, if we use compressed revealing message in the second phase, then it will occupy MS's more uplink transmission in the first phase to transmit more redundancy, which is used to decompress revealing message. Thus, the total bandwidth utility will be reduced.

To solve this problem, our new scheme is designed with the aid of uplink optimization. Assume that for MS$_i$, one part of revealing message r_{i1} is within coherence interval, and the other part r_{i2} is in uncoded channel. Since transmission of r_{i1} has sufficient channel information, assume that r_{i1} can be decoded correctly while r_{i2} needs error-correcting support. When there are two parity-check sets, minimal distance of a low-density parity-check code can be bounded by $D \leq 2 + \frac{2ln\frac{n_e}{2}}{ln(k-1)}$. If channel utilization in coherence interval is regarded as 1, then utilization rate in uncoded channel will be $1 - P_b$. The optimization problem can be described as,

$$\begin{aligned} \min \quad & \|r_{i1}\| + 2n_e + l + (1 - P_b)\|r_{i2}\|. \\ \text{s. t.} \quad & t_e > \frac{ln\frac{n_e}{2}}{ln(k-1)} + \frac{1}{2} \geq P_b\|r_{i2}\|. \\ & \|r_{i1}\| + \|r_{i2}\| = n \geq K. \end{aligned}$$

This optimization problem can be solved simply by some optimization tool since there are few variables. But the result is very interesting. As shown in Fig. 13, three facts can be learned when there is only a small group of MSs. First, when the minimum BER with outdated CSI is relatively low, the objective function can reach the minimum while $\|r_{i1}\|$ is very small. This means when K and P_b are low, most of the revealing message can be arranged outside current coherence interval. Second, when the minimum BER rises up, $\|r_{i1}\|$ gets larger for all MSs scales. This is reasonable because part of r_{i2} should be arranged into coherence time to avoid more redundance caused by the second part of our objective function when uncoded channel has very low SNR. Third, $\|r_{i1}\|$ shows unstable in low BER scenario when K reaches 22. This fluctuation will continue and even grow sharply when K keeps increasing. Thus, $\|r_{i1}\|$ values with regards to different

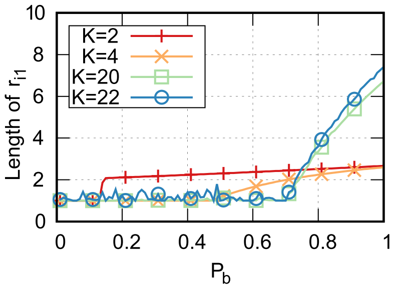


Fig. 13. $\|r_{i1}\|$ has a rough proportion to channel BER when the amount of MSs is small.

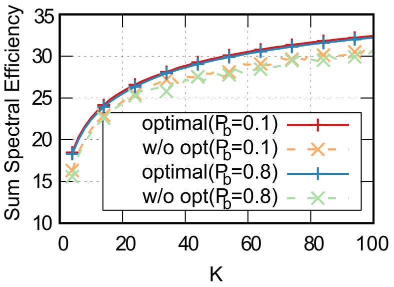


Fig. 15. Sum spectral efficiency with and without power allocation optimization.

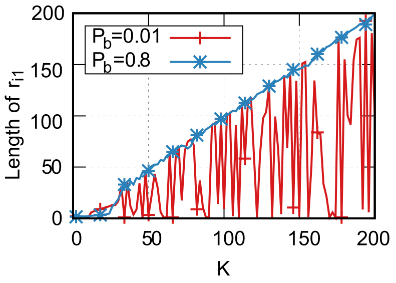


Fig. 14. When the amount of MSs is larger than 23, $\|r_{i1}\|$ is influenced by P_b to reach the minimum of objective function.

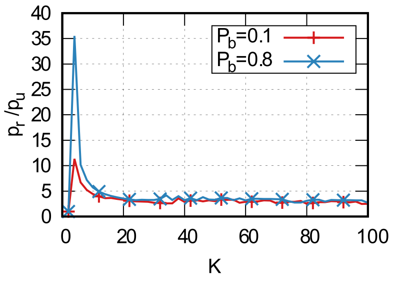


Fig. 16. Optimized power allocation ratio of p_r to p_u.

K are recorded in Fig. 14. When K is greater than 23, $\|r_{i1}\|$ begins to fluctuate to reach the minimum of the objective function for very small P_b. Although this figure just shows two possible values of P_b, we've observed that the fluctuation occurs when the minimum BER is low. When values of P_b is large enough, the fluctuation will disappear and $\|r_{i1}\|$ is approaching its maximum value. The overall tendency indicates that revealing messages should be arranged into coherence interval when the scale of MSs get big and especially when the BER is very high.

In the above optimization procedure, we mainly focus on the arrangement of training sequences and revealing messages while assuming that all MSs use the same transmitting power for training sequences, revealing messages and data. As reported in [55], energy and spectral efficiency of MU-MIMO can be further improved according to specific fading situation and coding scheme. We will use a similar optimization model as in [55] to maximize the sum spectral efficiency by selecting transmitting power for training sequences, revealing messages and data. Generally, we assume that total energy budget for the coherence interval is P_i for MS$_i$. If part of revealing message is arranged outside the coherence interval, it still needs transmission power. So we will also take the power allocation for this part into consideration of the coherence interval. Since the commitments of training sequences are always transmitted along with data, power for commitments is the same with data transmitting power. Assume that coherence interval length is T, data transmitting power is denoted by p_u and revealing messages transmitting power is denoted by p_r. For MS$_i$, we can bound the achievable uplink rate by,

$$\begin{aligned} R_i = \log_2\left(1 + \frac{p_u^2 p_r^2 (M-K)\beta_i^2 \|r_{i2}\|}{(p_u p_r \beta_i \|r_{i2}\| + p_u)\sum_{k=1}^{K} \frac{p_u \beta_k}{p_u \beta_k \|r_{i2}\| + 1} + \dots}\right), \\ \dots p_u p_r \beta_k \|r_{i2}\| + p_u, \end{aligned} \tag{19}$$

where β_i is priori knowledge that models shadow fading. The energy constraint will be $Tp_u + \|r_{r2}\|p_r \leq P_i$. Then we can maximize the sum spectral efficiency by

$$\begin{aligned} \max \quad & \left(1 - \frac{\|r_{i1}\| + n_e + l}{T}\right) \sum\nolimits_{i=1}^{K} R_i. \\ \text{s. t.} \quad & Tp_u + \|r_{i2}\| p_r \leq P_i \\ & p_u \geq 0, p_r \geq 0. \end{aligned}$$

Generally, 12 to 20 OFDM symbols are transmitted in one coherence interval. Assume that an OFDM symbol duration T_s is 70 μs. Then, we have $T = 1.4$ms for 20 symbols in coherence interval. Since $\|r_{i1}\|$ and $\|r_{i2}\|$ will be influenced by K seriously, we will also study this optimization problem in different MSs scales. Besides, we use fitted tendency to cancel the fluctuation of $\|r_{i1}\|$ when K is large. Similar environment settings are used here as in [55]. However we reduce M from 400 to 200 to avoid more unstable r_{i1} values. Optimization results are shown in Fig. 15. The bandwidth efficiency after optimization is evaluated for all users in different scales as sum spectral efficiency. Given total power P_i for MS$_i$, optimized power allocation can always achieve higher spectral efficiency than equally allocated power strategy. Much higher spectral efficiency can be achieved when K is larger. In Fig. 16, details of power allocation for p_r and p_u are recorded when the objective function is optimal. p_r is transmitting power for revealing messages while p_u is transmitting power for data. It is obvious that p_r needs much more power when P_b is high. However the overall allocation tendencies for $P_b = 0.1$ and $P_b = 0.8$ are very similar when K increases. This is because of the sharp increasing of $\|r_{i1}\|$ when K gets large.

9 Security Enhancement for Mismatch Channel

In TDD operation system, the channel reciprocity is the most important assumption. However, as reported in recent research work [8], [9], [56], the channel reciprocity assumption will be impaired when radio frequency (RF) circuit mismatch is taken into account. Recently many efforts have been done to study RF circuit mismatch in MU-MIMO networks. [10], [11], [12], [57]. The eavesdropping attack and the countermeasure proposed in our work are highly depended on the channel reciprocity. According to the feedback and our observations, the RF mismatch has significant impact on work. Thus, in this section we will introduce an enhancement scheme for mismatch channel. We will analyse the security of our enhancement scheme under RF mismatch channel. We will show that the attacker will be further weakened by our enhancement scheme under RF mismatch channel.

9.1 Estimation Enhancement Scheme

As reported in [58], [59], [60], [61], Channel mismatch in OFDM TDD systems will lead to serious destruction of channel reciprocity. This can be reflected in signal delay, phase mismatch, frequency mismatch and even all of these together. Hence, we will investigate how to ensure the correctness of our secure estimation scheme under a mismatch channel model which has been studied for various purposes [13], [62], [63]. When the unique characteristics of RF circuits in transmit and receive modules of both the BS and MS are taken into consideration, the whole communication channel will be composed of the propagation channel and the RF circuits at both the BS and MSs. Instead of Equation (1), the whole channel of downlink and uplink between the BS and MSs should now be expressed as,

$$\begin{aligned} \boldsymbol{G}_D &= \boldsymbol{U}_r \boldsymbol{H}_D \boldsymbol{B}_t, \\ \boldsymbol{G}_U &= \boldsymbol{B}_r \boldsymbol{H}_U \boldsymbol{U}_t, \end{aligned}$$

where $\boldsymbol{B}_t, \boldsymbol{B}_r$ are matrices of complex gains which are introduced by the BS's transmit and receive RF circuits respectively, $\boldsymbol{U}_t, \boldsymbol{U}_r$ are matrices of complex gains introduced by MS's transmit and receive RF circuits respectively, $\boldsymbol{H}_D$ and $\boldsymbol{H}_U$ are wireless channel propagation coefficients of downlink and uplink respectively. $\boldsymbol{B}_t, \boldsymbol{B}_r, \boldsymbol{U}_t, \boldsymbol{U}_r$ are diagonal matrices. Since in TDD system, the BS uses reciprocity of uplink training to encode downlink streams, the relation between downlink channel and uplink channel will be,

$$\boldsymbol{G}_D = \boldsymbol{U}_r \boldsymbol{U}_t^{-1} \boldsymbol{G}_U^T \boldsymbol{B}_r^{-1} \boldsymbol{B}_t. \tag{20}$$

Commonly, we assume that $\boldsymbol{H}_D = \boldsymbol{H}_U^T$ like we do in original analysis because of the channel reciprocity. But the RF circuits are mismatched, which means $\boldsymbol{U}_r \neq \boldsymbol{U}_t, \boldsymbol{B}_r \neq \boldsymbol{B}_t$. Then $\boldsymbol{G}_D \neq \boldsymbol{G}_U^T$. To address this problem, recent work has offered detailed analysis of MU-MIMO system under mismatch channel [10], [11], [12]. Based on these studies, a more robust precoding scheme or a more accurate estimation will be needed to conquer RF mismatch. Thus, we aim to propose an enhancement scheme for more accurate estimation based on our original secure estimation scheme.

To analyse the security of the enhancement scheme, we will use the same system model and threat model as original scheme. Recall our secure estimation scheme, when the commitment of training sequence is revealed, the linear MMSE estimator will be applied [64] to estimate uplink channel coefficients $\boldsymbol{G}_U$. Then the downlink channel estimation can be derived by the relation between uplink and downlink, i.e., $\hat{\boldsymbol{G}}_D = \boldsymbol{U}_r \boldsymbol{U}_t^{-1} \hat{\boldsymbol{G}}_U^T \boldsymbol{B}_r^{-1} \boldsymbol{B}_t$. Since RF mismatch coefficients vary slowly in time comparing with the propagation coefficients, it can be assumed that $\boldsymbol{B}_t, \boldsymbol{B}_r, \boldsymbol{U}_t, \boldsymbol{U}_r$ keep constant over several coherence intervals. It's also obvious that uplink and downlink are reciprocal when $\boldsymbol{U}_r = \boldsymbol{U}_t, \boldsymbol{B}_r = \boldsymbol{B}_t$. It has been proved that the downlink channel estimation error of kth MS can be modeled as [10],

$$e_k = \sqrt{M} B_t \boldsymbol{\Phi}_k^{\frac{1}{2}} \boldsymbol{e}_{vk} \boldsymbol{U}_{rk}, \tag{21}$$

where $\boldsymbol{\Phi}$ is wireless channel correlation matrix which is perfectly know by the BS, e_{vk} follows independent Gaussian distribution with zero-mean and variance $\frac{1}{M}\boldsymbol{I}$. Then the relation between downlink channel coefficients and estimation error can be determined by $\boldsymbol{G}_D = \hat{\boldsymbol{G}}_D + \boldsymbol{E}$, where $\boldsymbol{E} = [e_1, e_2, \ldots, e_K]^T$.

In the revealing phase, the real training sequences will be collected by the BS. The whole uplink channel estimation $\hat{\boldsymbol{G}}_U$ can be obtained by MMSE directly. Our secure estimation is based on linear ECC to recover benign training sequence. But to ensure that the commitment of training sequence can be correctly revealed under mismatch channel, the error introduced by $\boldsymbol{U}_t, \boldsymbol{B}_r$ must be taken into account. As reported in [8], [10], the RF circuit coefficients can be seen as known constants at the BS and the corresponding MS. The received training sequence at the BS from the ith MS will be $r_i \boldsymbol{g}_U + \boldsymbol{z}$, where $\boldsymbol{g}_U$ is the corresponding vector for MS_i from $\boldsymbol{G}_U$, $\boldsymbol{z}$ is the independently additive white Gaussian noise. Due to the existence of RF mismatch, the received training sequence $r_i \boldsymbol{B}_r \boldsymbol{H}_U \boldsymbol{U}_t$ cannot reveal the commitment which has been committed by MS_i for r_i. Because when ECC will try to recovery $r_i \boldsymbol{B}_r \boldsymbol{H}_U \boldsymbol{U}_t$ to be r_i, $\boldsymbol{B}_r, \boldsymbol{U}_t$ will change the distance between r_i and its codeword significantly. To address this problem, we use different training sequence design from the original scheme.

Algorithm 4. Enhanced Secure CSI Estimation Scheme Under Mismatch Channel

```
Commitment Phase:
for all $MS_i, i \in [1, K]$ do
    $MS_i$ generates random training sequence $r_i$,
    $MS_i$ computes $o_i = r_i \boldsymbol{B}_r \boldsymbol{U}_t$,
    $MS_i$ chooses codeword $c_i \in C$,
    $MS_i$ computes $F(c_i, o_i) = (\alpha_i, \delta_i) = (h(c_i), o_i - c_i)$.
    $MS_i$ sends $F(c_i, o_i)$ to the BS.
end for
Revealing Phase:
for all $MS_i, i \in [1, K]$ do
    the BS collects revealing message $r'_i$,
    if $\alpha_i = h(f(r'_i - \delta_i))$ then
        The BS recover $o_i = f(r'_i - \delta_i) + \delta_i$,
        The BS recover $r_i = \boldsymbol{B}_r^{-1} \boldsymbol{U}_t^{-1} o_i$,
    end if
end for
for all $r_i, i \in [1, K]$ do
    The BS uses $r_i$ to do replicate cheating detection,
    The BS estimates $\hat{\boldsymbol{H}}_D^T \leftarrow \hat{\boldsymbol{H}}_U^{MMSE}(o_i, r'_i)$.
    The BS estimates the whole downlink channel $\hat{\boldsymbol{G}}_D =$
    $\boldsymbol{U}_r \boldsymbol{U}_t^{-1} \hat{\boldsymbol{H}}_D^T \boldsymbol{B}_r^{-1} \boldsymbol{B}_t$.
end for
Adaptive security module.
```

As shown in the Algorithm 4, the ith MS will commit $o_i = r_i \boldsymbol{B}_r \boldsymbol{U}_t$ rather than simply random sequence r_i in the commitment phase. As a result, the BS should recover o_i first in the revealing phase. Then o_i and r_i' are used to estimate uplink channel propagation. Finally, the whole downlink coefficients can be obtained by the relation expressed in (20).

9.2 Security Analysis under Mismatch Channel

Let A_r, A_t denote the coefficients of receiver and transmitter of the eavesdropping RF. We still assume that MS_2 is coveting downlink message that MS_1 is downloading from the BS. Then, *in the attacker's view*, r_1', $F(c_1, o_1)$, $\boldsymbol{B}_r$, $\boldsymbol{B}_t$, $\boldsymbol{A}_r$, $\boldsymbol{A}_t$, $\boldsymbol{U}_{2r}$, $\boldsymbol{U}_{2t}$ are known. Specifically, r_1' eavesdropped by the attacker should be $r_1 \boldsymbol{A}_r \boldsymbol{h}_{1U} \boldsymbol{U}_{1t}$, and is one coherence interval ahead of the corresponding $F(c_1, o_1)$. We assume that the attacker can record all the commitments that have been received by the BS. In our enhancement scheme, if the attacker aims to eliminate the influence of RF circuits mismatch totally, the RF circuits coefficients must be known to the attacker. This is a rather strong assumption because it's almost impossible for the attacker to measure the RF circuits coefficients while all training sequences are protected by commitment scheme. Besides, the RF that the attacker uses to eavesdrop on the BS should be taken into consideration too.

For instantaneous CSI, the enhancement scheme can protect all MSs perfectly. This conclusion can be derived from the original scheme. It will be different when we analyse the security of our enhancement scheme under mismatch channel. To reveal MS_1's training sequence from $r_1 \boldsymbol{A}_r \boldsymbol{h}_{1U} \boldsymbol{U}_{1t}$, MS_2 needs MS_1's RF coefficients $\boldsymbol{U}_{1t}$. Without MS_1's RF coefficients, MS_2 cannot obtain correct codeword from $f(r_1' - \delta_1)$. It means that $\alpha_1 = h(f(r_1' - \delta_1))$ will not hold in the adversary's view. We assume that hash function h is secure. Hence, MS_2 cannot recover MS_1's benign training sequence by $r_1 = \boldsymbol{A}_r^{-1} \boldsymbol{U}_1^{-1}(f(r_1' - \delta_1) + \delta_1)$. Without training sequence r_1, MS_2 cannot estimate $\boldsymbol{h}_{1U}$ correctly.

For the relatively statistical CSI, the original estimation scheme may be broken because the attacker can obtain approximate CSI by eavesdropping. That's why we integrate adaptive security for the statistical CSI situation. As for the security analysis of enhancement scheme, we also assume that the adversary is capable of obtaining target's approximate CSI $\hat{\boldsymbol{h}}_{1U}$. Then the downlink CSI for MS_1 at the BS should be $\boldsymbol{U}_{1r} \boldsymbol{U}_{1t}^{-1} \hat{\boldsymbol{h}}_{1U}^T \boldsymbol{B}_r^{-1} \boldsymbol{B}_t$. So MS_1's download in mixed signals will be precoded by MS_1's download CSI at the BS. When MS_2 tries to decompose MS_1's download in received signals, MS_2 should consider RF mismatch of MS_1, MS_2 and the BS. The error rate of MS_2's decomposing result will be increased. Because MS_1's download in mixed signals is precoded by $\boldsymbol{B}_r, \boldsymbol{B}_t, \boldsymbol{U}_{1r}, \boldsymbol{U}_{1t}$ and $\boldsymbol{h}_{1D}$ while MS_2 cannot eliminate RF mismatch entirely even that we assume $\hat{\boldsymbol{h}}_{1U}$ is available in statistical CSI situation. This will result in a lower δ_{BER} to be introduced in our adaptive security module, which means that our enhancement scheme under RF mismatch channel actually weakens the adversary in statistical CSI situation.

10 Conclusions

In this paper, we have proposed a feasible approach to launch an aggressive eavesdropping attack against TDD MU-MIMO network with implicit CSI estimation. Based on the transforming of training sequences, we successfully implement eavesdropping attack in real world. To eliminate this threat, we have designed secure CSI estimation scheme for instantaneous CSI condition. With the help of fuzzy commitment scheme, our countermeasure can effectively stop the attacker from eavesdropping. Then, considering the situation where CSI may change slowly, we integrate an adaptive security approach to our scheme. This approach can guarantee users' security in different CSI scenarios. Since additional bit errors are introduced dynamically, the integrated scheme will cause extra overhead and bandwidth loss. However, it has been proved acceptable even in some undesirable conditions.

In order to deploy our scheme for massive MIMO networks, we redesign uplink scheme to achieve higher transmission rate by bounding revealing phase in a short time window. Furthermore, to meet the energy budget of mobile devices, we also optimize power allocation for our secure training sequence scheme. Combined two optimization procedures, we conclude that revealing phase should be arranged outside coherence interval when MSs scale is small and Power allocated for revealing messages should be higher than data in coherence interval, especially when SNR is relatively low. Last, when the reciprocity of channels is broken by RF mismatch in TDD system, implicit CSI estimation scheme will be at stake. To guarantee the correctness and security of our CSI estimation scheme, we have introduced an enhancement which can deal with mismatch channel situation and secure CSI estimation as usual.

Acknowledgments

The authors would like to thank the editor and the reviewers for the time and efforts they have kindly made on this paper. Their insightful comments help us make this paper more comprehensive and precise. This work was supported in part by National Key R&D Program of China (2018YFB1004301), NSFC-61872179, NSFC-61300235, NSFC-61425024, and NSFC-61872176. A preliminary version of this work was presented at CCS 2016 [1].

References

[1] Y. Mao, Y. Zhang, and S. Zhong, "Stemming downlink leakage from training sequences in multi-user MIMO networks," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2016, pp. 1580–1590.

[2] Y. Wu, C. Xiao, Z. Ding, X. Gao, and S. Jin, "A survey on MIMO transmission with finite input signals: Technical challenges, advances, and future trends," in *Proc. IEEE*, Oct. 2018, vol. 106, no. 10, pp. 1779–1833.

[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots," in *Proc. 51st Annu. Allerton Conf. Commun. Control Comput.*, 2013, pp. 293–298.

[4] P. Ting, C. K. Wen, and J. T. Chen, "An efficient CSI feedback scheme for MIMO-OFDM wireless systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2012–2015, Jun. 2007.

[5] R. Abu-alhiga and H. Haas, "Implicit pilot-borne interference feedback for multiuser MIMO TDD systems," in *Proc. IEEE 10th Int. Symp. Spread Spectr. Techn. Appl.*, 2008, pp. 334–338.

[6] Y.-C. Tung, S. Han, D. Chen, and K. G. Shin, "Vulnerability and protection of channel state information in multiuser mimo networks," in *Proc. ACM SIGSAC Conf. Comput. Commun. Security*, 2014, pp. 775–786.

[7] A. Juels and M. Wattenberg, “A fuzzy commitment scheme,” in *Proc. 6th ACM Conf. Comput. Commun. Security*, 1999, pp. 28–36.

[8] W. Zhang, H. Ren, C. Pan, M. Chen, R. C. de Lamare, B. Du, and J. Dai, “Large-scale antenna systems with UL/DL hardware mismatch: Achievable rates analysis and calibration,” *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1216–1229, Apr. 2015.

[9] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, “Relative channel reciprocity calibration in MIMO/TDD systems,” in *Proc. Future Netw. Mobile Summit*, Jun. 2010, pp. 1–10.

[10] Y. Chen, X. Gao, X. G. Xia, and L. You, “A robust precoding for RF mismatched massive MIMO transmission,” in *Proc. IEEE Int. Conf. Commun.*, May 2017, pp. 1–6.

[11] O. Raeesi, A. Gokceoglu, Y. Zou, E. Björnson, and M. Valkama, “Performance analysis of multi-user massive MIMO downlink under channel non-reciprocity and imperfect CSI,” *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2456–2471, Jun. 2018.

[12] X. Wang, Y. Wang, W. Ni, R. Sun, and S. Meng, “Sum rate analysis and power allocation for massive MIMO systems with mismatch channel,” *IEEE Access*, vol. 6, pp. 16997–17009, Mar. 2018.

[13] D. Mi, M. Dianati, L. Zhang, S. Muhaidat, and R. Tafazolli, “Massive MIMO performance with imperfect channel reciprocity and channel estimation error,” *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3734–3749, Sep. 2017.

[14] H. Wei, D. Wang, H. Zhu, J. Wang, S. Sun, and X. You, “Mutual coupling calibration for multiuser massive MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 606–619, Jan. 2016.

[15] M. Jiang and L. Hanzo, “Multiuser MIMO-OFDM for next-generation wireless systems,” *Proc. IEEE*, vol. 95, no. 7, pp. 1430–1469, Jul. 2007.

[16] R. Liao, B. Bellalta, M. Oliver, and Z. Niu, “MU-MIMO MAC protocols for wireless local area networks: A survey,” *IEEE Commun. Surveys Tutorials*, vol. 18, no. 1, pp. 162–183, Jan.-Mar. 2016.

[17] A. Mukhopadhyay, N. B. Mehta, and V. Srinivasan, “Acknowledgement-aware MPR MAC protocol for distributed WLANs: Design and analysis,” in *Proc. IEEE Global Commun. Conf.*, 2012, pp. 5087–5092.

[18] K. Tan, H. Liu, J. Fang, W. Wang, J. Zhang, M. Chen, and G. M. Voelker, “SAM: Enabling practical spatial multiple access in wireless LAN,” in *Proc. 15th Annu. Int. Conf. Mobile Comput. Netw.*, 2009, pp. 49–60.

[19] T. H. Lin and H. T. Kung, “Concurrent channel access and estimation for scalable multiuser MIMO networking,” in *Proc. IEEE INFOCOM*, 2013, pp. 140–144.

[20] W. L. Huang, K. B. Letaief, and Y. J. Zhang, “Joint channel state based random access and adaptive modulation in wireless lans with multi-packet reception,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4185–4197, Nov. 2008.

[21] T. Tandai, H. Mori, K. Toshimitsu, and T. Kobayashi, “An efficient uplink multiuser MIMO protocol in IEEE 802.11 WLANs,” in *Proc. IEEE 20th Int. Symp. Pers. Indoor Mobile Radio Commun.*, 2009, pp. 1153–1157.

[22] H. Li, K. Wu, Q. Zhang, and L. M. Ni, “CUTS: Improving channel utilization in both time and spatial domain in WLANs,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 6, pp. 1413–1423, Jun. 2014.

[23] H. Shirani-Mehr, D. N. Liu, and G. Caire, “Channel state prediction, feedback and scheduling for a multiuser MIMO-OFDM downlink,” in *Proc. 42nd Asilomar Conf. Signals Syst. Comput.*, 2008, pp. 136–140.

[24] J. Choi, D. J. Love, and P. Bidigare, “Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 802–814, Oct. 2014.

[25] M. Kobayashi, N. Jindal, and G. Caire, “Training and feedback optimization for multiuser MIMO downlink,” *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug. 2011.

[26] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, “Secure massive MIMO transmission with an active eavesdropper,” *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 3880–3900, Jul. 2016.

[27] S. Fang, Y. Liu, W. Shen, and H. Zhu, “Where are you from?: Confusing location distinction using virtual multipath camouflage,” in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 225–236.

[28] B. Akgun, M. Krunz, and O. Ozan Koyluoglu, “Vulnerabilities of Massive MIMO Systems to Pilot Contamination Attacks,” in *IEEE Trans. Inf. Forensics Secur.*, May 2019, vol. 14, no. 5, pp. 1251–1263.

[29] U. Schilcher, C. Bettstetter, and G. Brandner, “Temporal correlation of interference in wireless networks with rayleigh block fading,” *IEEE Trans. Mobile Comput.*, vol. 11, no. 12, pp. 2109–2120, Dec. 2012.

[30] M. Costa, “Writing on dirty paper (corresp.),” *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[31] M. Biguesh and A. Gershman, “Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals,” *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[32] D. Schurmann and S. Sigg, “Secure communication based on ambient audio,” *IEEE Trans. Mobile Comput.*, vol. 12, no. 2, pp. 358–370, Feb. 2013.

[33] G. Revadigar, C. Javali, W. Xu, A. V. Vasilakos, W. Hu, and S. Jha, “Accelerometer and fuzzy vault-based secure group key generation and sharing protocol for smart wearables,” *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 10, pp. 2467–2482, Oct. 2017.

[34] S. Barman, A. K. Das, D. Samanta, S. Chattopadhyay, J. J. P. C. Rodrigues, and Y. Park, “Provably secure multi-server authentication protocol using fuzzy commitment,” *IEEE Access*, vol. 6, pp. 38 578–38 594, Jul. 2018.

[35] E. J. C. Kelkboom, J. Breebaart, T. A. M. Kevenaar, I. Buhan, and R. N. J. Veldhuis, “Preventing the decodability attack based cross-matching in a fuzzy commitment scheme,” *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 107–121, Mar. 2011.

[36] J. Ma, R. Shrestha, J. Adelberg, C.-Y. Yeh, Z. Hossain, E. Knightly, J. M. Jornet, and D. M. Mittleman, “Security and eavesdropping in terahertz wireless links,” *Nature*, vol. 563, no. 7729, 2018, Art. no. 89.

[37] S. Chang and Y. Hu, “SecureMAC: Securing wireless medium access control against insider denial-of-service attacks,” *IEEE Trans. Mobile Comput.*, vol. 16, no. 12, pp. 3527–3540, Dec. 2017.

[38] Y. Li, L. Xiao, H. Dai, and H. V. Poor, “Game theoretic study of protecting MIMO transmissions against smart attacks,” in *Proc. IEEE Int. Conf. Commun.*, 2017, pp. 1–6.

[39] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, “Safeguarding 5G wireless communication networks using physical layer security,” *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, Apr. 2015.

[40] M. Ghaderi, D. Goeckel, A. Orda, and M. Dehghan, “Minimum energy routing and jamming to thwart wireless network eavesdroppers,” *IEEE Trans. Mobile Comput.*, vol. 14, no. 7, pp. 1433–1448, Jul. 2015.

[41] D. Guo, Y. Wu, S. S. Shitz, and S. Verdu, “Estimation in gaussian noise: Properties of the minimum mean-square error,” *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2371–2385, Apr. 2011.

[42] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, “Scheduling and pre-conditioning in multi-user MIMO TDD systems,” in *Proc. IEEE Int. Conf. Commun.*, 2008, pp. 4100–4105.

[43] K. Appaiah, A. Ashikhmin, and T. L. Marzetta, “Pilot contamination reduction in multi-user TDD systems,” in *Proc. IEEE Int. Conf. Commun.*, 2010, pp. 1–5.

[44] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, “Analysis of transmit-receive diversity in rayleigh fading,” *IEEE Trans. Commun.*, vol. 51, no. 4, pp. 694–703, Apr. 2003.

[45] Y. Chen and C. Tellambura, “Performance analysis of maximum ratio transmission with imperfect channel estimation,” *IEEE Commun. Lett.*, vol. 9, no. 4, pp. 322–324, Apr. 2005.

[46] H. Viswanathan and J. Balakrishnan, “Space-time signaling for high data rates in EDGE,” *IEEE Trans. Veh. Technol.*, vol. 51, no. 6, pp. 1522–1533, Nov. 2002.

[47] R. Gallager, “Low-density parity-check codes,” *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, 1962.

[48] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, “Design of capacity-approaching irregular low-density parity-check codes,” *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[49] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels,” *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[50] “Gnu radio,” 2016. [Online]. Available: http://gnuradio.org/redmine/projects/gnuradio/wiki

[51] “Usrp,” 2016. [Online]. Available: https://www.ettus.com/

[52] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive MIMO: Benefits and challenges,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[53] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[54] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, “The role of small cells, coordinated multipoint, and massive MIMO in 5G,” *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 44–51, May 2014.

[55] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, “Energy and spectral efficiency of very large multiuser MIMO systems,” *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[56] O. Raeesi, Y. Zou, A. Tölli, and M. Valkama, "Closed-form analysis of channel non-reciprocity due to transceiver and antenna coupling mismatches in multi-user massive MIMO network," in *Proc. IEEE Globecom Workshops*, Dec. 2014, pp. 333–339.

[57] C. Shan, Y. Zhang, L. Chen, X. Chen, and W. Wang, "Performance analysis of large scale antenna system with carrier frequency offset, quasi-static mismatch and channel estimation error," *IEEE Access*, vol. 5, pp. 26 135–26 145, 2017.

[58] J.-C. Guey and L. D. Larsson, "Modeling and evaluation of MIMO systems exploiting channel reciprocity in TDD mode," in *Proc. IEEE 60th Veh. Technol. Conf.*, 2004, vol. 6, pp. 4265–4269.

[59] A. D. Dabbagh and D. J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," *IEEE Trans. Commun.*, vol. 56, no. 11, pp. 1859–1868, Nov. 2008.

[60] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[61] M. Petermann, M. Stefer, F. Ludwig, D. Wubben, M. Schneider, S. Paul, and K. Kammeyer, "Multi-user pre-processing in multi-antenna OFDM TDD systems with non-reciprocal transceivers," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3781–3793, Sep. 2013.

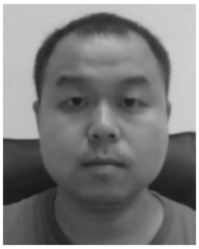
[62] Y. Zou, O. Raeesi, R. Wichman, A. Tolli, and M. Valkama, "Analysis of channel non-reciprocity due to transceiver and antenna coupling mismatches in TDD precoded multi-user MIMO-OFDM downlink," in *Proc. IEEE 80th Veh. Technol. Conf.*, 2014, pp. 1–7.

[63] O. Raeesi, A. Gokceoglu, and M. Valkama, "Estimation and mitigation of channel non-reciprocity in massive MIMO," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2711–2723, May 2018.

[64] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

Yunlong Mao received the BS and PhD degrees in computer science from Nanjing University, in 2013 and 2018, respectively. He is currently an assistant researcher with the Department of Computer Science and Technology in Nanjing University. His current research interests include security, privacy and machine learning.

Ying He receive the BS degree in computer science from Nanjing University, in 2019. She will become a doctoral student in computer science in 2019. Her research interests include security, privacy, and network transmission.



Yuan Zhang received the BS degree in automation from Tianjin University, in 2005, the MSE degree in software engineering from Tsinghua University, in 2009, and the PhD degree in computer science from the State University of New York at Buffalo, in 2013. His current research interests include security, privacy, and economic incentives.

Jingyu Hua received the BE and ME degrees in software engineering from the Dalian University of Technology, China, in 2007 and 2009, respectively, and the PhD degree in informatics from Kyushu University, Japan, in 2012. His current research interests include security and privacy in mobile computing, and system security.

Sheng Zhong received the BS and MS degrees in computer science from Nanjing University, in 1996 and 1999, respectively, and the PhD degree in computer science from Yale University, in 2004. He is interested in security, privacy, and economic incentives.