Imbalanced Beamforming by a Multi-antenna Source for Secure Utilization of an Untrusted Relay

Amitav Mukherjee, Member, IEEE

Abstract—We investigate a relay network where a multi-antenna source can potentially utilize an unauthenticated (untrusted) relay to augment its direct transmission of a confidential message to the destination. Since the relay is untrusted, it is desirable to protect the confidential data from it while simultaneously making use of it to increase the reliability of the transmission. We present a low-complexity scheme denoted as imbalanced beamforming based on linear beamforming and constellation mapping that ensures perfect physical-layer security even while utilizing the untrusted relay. Furthermore, the security of the scheme holds even if the relay adopts the conventional decode-and-forward protocol, unlike prior work. Simulation results show that the proposed imbalanced signaling maintains a constant BER of 0.5 at the eavesdropper at any SNR and number of source antennas, while maintaining or improving the detection performance of the destination compared to not utilizing the relay or existing security methods.

Index Terms—Physical-layer security, untrusted relay.

I. INTRODUCTION

There has been a resurgence of interest in covert wireless communications that are secure from eavesdropping at the physical layer without relying on higher-layer encryption. Information security is an important issue for both single-hop links [1]-[3] and relay networks [4], in which secure transmissions may be compromised by external eavesdroppers that are distinct from the source and the relay nodes. Relays can deploy well-known protocols such as amplify-and-forward (AF), decode-and-forward (DF), and compress-and-forward (CF) to aid the secure transmission of messages in the presence of external eavesdroppers.

However, even if external eavesdroppers are absent, it may still be desirable to keep the source signal confidential from the relay node itself in spite of its assistance in forwarding the data to the destination [5]. For example, the unauthenticated relay may belong to a heterogeneous network without the same security clearance as the source and destination. This scenario has also been denoted as cooperative communication via an untrusted relay in [6], [7], where the authors presented bounds on the achievable secrecy rate.

Prior art on untrusted relays has focused on computing either the information-theoretic system secrecy rate [5]-[8] or the probability that it is in outage [9]. However, information-theoretic measures such as secrecy rate generally rely on idealized assumptions such as continuous (e.g., Gaussian) input distributions and the existence of ‘good’ random coding schemes with asymptotically long block-lengths. Therefore, in order to account for practical codes with finite block lengths and discrete modulation alphabets, the authors of [10] advocate the use of bit-error rate (BER) as a security metric, and define physical-layer security to be the enforcement of 0.5 BER at the eavesdropper while the destination BER is below a reliability threshold. In [10], this was achieved by appropriate puncturing of the output of a low-density parity check encoder at the source for a single-hop network with an external eavesdropper. In contrast, we present a low-complexity scheme for uncoded MISO systems with an untrusted relay based on a combination of transmit beamforming and constellation mapping for M-ary phase-shift keying (PSK) or quadrature amplitude modulation (QAM), that ensures the relay has a constant BER of 0.5 even when it is utilized.

The contributions of this letter can therefore be summarized as follows: (1) we propose a novel and low-complexity imbalanced beamforming (IBF) scheme where separate beamforming vectors are applied to the real and imaginary components of the transmitted symbol; (2) we define a class of bit-labeling schemes for various PSK and QAM constellations, that combined with IBF provide perfect physical-layer security while using an untrusted relay, (3) perfect security is obtained regardless of the relaying protocol, while previously, untrusted relays could not have employed DF, and (4) the proposed scheme outperforms existing schemes based on generalized eigenvector beamforming, artificial noise transmission by the source, and cooperative jamming by the destination.

II. MATHEMATICAL MODEL

A. Network Model

The network under consideration comprises a transmitter with N antennas, while the intended destination and an unauthenticated relay have a single antenna each. A direct link is assumed to exist between source and destination, and all nodes operate in half-duplex mode. The transmitter wishes to convey a confidential bit sequence $m = \{m_0, \ldots, m_{K-1}\} \in [0, 1]^K$ to the destination, where each bit is assumed to be equally likely. The transmitter can choose to utilize the relay to augment its transmission to the destination, but must ensure that the relay cannot decode the message $m$ itself. Alternatively, the source can treat the relay purely as an eavesdropper and only rely on the direct link to the destination.

The transmitter adopts a memoryless two-dimensional $M$-ary modulation scheme such as PSK or QAM, by mapping $q = \log_2 M$ bits at a time from $m$ to a unit-power complex symbol $s = s_R + js_I$. Let the alphabets of the real and imaginary components of $s$ be denoted as $S_R = \{s_{R,1}, \ldots, s_{R,|S_R|}\}$ and $S_I = \{s_{I,1}, \ldots, s_{I,|S_I|}\}$, i.e., $s_R \in S_R$ and $s_I \in S_I$. The $i^{th}$
2D constellation point is mapped to a unique length-\(q\) binary vector or bit pattern \(b_i = [b_{i,1}, \ldots, b_{i,q}], i = 1, \ldots, M\).

In order to isolate the security gain with the proposed scheme we consider an uncoded system, although a channel encoder can be incorporated without modification. In the scenario where the source utilizes the relay, the overall transmission time per symbol is split into two slots. In the first time slot the source broadcasts signal \(x_t \in \mathbb{C}^{N \times 1}\) to the relay and destination; in the second time slot it remains silent while the relay transmits \(x_r\) to the destination. Signals \(x_t\) and \(x_r\) are functions of \(s\) and \(y_r\), respectively. The source and relay each have transmit power constraints \(E\left\{ \|x_t\|^2 \right\} \leq P_t\) and \(E\left\{ \|x_r\|^2 \right\} \leq P_r\), respectively. The received signals at the relay and destination over two time slots are then given by

\[
y_r = h_rx_t + n_r \tag{1}
y_{d,1} = h_dx_t + n_{d,1} \tag{2}
y_{d,2} = g_dx_r + n_{d,2} \tag{3}
\]

where \(h_r \in \mathbb{C}^{1 \times N}\), \(h_d \in \mathbb{C}^{1 \times N}\), and \(g_d\) are the source-to-relay, source-to-destination, and relay-to-destination channels, and \(n_r, n_{d,1}, n_{d,2}\) are independent complex additive white Gaussian noise samples with variance \(N_0/2\) per real and imaginary dimensions. All channels and noise samples are mutually independent. We assume the transmitter and relay both have knowledge of all channel realizations, while the destination has perfect receive channel state information (CSI) in both slots. This is feasible since the relay is a cooperating node and not a conventional passive eavesdropper, and has also been assumed in [6]-[9]. Furthermore, both relay and destination know the constellation mapping used by the transmitter. Therefore, the security of the IBF scheme is not dependent upon the use of CSI or constellation mapping as a shared secret key.

B. Existing Approaches

In the conventional case where the source does not utilize the relay and treats it purely as an eavesdropper, the optimal transmit signal that maximizes the secrecy rate\(^1\) is given by [2, Theorem 2]

\[
x_t = \psi_{max}s \tag{4}
\]

where \(\psi_{max}\) is the generalized eigenvector corresponding to the largest generalized eigenvalue of the matrix pencil \((I + P_r h_r^H h_r, I + P_h h_d^H h_d)\). In this scenario the relay stays silent and the source transmits a new symbol in every time slot. This scheme is referred to as generalized eigenvector beamforming (GBEF) in Sec. [IX].

Popular alternatives for secure transmission are schemes based on artificial noise (AN) or destination jamming (DJ), where the legitimate nodes deliberately transmit noise-like signals to jam eavesdroppers [1, 6, 11]. In the AN scheme, the transmitter splits its total power between the data symbol and a zero-mean unit-variance Gaussian AN symbol \(w\) in time slot 1. Define the maximum ratio transmission (MRT) beamforming vector as \(\tilde{h}_d^H = h_d^H/\|h_d\|\). The data is sent via MRT, while the AN is transmitted in the orthogonal space of the intended receiver’s channel:

\[
x_t = \sqrt{\beta} \tilde{h}_d^H s + \sqrt{P_t - \beta} (h_d^H)^H w \tag{5}
\]

where scalar \(0 \leq \beta \leq P_t\) controls the power allocation between data and AN. However, the destination cannot cancel out the AN retransmitted to it in the second phase by the relay for the AF scenario.

Under DJ, the transmitter sends only data via MRT while other nodes jam eavesdroppers. In the absence of external helpers, the destination node jams the untrusted relay in time slot 1 and therefore does not receive information in this phase, i.e., \(y_{d,1} = 0\). The relay receives data along with the jamming signal as

\[
y_r = h_rx_t + f_rw + n_r
\]

where \(f_r\) is the destination-to-relay channel coefficient. When the AF protocol is in effect, the destination can cancel out its own jamming signals from \(y_{d,2}\).

III. IMBALANCED BEAMFORMING

A. Beamforming Design

The two major ingredients of the proposed imbalanced signaling scheme are the beamforming design and constellation mapping at the source. The main idea for the beamforming design is to have the source construct \(x_t\) such that only the real (or imaginary) part of the confidential symbol \(s\) is received by the relay, whereas the destination receives \(s\) in its entirety in the first time slot. This is achieved as follows. Let \(Z\) represent the orthonormal basis of the nullspace of \(h_r\), which implies \(h_r Z = 0\). Given \(h_r\) is rank-1, \(Z\) is comprised of \(N - 1\) orthogonal vectors. Denoting an arbitrary column of \(Z\) by \(h_r^\perp \in \mathbb{C}^{N \times 1}\), the transmit signal is designed as

\[
x_t = \left[ \sqrt{\alpha} \tilde{h}_d^H \sqrt{P_t - \alpha} R_j + j \sqrt{P_t - \alpha} R_j \right] \begin{bmatrix} s_R & j s_I \end{bmatrix}^T
\]

\[
= \sqrt{\alpha} \tilde{h}_d^H s_R + j \sqrt{\alpha} \tilde{h}_d^H s_I \tag{6}
\]

where the scalar \(0 < \alpha \leq P_t\) is used to preserve the power constraint \(P_t\). Therefore, separate beamformers (MRT and nullspace) are applied to the real and imaginary components of the confidential symbol to maximize and zero-force their reception at the intended destination and relay, respectively.

B. Optimal Relay Eavesdropping

Based on the preceding discussion on the transmit beamforming design, the relay observes only the real part of the confidential symbol:

\[
y_r = \sqrt{\alpha} h_d \tilde{h}_d^H s_R + n_r \tag{7}
\]

Let \(\tilde{y}_r = \text{Re}\{y_r\}\) and \(y_r = \sqrt{\alpha} \text{Re}\left\{h_r \tilde{h}_d^H\right\}\). From the perspective of physical-layer security, we make the worst-case assumption that the untrusted relay utilizes an optimal bitwise demodulator that minimizes its (uncoded) BER [12]:

\[
L_j(\tilde{y}_r) = \log \frac{\Pr (M_j = 1 | Y_r = \tilde{y}_r)}{\Pr (M_j = 0 | Y_r = \tilde{y}_r)} = \log \frac{\sum_{s \in S_{y_r}} e^{-(y_r - g, s)^2/N_0}}{\sum_{s \in S_{\tilde{y}_r}} e^{-(y_r - g, s)^2/N_0}}, j = 1, \ldots, q, \tag{8}
\]

\[1\]This scheme assumes a Gaussian input distribution; a closed-form solution for the optimal beamformer with finite inputs remains an open problem.
where $S_{R,t} = \{s_{R,i} \in S_R : b_{i,t} = k, \forall i\}$, $q = \log_2 M$ as before, and we have exploited that the likelihood function of the received signal in (7) is Gaussian when conditioned on the message and channel. The likelihood ratio test at the relay is effectively the optimal bitwise 1D pulse-amplitude modulation detector [12] Sec. II-B] with alphabet $S_R$, since only the real dimension is observed. Note that (8) would need to be modified when considering a coded system. By correctly detecting $s_R$ the relay narrows the number of possible values of the symbol $s \in M/2$, but with appropriate constellation mapping the a posteriori uncertainty per individual bits can remain unchanged, as explained next.

C. Constellation Mapping

The second facet of the IBF scheme is the assignment of bit patterns to constellation points. There are $M^q$ possible mappings of bit patterns to constellation points at the transmitter, with Gray coding being a classical example. In the sequel we assume $S_R$ and $S_T$ are each 1D PAM alphabets. Let the $(d_c \times q)$ binary matrix $B_c$ represent all the bit patterns associated with the $c^{th}$ element of $S_R$, $c = 1, \ldots, |S_R|$, where the $k^{th}$ row of $B_c$ is the binary vector mapped to the complex symbol $s_{R,c} = j k_{R,c}$. For QPSK/8-PSK we have $d_c = 2$, and $d_c = |S_T|$ for square QAM constellations, $\forall c$. For rectangular 32-QAM and 128-QAM constellations, $d_c = 4$ and $d_c = 8$ for extremal elements of $S_R$; $d_c = |S_T|$ for all other elements.

Definition 1: An E-map is defined as a constellation bit mapping wherein for every real-dimensional coordinate $s_{R,c} \in S_R$, the mapping matrix $B_c$ has an equal number of 1s and 0s in each of its $q$ columns.

The above definition implies that an E-map can be constructed only when $d_c$ is even. Fortunately, E-maps can be easily constructed for commonly-used 2D M-ary constellations, and several existing constellation labels satisfy this property:

- QPSK: Anti-Gray mapping.
- 8-PSK: The mapping proposed in Fig. [1].
- 16 QAM: Anti-Gray mapping.
- 32-QAM/128-QAM: Approximate Anti-Gray mapping.
- 64 QAM: The “DS” mapping in [13].
- 256-QAM: The TV-map in [14] pg. 5046.

We reiterate that the relay has complete knowledge of all CSI and the E-map used by the source.

Proposition 1: The imbalanced beamforming scheme with E-maps provides perfect physical-layer security at any SNR and number of transmit antennas $N$, regardless of the relaying protocol.

Proof: The detection of $s_R$ by the relay in (7) is equivalent to $|S_R|$-PAM demodulation. Conditioned on correct PAM detection, for every transmitted symbol the relay must attempt to determine the associated bit sequence by choosing the correct row from the associated mapping matrix $B_c$. Without knowledge of imaginary component $s_I$, the best the relay can do is to randomly pick a row from $B_c$. By definition, under E-maps each column of $B_c$ has an equal number of 1s and 0s, and therefore the relay will correctly pick each individual transmit bit with probability 1/2. Hence, the relay BER is constantly 0.5 at any SNR, whereas the BER at the destination is bounded away from 0.5 as $P_t, P_r, N$ increase since it receives both real and imaginary components of $s$.

To complete the description of the relay system, we outline the destination processing for AF and DF relay protocols. When the AF protocol is deployed by the relay, its transmit signal is given by $x_r = \sqrt{\gamma} \Re\{y_d\}$, where $\gamma$ enforces the relay power constraint $P_r$. The destination aggregates the signals received over two time slots into $y_d = [y_{d,1}, y_{d,2}]^T$, equivalent to

$$y_d = \begin{bmatrix} \sqrt{\gamma} h_d \tilde{h}_d^H & \sqrt{P_r - \alpha h_d h_d^H} & 0 \end{bmatrix} \begin{bmatrix} s_r \ j s_I \end{bmatrix} + \begin{bmatrix} n_{d,1} \ n_{d,2} \ \tilde{n}_{d,2} \end{bmatrix}$$

where the effective noise term is $\tilde{n}_{d,2} = \sqrt{\gamma} g_d \Re\{c_r\} + n_{d,2}$. The destination retrieves the MMSE estimate of $[s_{R}, j s_I]^T$ as

$$\hat{s}_t = \begin{bmatrix} H_e H_e^H + \begin{bmatrix} 2N_0 & 0 \\ 0 & 2P_r |g_d|^2 + 2N_0 \end{bmatrix} \end{bmatrix}^{-1} H_e^H y_d$$

where $H_e$ is the $(2 \times 2)$ equivalent channel in (10), and then performs minimum-distance detection for the associated E-map. Under the DF protocol, we have $x_r = \sqrt{\gamma} s_R$, where $s_R$ is the relay’s estimate of the real component of $s$ with corresponding BER as in [15], and the extension of the destination detection under AF is straightforward. When the relay is not utilized as in GEBF, $y_{d,2} = 0$ and the destination performs minimum-distance detection based on $y_{d,1}$.

We conclude this section by briefly describing how imbalanced signaling can be applied to the scenarios where all terminals have a single antenna, or where the relay has more antennas than the source. Since no spatial degrees-of-freedom are now available to conceal the imaginary component $s_I$ from the relay, we must resort to non-linear transmission schemes such as dirty-paper coding [16] to achieve this. In all cases, the E-map design would remain unchanged, and perfect physical-layer security can still be ensured.

IV. NUMERICAL RESULTS

We present some examples that validate the security of the proposed IBF scheme in comparison with the existing artificial noise (AN) and destination jamming (DJ) schemes.

![Fig. 1. E-map for 8-PSK. $B_e$ is of dimension $(2 \times 3)$ $\forall c$.](image-url)
from Sec. [11B] using E-maps in all cases. The channels and background noise samples are assumed to be composed of independent, zero-mean Gaussian random variables with unit variance. In AN schemes, it has been observed previously that allocating power roughly equally between AN and data is a good rule of thumb for ergodic secrecy rate maximization, and performance is not very sensitive to small variations in $\beta$ [3]. We roughly follow this rule of thumb with fractionally a little more power ($\beta = 0.7P_t$) allocated for the data. For the IBF scheme, the source allocates equal power to the real and imaginary components by setting $\alpha = 0.5P_t$. For DJ, the extra destination jamming power is set to $0.5P_t$ in addition to transmitter power $P_t$, i.e., there is no power sharing between transmitter and destination.

V. CONCLUSIONS

In this letter we presented a low-complexity scheme denoted as imbalanced beamforming based on linear beamforming and constellation mapping that ensures perfect physical-layer security even while utilizing an untrusted relay, when a direct link to the destination is also available. Simulations showed that the proposed imbalanced signaling maintains a constant BER of 0.5 at the eavesdropper at any SNR and number of source antennas, while improving the detection performance of the destination compared to not utilizing the relay.

REFERENCES

[1] S. Goel and R. Negi, “Guaranteeing secrecy using artificial noise,” IEEE Trans. Wireless Commun., vol. 7, no. 6, pp. 2180-2189, June 2008.
[2] A. Khisti and G. Wornell, “Secure transmission with multiple antennas I: The MISOME wiretap channel,” IEEE Trans. Inf. Theory, vol. 56, no. 7, pp. 3088-3104, July 2010.
[3] A. Mukherjee and A. L. Swindlehurst, “Robust beamforming for secrecy in MIMO wiretap channels with imperfect CSI,” IEEE Trans. Signal Process., vol. 59, no. 1, pp. 351-361, Jan. 2011.
[4] L. Lai and H. Gamal, “The relay-eavesdropper channel: Cooperation for secrecy,” IEEE Trans. Inf. Theory, vol. 54, pp. 4005-4019, Sep. 2008.
[5] Y. Ohama, “Coding for relay channels with confidential messages,” in Proc. IEEE Inf. Theory Workshop, Sep. 2001, pp. 87-89.
[6] X. He and A. Yener, “Cooperation with an untrusted relay: A secrecy perspective,” IEEE Trans. Inf. Theory, vol. 56, pp. 3807-3827, Jul. 2010.
[7] X. He and A. Yener, “Two-hop secure communication using an untrusted relay, EURASIP J. Wireless Commun. Netw., vol. 2009, Article ID 305146.
[8] C. Jeong, I.-M. Kim, and D. I. Kim, “Secure beamforming design at the source and the relay for an amplify-and-forward MIMO untrusted relay system,” IEEE Trans. Signal Process., vol. 60, pp. 310-325, Jan. 2012.
[9] J. Huang, A. Mukherjee, and A. L. Swindlehurst, “Secure communication via an untrusted non-regenerative relay in fading channels,” IEEE Trans. Signal Process., vol. 61, no. 5, pp. 2536-2550, May 2013.
[10] D. Kline, J. Ha, S. W. McLaughlin, J. Barros, and B.-J. Kwak, “LDPC codes for the Gaussian wiretap channel,” IEEE Trans. Inf. Forensics Security, vol. 6, no. 3, pp. 532-540, Sep. 2011.
[11] Z. Ding, K. Leung, D. L. Goeckel and D. Towsley, “Opportunistic relaying for secrecy communications: Cooperative jamming vs relay chatting,” IEEE Trans. Wireless Commun., vol. 10, no. 6, pp.1725-1729, June 2011.
[12] M. K. Simon and R. Annanavajala, “On the optimality of bit detection of certain digital modulations,” IEEE Trans. Commun., vol. 53, no. 2, pp. 299-307, Feb. 2005.
[13] S. Pfletsching and F. Sanzi, “Error floor removal for bit-interleaved coded modulation with iterative detection,” IEEE Trans. Wireless Commun., vol. 5, no. 9, pp. 3174-3181, Sep. 2006.
[14] D. Torrieri and M. Valenti, “Constellation labeling maps for low error floors,” IEEE Trans. Wireless Commun., vol. 7, pp. 5401-5407, Dec. 2008.
[15] I. Lee and D. Kim, “BER analysis for decode-and-forward relaying in dissimilar Rayleigh fading channels,” IEEE Commun. Lett., vol. 11, no. 1, pp.52-54, Jan. 2007.
[16] S. A. A. Fakoorian and A. L. Swindlehurst, “Dirty paper coding versus linear GSV-based precoding in MIMO broadcast channel with confidential messages,” in Proc. IEEE GLOBECOM, Dec. 2011.