PHYSICAL LAYER ANONYMOUS COMMUNICATIONS: AN ANONYMITY ENTROPY ORIENTED PRECODING DESIGN

Zhongxiang Wei †, Christos Masouros ‡ and Sumei Sun *

† College of Electronic and Information Engineering, Tongji University, Shanghai, China
‡ Department of Electronic and Electrical Engineering, University College London, London, UK
* Institute of Infocomm Research, Agency for Science, Technology and Research, Singapore.

ABSTRACT

Different from traditional security-oriented designs, the aim of anonymizing techniques is to mask users’ identities during communication, thereby providing users with unidentifiability and unlinkability. The existing anonymizing techniques are only designated at upper layers of networks, ignoring the risk of anonymity leakage at physical layer (PHY). In this paper, we address the PHY anonymity design with focus on a typical uplink scenario where the receiver is equipped with more antennas than the sender. With the increased degrees-of-freedom at the receiver side, we first propose a maximum likelihood estimation (MLE) signal trace-back detector, which only analyzes the signaling pattern of the received signal to disclose the sender’s identity. Accordingly, an anonymity entropy anonymous (AEA) precoder is proposed, which manipulates the transmitted signalling pattern to counteract the receiver’s trace-back detector and meanwhile to guarantee high receive signal-to-interference-plus-noise ratio for communication. More importantly, more data streams can be multiplexed than the number of transmit antennas, which is particularly suitable for the strong receiver configuration. Simulation demonstrates that the proposed AEA precoder can simultaneously provide high anonymity and communication performance.

Index Terms— Signal Trace-back Detection, Anonymous Precoding, Physical Layer Anonymity, Strong Receiver

1. INTRODUCTION

In the era of big data, threats arise from two aspects, namely security and privacy [1]. The target of security is to keep the signal unbreakable towards external adversaries [2], and a series of secure techniques have also been proposed, including but not limited to encryption [3], authentication [4], secure precoding [5], artificial noise [6], and other systematic secure protocols, from the upper layers to the PHY layer of the networks. Nevertheless, in the emerging edge/cloud computing and storage, users need to disclose its data towards trusted but curious receivers, in order to receive utility [7]. Hence, departing from the conventional secure design, the aim of privacy protection is to mask part of data [8], or conceal users’ identities during communication [9] [10]. The latter one, which is the focus of this work, is known as anonymous communications. An example of the anonymous communications arises from e-Health communications, where patients wish to get their physiological signal analyzed by the medical edge, with the patients’ identities kept unknown.

Though a number of anonymizing approaches has been proposed, from the media access control (MAC) layer to the application layer, there are still open challenges. i) The anonymous authentication and encryption techniques [11] [12] use pseudo ID or anonymous account index, instead of the users’ real IP/MAC addresses, for user authentication and data encryption. As a result, the curious receiver is unable to identify the origin of the received signal during authentication and encryption. However, they are only able to mask users’ identities during authentication and data encryption at the upper layers, ignoring the threat in the subsequent signal transmission phases. ii) The agent-based anonymous protocols deploy external cooperative nodes for data relaying, such as the well-known onion router [13] [14], where the receiver is unaware of the signal source and routing paths. However, its anonymity is breakable by the agents along the routing path. iii) The threat of identity-leaking at the PHY has not been addressed. This threat starts from the acquisition of signal. Since the PHY contains critical information that can be utilized to extract the users’ identities, the receiver can analyze the signaling patterns to unmask the users at the PHY directly. PHY layer anonymizing techniques have been investigated in our previous work [9]. A strong sender case was considered, where the receiver has less antennas than the users. Due to the low degrees-of-freedom (DoF)s at the receiver, the detection error rate (DER) of the signal trace-back detector demonstrates a strong error floor (Fig. 1 in Ref. [9]). In practice, the receiver usually has more antennas than the users at uplink, and thus a higher DoF for detection enables a more accurate signal trace-back mechanism, making the provision of the user’s anonymity more challenging.

To address the challenges, in this paper we investigate the signal trace-back detection and its counterpart anonymous
precoder with the strong receiver configuration, where the main contributions are summarized as follows. i) Focusing on the strong receiver configuration at uplink, we first propose a maximum likelihood estimation (MLE) based signal track-back detector. It only exploits the PHY information of the received signal to break the sender’s anonymity, and the DER performance of the MLE based detector is significantly enhanced over that in [9]. ii) To counteract the receiver’s enhanced detection performance, we further propose a novel anonymity entropy based anonymous (AEA) precoder, which manipulates the dissipated signaling pattern and meanwhile provides high receive signal-to-interference-plus-noise-ratio (SINR) for communication. To be specific, based on the concept of interference exploitation, the intended signal is designated to be located at constructive regions, instead of being in the proximity region around the constellation point. Hence, it enables multiplexing more data streams than the number of transmit antennas, and also utilizes inter-antenna interference as a beneficial element to enhance receive quality without loss of the sender’s anonymity.

2. SYSTEM MODEL AND ANONYMITY METRIC

We consider a multiple-input and multiple-output (MIMO) system at uplink, where $K$ users anonymously transmit signals to an access point (AP) without revealing their identities. The AP is equipped with $N_r$ receive antennas while each user has $N_t$ transmit antennas ($N_r > N_t$). In the training phase, all the active users send pilot signals to the AP and channel estimation is performed at the AP side; then the channel state information (CSI) is fed back to the users for use in precoding design, as that in generic MIMO communications. It is important to note however that the CSI estimation process does not jeopardize the anonymity aims of our work. It is because the aim of our work is to obstruct the AP, that has all users’ IDs, from linking the data received to the correct user ID. Assume a time-division-multiple-access fashion and the $k$-th user as the sender without loss of generality [15]. Define $H_k \in \mathbb{C}^{N_r \times N_t}$ as the MIMO channel between the $k$-th user and AP, $W_k$ as the precoder, and $s_k$ as the symbols to be transmitted. The received signal at the AP is given by

$$y = H_k W_k s_k + z,$$

where $z \sim (0, \sigma^2 I_{N_r})$ denotes the Gaussian noise at the AP.

Anonymity can be quantified by an entropy-based metric [16], which in fact measures the uncertainty of a system. Define a set $\mathbb{K} = \{K\}$, and let $p_k$ denote the probability that the AP estimates the $k$-th user as the real sender. Hence, the anonymity entropy can be calculated as $A(\mathbb{K}) = -\sum_{k \in \mathbb{K}} p_k \log_2 p_k$. Evidently, a favorable anonymous precoder needs to inhibit the AP’s detection, while guaranteeing reasonable SINR quality for communication. In the following, we will first design the signal trace-back detector at the AP in Section III, and then the anonymous precoder at the sender is given in Section IV.

3. SIGNAL TRACE-BACK DETECTOR DESIGN

The signal trace-back detection is formulated as a multiple hypotheses testing (MHT) problem at the AP-side, given by

$$\gamma = \begin{cases} H_0 : & z, \\ H_1 : & H_1 W_1 s_1 + z, \\ \vdots \\ H_K : & H_K W_K s_K + z, \end{cases}$$

where the hypothesis $H_0$ means that only noise exists, while hypothesis $H_k$ means the presence of a signal sent from the $k$-th user. Evidently, the AP can first detect the hypothesis $H_0$, and only clarifies the origin of the signal if $H_k$ is decided as a false hypothesis. The clarification of $H_k$ leads to a classic energy detection with the test statistic $T(y) = |y|^2$, which is compared against a threshold $\beta$ for declaring the presence of an incoming signal. Define the probability of false alarm as the probability of the AP falsely clarifying the presence of an incoming signal. The probability of false alarm is calculated as $P_{FA}(\beta|H_0) = 1 - F_{\chi^2(2N_r)}\left(\frac{2\beta N_r}{\sigma^2}\right)$, where $F_{\chi^2(N)}(\cdot)$ denotes the cumulative distribution function (cdf) of a chi-square random variable with $2N_r$ DoFs. There is a series of other energy detectors, and we refer readers to [17] for the discussion herein. Subsequently, the AP can turn to clarify the origin of the signal. Since the propagation channel is the unique PHY identity of a specific user within certain time, the detection of the sender’s identity is equivalent to the identification of the propagation channel. With the knowledge of CSI set $H_k, \forall k \in \mathbb{K}$, the AP is able to apply the MLE to estimate the transmitted vector $x_k = W_k s_k \in \mathbb{C}^{N_t \times 1}$, in the form of $x_k = H_k^\dagger y = W_k s_k + H_k^\dagger z$, where $H_k^\dagger = (H_k^H H_k)^{-1} H_k^H$ denotes the pseudo-inverse of the channel $H_k$. Then, the estimated vector $\hat{x}_k$ is left multiplied by $H_k$ to imitate that it propagates through MIMO channel $H_k$, leading to a re-constructed signal $\hat{y}_k$ as $\hat{y}_k = H_k \hat{x}_k = H_k W_k s_k + H_k H_k^\dagger z$. Evidently, if the received signal comes from the $k$-th user, there is a high probability that the re-constructed signal $\hat{y}_k$ built on $H_k$ has the smallest Euclidean distance to the actual signal $y$. Hence, the AP can calculate the Euclidean distance between the actual signal and different re-constructed signal in sequence, and clarify the one with the minimum value as the real sender. Finally, the MLE-based signal trace-back detector is finally written as

$$D_{MLE} = \min_{k \in \mathbb{K}} \{|| (I_{N_r} - H_k H_k^\dagger) y ||^2, ..., || (I_{N_r} - H_K H_K^\dagger) y ||^2 \},$$

Note that $H_k H_k^\dagger \neq I_{N_r}$, when $N_r > N_t$. In our previous work [9], the so-called MHT detector was proposed, which
demonstrates a strong DER error floor (we refer readers to [9] for the discussion herein). In contrast, the MLE detector is able to clarify the real sender with a low level of DER, which will be shown in our simulation part. As a result, a great challenge is imposed for the sender’s anonymous precoder design against the AP’s detection.

4. ANONYMITY ENTROPY BASED PRECODER

Problem formulation of the anonymous precoder design and its optimal solution are presented in subsections IV-A and B.

4.1. Problem Formulation

By the MLE detector, since the AP calculates the norm in (3) in sequence and considers the one with the minimum value as the sender, we can select another user \(k', k' \neq k\), from the set \(K\) as an alias, and confines the following inequality as

\[
|| (H_{k'}H_{k'}^\dagger - H_kH_k^\dagger)H_kW_k s_k ||^2 \leq \delta, \forall k', k' \in K, (4)
\]

With a small-valued threshold \(\delta\), (4) lets the \(k\)-th and \(k'\)-th users equally suspicious to the AP, and hence the AP fails to declare who is the real sender as well as the exact channel that the signal comes from, further indicating that a correct equalizer would be impossible. Hence, we need to treat each receive antenna of the AP as an individual receiver and guarantee per-antenna SINR requirement for multiplexing data streams, so that the AP can decode the signal directly without the help of the channel equalizer. Note that the principle is to manipulate the dissipated signal pattern to inhibit the AP’s detection, instead of letting the alias \(k'\) transmit artificial noise. Now, we target to maximize the per-antenna SINR under the sender’s anonymity entropy requirement, written as

\[
P_1 : \max_{W_k} \Gamma, \tag{5}
\]

s.t. (C1) : \(||W_k s_k||^2 \leq p_{\text{max}}, \) (C2) : \(\forall r \in N_r, \Gamma_r \geq \Gamma, \) (C3) : \(|| (H_{k'}H_{k'}^\dagger - H_kH_k^\dagger)H_kW_k s_k ||^2 \leq \delta, \forall k', \) \(\forall k' \in K, (5)\)

where (C1) denotes the power budget constraint. (C2) guarantees the per-antenna SINR \(\Gamma_r\) higher than the lower-bound \(\Gamma\). (C3) suppresses the test norm to inhabit the AP’s detection.

4.2. Optimization Solution

The difficulties of solving P1 lie in multiplexing \(N_r\) data streams (where \(N_r > N_f\)), and meanwhile guaranteeing the sender’s anonymity requirement. In [9], we have proposed an interference-suppression based anonymous (ISA) precoder, whereas it is only applicable for the strong sender case with \(N_r \leq N_f\). More importantly, by the ISA precoder, the transmitted data is considered as Gaussian signal, and thus the inter-antenna interference is always treated as a harmful element to be strictly suppressed, which significantly reduces the DoFs of the precoder design.

In practice, the constellation size is limited and the input may not always be Gaussian signal. Since the transmitted symbols are known by the sender, based on the concept of interference exploitation (IE) [18], the inter-antenna interference can be utilized as a beneficial element to enhance the per-antenna receive performance. Let us start by briefly discussing the concept of IE and then we elaborate IE for addressing anonymity. We use PSK modulation for notation simplicity, nevertheless the following is applicable to multi-level modulations [19]. Write the intended symbol of the \(r\)-th receive antenna as \(s_k \in \mathbb{C}\) by M-PSK modulation, where the symbol amplitude is normalized to \(|d| = 1\). It can be expressed as a rotated version of another symbol, such that \(s_k = e^{j\phi} d\) to \(s_k \in \mathbb{C}\). Denotes \(w_r \in \mathbb{C}^{N_r \times 1}\) as the precoder vector for the intended symbol \(s_k\), where we have \(W_k = [w_1^T, ..., w_r^T, ..., w_{N_r}^T]^T\) and \(s_k = [(s_k)_1, ..., (s_k)_r, ..., (s_k)_{N_r}]\). Hence, the received signal of the \(r\)-th receive antenna is written as

\[
y_r = h_r \sum_{r'=1}^{N_r} w_r (s_k)_{r'} e^{j(\phi_r - \phi)} + z_r, \tag{6}
\]

where \(h_r \in \mathbb{C}^{1 \times N_r}\) denotes the channel from the sender to the \(r\)-th receive antenna, and \(z_r\) is the noise at the \(r\)-th receive antenna. Taking \(s_1\) as a reference symbol (which can be arbitrary), the received signal is re-written as \(y_r = h_r e^{j(\phi_1 - \phi)} \sum_{r'=1}^{N_r} (w_r e^{j(\phi_r - \phi_1)})(s_k)_{r'} + z_r\), which denotes that the original inter-antenna interference channel reduces to a virtual multicast channel with common messages \((s_k)_{r'}\) to all receive antennas [19], and hence the inter-antenna interference can be turned into a beneficial element. Write \(\bar{y}_r\) as the noise-excluded signal on the \(r\)-th receive antenna. To make the inter-antenna interference constructive, one can rotate \(\bar{y}_r\) with an angle \(\alpha^R_r = 3\bar{y}_r (s_k)_{r'}^*\) and real axis \(\alpha^R_r = \Re\{\bar{y}_r (s_k)_{r'}^*\}\), respectively. As shown in in Fig. 1(c), \(\bar{y}_r\) falls into a constructive region (in Fig. 1(b)), if and only if \(|\alpha^R_r| \leq (\alpha^R - \gamma)\tan(\pi M)\) holds, where \(M\) denotes constellation size. In particular, \(\gamma\) physically denotes

![Fig. 1. The conventional precoders locates the intended signal in the proximity region around the constellation point (in Fig. 1(a)). In contrast, the IE-based precoder locates the desired symbol \((\frac{1+j}{\sqrt{2}}\) with QPSK modulation for illustration) into a constructive region (green area in Fig. 1(b)), by exploiting the geometric interpretation in Fig. 1(c).](image-url)
the Euclidean distance in the signal constellation between the constructive region and the decision thresholds, which directly relates to SINR performance. Substituting $\alpha_r$ and $\alpha_r^\Re$ into the inequality above yields

$$|\Im\{h_r W_k s_k (s_k)^*\}| \leq (\Re\{h_r W_k s_k (s_k)^*\} - \gamma_r) \tan(\pi M),$$ (7)

Now, P1 can be readily transformed into an equivalent optimization problem

$$P2 : \max \gamma, \text{ s.t. } (C1) : ||W_k s_k||^2 \leq p_{max}, (C2) : (\forall \gamma) , (C3) : ||(H_k H_k^H - H_k H_k^H) H_k W_k s_k||^2 \leq \delta, (C4) : k \neq k', k' \in K.$$ (8)

Though P2 is standard convex problem, if one selects all the $k' \in \mathbb{K}$ as aliases (hence there are $K-1$ constraints in (C3)), all the users will be equally suspicious to the AP with the highest level of anonymity $A(\mathbb{K}) = \log_2 K$. Nevertheless, it significantly constrains the DoFs for precoder design, further degrading the per-antenna SINR performance. To make a good compromise between the communication SINR and sender’s anonymity, we can randomly select a user $k'$ from $\mathbb{K}$ as the alias. As a result, there will be only 1 constraint in (C3) without significantly confining DoFs. Also, the AP still fails to declare the correct sender, as the sender $k$ and the alias $k'$ are equally suspicious. The whole algorithm is finally summarized as follows.

**Algorithm 1 The AEA Precoder Design**

**Input:** CSI $H_k$, power budget $p_{max}$, intended symbol $s_k$.  
1: Select a random user $k'$ from $\mathbb{K}$ as the alias. Then solve P2.  
**Output:** Optimal precoding design $W^*_k$.

**Remark 1:** Based on the AEA precoder, the received signal of each antenna has been directly located into constructive regions of the constellation. Hence, the AP can demodulate the received signal based on the amplitude and phase of the received signal. As a result, it removes the need for receiver or transmitter equalization, while utilizing inter-antenna interference as a beneficial element without loss of anonymity.

5. SIMULATION RESULTS

Now, we present the simulation results. Each block includes 50 symbols, and Rayleigh block-fading is adopted for channel modeling. The following classic precoders are selected as comparison algorithms: 1) IE precoder [20], where the sender’s anonymity is not considered. 2) MMSE precoder [21]. 3) SVD precoder [22]. In particular, the AP first tries to reveal the real sender, and then the precoder and equalizer are calculated based on the declared hypothesis. Note that the

![Fig. 2. DER and SER performance vs. different receive SNRs, where $\delta = 0.03$, $N_r = 10$, $N_t = 9$, $K=5$, $p_{max} = 1$ Watt, and $\beta = 10^{-2}$. QPSK modulation is adopted.](image)

ISA precoder in [9] is not plotted as it is inapplicable for the strong receiver configuration.

As can be observed in Fig. 2, since the AEA precoder controls the pattern of the waveform to scramble the AP’s signal trace-back detector, the AP’s DER is pushed to a high level, thereby ensuring the sender’s anonymity. In comparison, with the conventional SVD, MMSE and IE precoders, the AP is able to reveal the correct sender with a low DER performance, i.e., $\text{DER}=10^{-3}$ at 15 dB SNR regime. On the other hand, since the IE precoder utilizes the inter-antenna interference without anonymous constraints, its high DoFs endorse the lowest SER performance [20]. However, the proposed AEA precoder achieves a close SER to the IE precoder, and outperforms the SVD and MMSE at moderate/high SNR regimes. Hence, the AEA precoder makes a good compromise between guaranteeing reasonable receive quality for communication and providing senders anonymity.

6. CONCLUSIONS

Focusing on the strong receiver configuration at uplink, we first have proposed a novel MLE signal trace-back detector for the AP, which only analyzes the signaling patterns of the received signal to disclose the sender’s identity and demonstrates a significantly enhanced detection ability. Accordingly, we have further proposed an AEA precoder to provide anonymity for the sender, and simultaneously maintain high per-antenna SINR performance for communication. In particular, the AEA precoder is able to multiplex more data streams than the number of the transmit antennas. Compared to the classic IE, MMSE, SVD precoders, simulation results have demonstrated that the proposed AEA precoder can scramble the AP’s anonymity-violating detector, where the DER of is pushed to a high level. Also, high per-antenna receive SINR is guaranteed, striking a good compromise between the anonymity and communication performance.
7. REFERENCES


