A Covert Communication System Using Non-zero Mean Normal Distributions

Zhijiang XU [1,3], Weidang LU [2,3], Yi GONG [3], Jingyu HUA [4], Wenbing JIN [1]

[1] School of Automation, Zhejiang Institute of Mechanical & Electrical Engineering, Binwen Rd. 528, 310052 Hangzhou, China
[2] College of Information Engineering, Zhejiang University of Technology, Liuhe Rd. 288, 310023 Hangzhou, China
[3] Univ. Key Laboratory of Advanced Wireless Communications of Guangdong Province, Dept. of Electrical and Electronic Engineering, Southern Univ. of Science and Technology, Xueyuan Blvd. 1088, 518055 Shenzhen, China
[4] School of Information & Electronic Engineering, Zhejiang Gongshang University, Xuezheng Street 18, 310018 Hangzhou, China

{xuzhijiang, jinwenbing}@zime.edu.cn, {eehjy, luweid}@zjut.edu.cn, gongy@sustech.edu.cn

Submitted March 17, 2020 / Accepted June 15, 2020

Abstract. *A covert communication system is proposed in this study, in which a non-zero mean Gaussian sequence is used as a random carrier and its mean is modulated by a covert binary bit. The aperiodic transmitted signal exhibits the same statistical characteristics as the ambient noise to confuse an eavesdropper. The received signal is multiplied with the pseudo-random sequence synchronized with the transmitter to recover these positive and negative mean Gaussian sequence. The sample mean estimator and hard decision are used to determine the covert message, and accordingly, theoretical bit error rate in additive white Gaussian noise channel is also derived. Simulation results are very consistent with the theoretical derivation. The proposed system works in the physical layer with the advantages of simple structure, strong concealment, good BER performance and very suitable for low-cost, resource-limited and low-rate transmission devices.*

Keywords

Normal distribution, Kullback-Leibler divergence, covert communication

1. Introduction

Recently, the five generation communication and beyond have attracted much attentions [1], in which many key technologies have been proposed, such as the HetNet [2], massive MIMO [3], cognitive network [4] and vehicular communications [5]. Moreover, the future wireless communication introduces new demands, i.e., higher requirements of ultra-low latency, ultrahigh efficiency, ultra-high reliability and ultra-high density connection, etc. Driven by potentially great market and huge profit, wireless devices are deployed quickly and are affecting every aspect of the world. In the meantime, the massive use of wireless devices will expose new vulnerabilities in system security and bring more challenging privacy and reliability issues. In our study, we focus on the issue of information security.

It is generally believed that the Secure Socket Layer (SSL) protocol developed by Netscape uses data encryption technology to ensure that data will not be intercepted and eavesdropped during the transmission process on the network. However, there are two main drawbacks with using SSL to ensure data security. On the one hand, encryption only prevents unauthorized parties from decoding the communication. Further, the protocol-based security mechanisms may be insufficient to secure communication with advances in emerging hardware and software technologies. Unfortunately, most standard security solutions designed for enterprise systems may be not suitable for many Internet of Things (IoT) devices because of the limited resources of these low-cost IoT terminals. On the other hand, in many cases the simple existence of communication or a change in communication mode, such as an increase in the frequency of messages, is sufficient to cause suspicion and reveal the onset of events [6].

The covert communication technology with signal camouflage has become a research hotspot. The method of covert communication is to hide the very existence of the communication, i.e., an eavesdroppers can detect or intercept the presence of communications with very low probability. Salberg in [7], [8] proposed an artificially generated noise-like stochastic process shift keying (SPSK) to achieve digital security communication system. Specifically, during each bit interval, a realization of the stochastic process $X_0(t)$ to represent logic '0' whereas a realization of another stochastic process $X_1(t)$ to represent logic '1'. In order to achieve the purpose of confusing the eavesdropper, the two stochastic processes in SPSK method try to have the same statistical characteristics as the ambient noise.

Due to the complicated electromagnetic propagation environment generated by natural and human-made sources, many wireless communications systems, such as Wi-Fi, cognitive radio, cellular and LTE, are rapidly becoming interference limited [9]. In a multi-user network with the power-law path loss, studies including theoretically derived in [10–12] and actually measured in [13], [14] showed that the aggregate multiple access interference might result in a α-stable distribution. The characteristic exponent α of a stable distributions sequence modulated by a covert binary bit to achieve a secure communication is proposed in [15]. Further, the received signal is demodulated by the logarithmic moments based characteristic exponent estimator to recover the binary message, and the optimal decision threshold and the minimum BER are also derived in [16]. However, the waveform of the α-stable distribution sequence of different characteristic exponent has significant difference in time domain, which is easy to be detected by eavesdroppers. In [17], a covert method utilizing the random signals with skewed α-stable distributions exhibiting antipodal characteristics to encode the binary information is proposed. However, this method requires strict synchronization, which is difficult to implement in practice.

There is no doubt that additive Gaussian noise is the most commonly used noise model in communication systems. Xu et al. in [18] proposed a covert communication scheme, in which the correlation coefficient of two consecutive Gaussian sequences is modulated by a covert bit. However, a large number of samples is needed to obtain a good correlation coefficient estimation. In this study, we propose a new covert communication scheme, in which the transmitted waveform modulated by the covert bit is statistically similar to the Gaussian channel noise to achieve security. The proposed scheme on the physical layer protects the header within each packet as well the data, whereas the encryption algorithm on the application layer only protects the data. More importantly, the proposed scheme has a simple structure and is easy to implement on IoT devices with limited resources.

The remaining parts of this study are organized as follows. In Sec. 2, a covert communication scheme using the mean of a Gaussian noise sequence modulated by a covert binary bit is proposed. In Sec. 3, both the probability density function (PDF) of the mean estimator and the corresponding theoretical bit error rate (BER) in additive white Gaussian noise (AWGN) channel are derived. To evaluate the performance of the proposed communication system, we present many Monte Carlo simulation results in Sec. 4 and draw our concluding remarks in Sec. 5.

2. Covert Communication Scheme

In many communication scenarios, ambient noise can reasonably be assumed to obey the normal distribution. If the transmitted signal is statistically consistent with the ambient noise, the eavesdropper can not distinguish the transmitted signal from the ambient noise to achieve the purpose of covert communication. In this section, we first put forward a covert communication system using normal distributions with non-zero mean, then demonstrate that the transmitted signal under certain conditions follows a normal distribution with zero mean, which is statistically consistent with the assumed ambient noise.

2.1 Proposed Covert Communication System

In the proposed covert communication system, as shown in Fig. 1, the binary message $b \in \{0,1\}$ is encoded by the mean parameter of the normally distributed noise generator. Specifically, in each bit period T_b, if the message bit is logical 'b', then an independent identically distribution (i.i.d) Gaussian sequence of length n, $\{x_i, i = 1,2,\cdots,n\}$, is generated by the noise generator. The sequence $\{X_i\}$ follows a normal distribution with mean $(-1)^b\mu$ and variance σ^2, denoted as $X \sim \mathcal{N}\left((-1)^b\mu, \sigma^2\right)$, where μ and σ are real numbers greater than 0. Similarly, the pseudo-random generator generates a bipolar equiprobability pseudo-random sequence of length n, $\{m_i, i = 1,2,\cdots,n\}$, its probability distribution is listed as $\Pr(m_i = -1) = \Pr(m_i = +1) = \frac{1}{2}$. The pseudo-random sequence can be either fixed or variable within each bit period. After multiplying the non-zero mean normal random sequence $\{x_i\}$ with the pseudo random sequence $\{m_i\}$, i.e., the sequence $\{s_i = m_i x_i, i = 1,2,\cdots,n\}$ is transmitted to the channel. In the following Sec. 2.2, it is demonstrated that the transmitted signal S follows a normal distribution of zero mean under certain conditions.

From the above description of the proposed covert transmitter, it can be known that the transmitted signal is observed to exhibit the same statistical characteristics as the ambient noise. In other words, the mean of the transmitted signal is 0 regardless of whether the covert bit is '0' or '1'. The transmitted noise sequence is shown in Fig. 2 together with the random covert bit stream with the sample length $n = 100$ in each bit period. The normally distributed noise has a mean of 0.8 and a standard deviation of 1, i.e., $\mu = 0.8$, $\sigma = 1$, and the Generator polynomial of the pseudo-random generator is: [7 6 0]. By comparing Fig. 2(b) from Fig. 2(c), we find that the modulated sequence in each bit period has a significant deviation from 0, while the transmitted sequence after being multiplied by the pseudo-random sequence has an obvious symmetry about 0.

In the receiver, the received sequence $\{r_i, i = 1,2,\cdots,n\}$ is multiplied by a bipolar equiprobability pseudo-random sequence $\{m_i, i = 1,2,\cdots,n\}$, which is perfectly synchronized with that of the transmitter. The sample mean estimator is used to estimate the mean of the samples $\{y_i = r_i m_i, i = 1,2,\cdots,n\}$ with length of n,

$$\hat{\mu} = h([y_1, y_2, \cdots, y_n]) \tag{1}$$

and the hard decision is used to determine the covert bit

$$\hat{b} = \begin{cases} 1, & \hat{\mu} \geq 0 \\ 0, & \hat{\mu} < 0 \end{cases} \tag{2}$$

In the following subsection, the probability density function of the transmitted signal S is derived. More importantly, on this basis, we obtain the condition for the signal S to obey the normal distribution with zero mean.

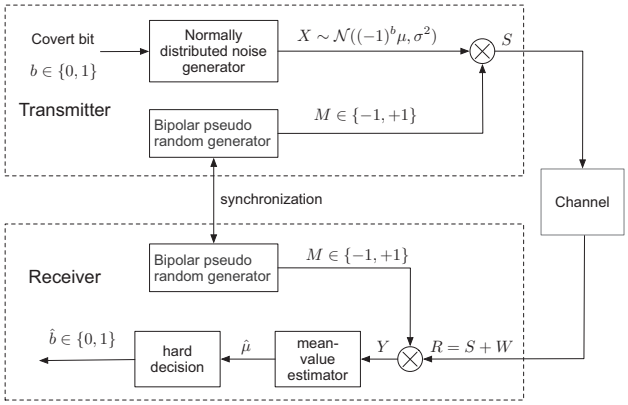
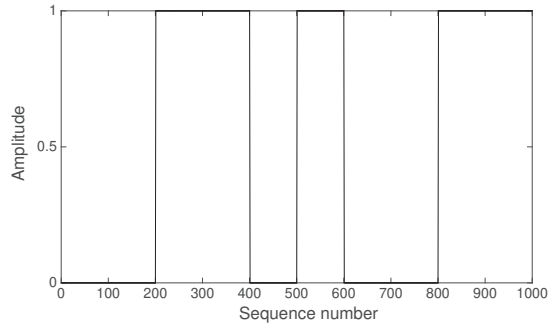
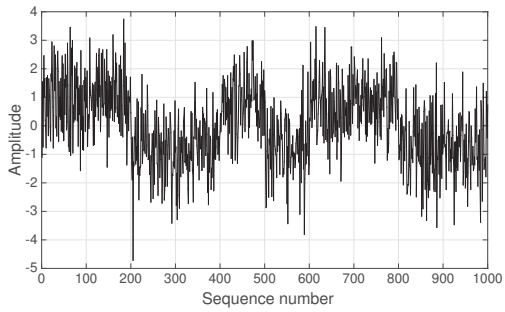


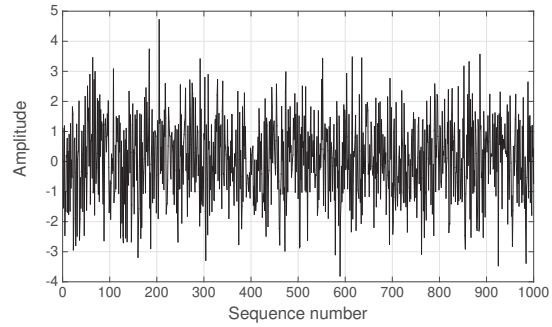
Fig. 1. Block diagram of the proposed covert communication system.



(a) Covert bit stream, $b \in \{0, 1\}$



(b) Modulated noise sequence, $X \sim \mathcal{N}((-1)^b\mu, \sigma^2)$



(c) Transmitted sequence S

Fig. 2. Covert bit stream, modulated signal and transmitted signal in time domain, where the length of each bit sequence $n = 100$, the mean $\mu = 0.8$ and the variance $\sigma^2 = 1$ of the normal distribution noise.

2.2 PDF of the Transmitted Signal

In this subsection, we first give the probability density function of the transmitted signal in the form of theorem, and then we demonstrate that the approximate PDF under certain condition is very close to the exact PDF from two aspects of Kullback-Leibler (KL) divergence and relative error. The approximate PDF shows that the transmitted signal S can be reasonably considered to be a normal random variable with zero mean.

Theorem 1 *Let $M \in \{+1, -1\}$ be a discrete random variable of equal probability, i.e., $Pr(M = +1) = Pr(M = -1) = \frac{1}{2}$. Let X be a continuous random variable and its probability density function be $f_X(x)$. Further, M and X are independent of each other. If let $S = MX$, then S is a continuous random variable and its corresponding PDF, $f_S(s)$, is given by*

$$f_S(s) = \frac{1}{2}\left(f_X(s) + f_X(-s)\right) . \tag{3}$$

Proof 1 *The distribution function of the random variable S is*

$$\begin{aligned} F_S(s) &= Pr(S \le s) = Pr(XM \le s) \\ &= Pr(M = 1, X \le s) + Pr(M = -1, X \ge -s) \\ &= Pr(M = 1)Pr(X \le s|M = 1) + \\ &\quad Pr(M = -1)Pr(X \ge -s|M = 1) \\ &= \frac{1}{2}\int_{-\infty}^{s} f_X(x)\mathrm{d}x + \frac{1}{2}\left(1 - \int_{-\infty}^{-s} f_X(x)\mathrm{d}x\right) . \end{aligned} \tag{4}$$

The derivative of the distribution function is its probability density function, i.e., the PDF of S is

$$f_S(s) = \frac{\mathrm{d}}{\mathrm{d}y}F_S(s) = \frac{1}{2}\left(f_X(s) + f_X(-s)\right) . \tag{5}$$

Moreover, $f_S(s)$ is symmetrically distributed about $S = 0$ since $f_S(-s) = f_S(s)$.

In particular, when X is a random variable of normal distribution with non-zero mean μ and standard deviation σ, denoted as $X \sim \mathcal{N}(\mu, \sigma^2)$, i.e., the PDF of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \tag{6}$$

the PDF of the corresponding transmitted signal S is

$$f_S(s) = \frac{1}{2}\left(\frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(s-\mu)^2}{2\sigma^2}} + \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-\frac{(s+\mu)^2}{2\sigma^2}}\right). \tag{7}$$

Let k be the mean-to-standard-deviation ratio (MSR), which means the ratio of the mean μ to the standard deviation σ of the normally distributed random variable X, i.e.,

$$k = \frac{\mu}{\sigma} . \tag{8}$$

From the following two aspects of Kullback-Leibler divergence and relative error, it is shown that when the MSR k

is less than 0.4, the random variable S can be considered to be normally distributed with zero mean. Specifically, when $0 < k \leq 0.4$, $f_S(s)$ can be approximated by the PDF $g_S(s)$ of a normal distribution, i.e., $g_S(s) \simeq f_S(s)$, and $g_S(s)$ is given by

$$g_S(s) = \frac{1}{\sqrt{2\pi}\sigma_S} \mathrm{e}^{-\frac{s^2}{2\sigma_S^2}} \tag{9}$$

where σ_S^2 is the variance of S and $\sigma_S^2 = \mu^2 + \sigma^2$.

2.2.1 Kullback-Leibler Divergence

For distributions P and Q of a continuous random variable, the Kullback-Leibler divergence [19] is used to measure in statistics that quantifies how close P is to an approximated distribution Q. The Kullback-Leibler divergence of Q from P is defined to be

$$D_{\mathrm{KL}}(P||Q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} \mathrm{d}x \tag{10}$$

where $p(x)$ and $q(x)$ denote the densities of P and Q. Here, we use the KL divergence to measure the exact PDF and approximate PDF of the random variable S, that is, how close (7) and (9) are. Their KL divergence is

$$\begin{aligned} D_{\mathrm{KL}}\left(g_S(s), f_S(s)\right) &= \int_{-\infty}^{\infty} g_S(s) \log \frac{g_S(s)}{f_S(s)} \mathrm{d}s \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1+k^2}\sigma} \mathrm{e}^{-\frac{s^2}{2(1+k^2)\sigma^2}} \times \\ &\quad \log \frac{\frac{2}{\sqrt{1+k^2}} \exp\left(-\frac{s^2}{2(1+k^2)\sigma^2}\right)}{\exp\left(-\frac{(s-k\sigma)^2}{2\sigma^2}\right) + \exp\left(-\frac{(s+k\sigma)^2}{2\sigma^2}\right)} \mathrm{d}s \\ &\overset{x=s/\sigma}{=\!=\!=} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1+k^2}} \mathrm{e}^{-\frac{x^2}{2(1+k^2)}} \times \\ &\quad \log \frac{\frac{2}{\sqrt{1+k^2}} \exp\left(-\frac{x^2}{2(1+k^2)}\right)}{\exp\left(-\frac{(x-k)^2}{2}\right) + \exp\left(-\frac{(x+k)^2}{2}\right)} \mathrm{d}x \,. \end{aligned} \tag{11}$$

The equation (11) shows that the KL divergence only depends on the MSR k, not on the variance σ^2. It means that the KL divergence is the same whether the variance of the noise X is large or small. Unfortunately, there is no closed-form analytical expression to the (11). Therefore, the numerical integration is used to calculate the value of the (11), as shown in Fig. 3. It can seen that the KL divergence increases with increasing the MSR k. More importantly, when k is greater than 0.2, the KL divergence increases very quickly. Specifically, the KL divergence is equal to 3.3×10^{-5} when $k = 0.4$, and drops substantially to 1.8×10^{-7} when $k = 0.2$. Therefore, when the MSR k is less than 0.4, it can be reasonably considered that the random variable $S = MX$ obeys the normal distribution with zero mean. Furthermore, the variance of the random variable S is obtain easily by

$$\mathbb{E}\left[S^2\right] = \mathbb{E}\left[X^2\right] \mathbb{E}\left[M^2\right] = \mu^2 + \sigma^2 \tag{12}$$

where $\mathbb{E}[X]$ is the expectation of the random variable X.

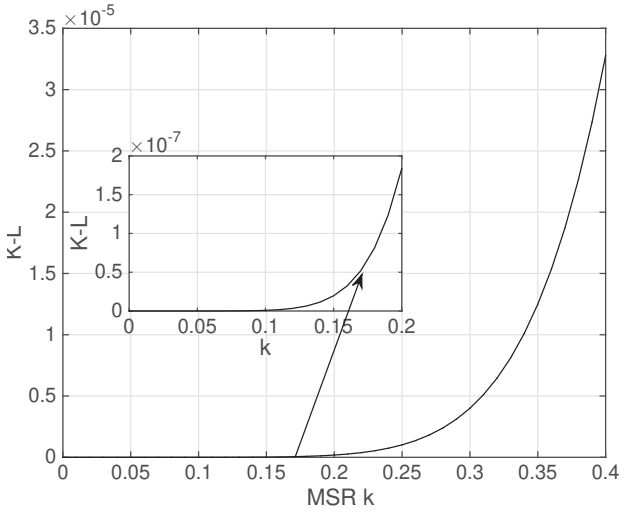


Fig. 3. Kullback-Leibler divergence between the exact PDF and approximate PDF of the random variable S.

2.2.2 Relative Error

The relative error is an important role to quantify the accuracy of an approximation expression. In particular, the relative error η is defined to be

$$\begin{aligned} \eta &= \frac{f_S(s) - g_S(s)}{f_S(s)} \times 100\% = \left(1 - \frac{g_S(s)}{f_S(s)}\right) \times 100\% \\ &= \left(1 - \frac{\frac{2}{\sqrt{1+k^2}} \exp\left(-\frac{s^2}{2(1+k^2)}\right)}{\exp\left(-\frac{(s-k)^2}{2}\right) + \exp\left(-\frac{(s+k)^2}{2}\right)}\right) \times 100\% \,. \end{aligned} \tag{13}$$

Similar to the KL divergence, the relative error η is also independent of the standard deviation σ of the noise X.

The relative error curves of the random variable S in the range of $[0, 3\sigma_S]$ is plotted in Fig. 4 where the MSR k is from 0.1 to 0.4 with step 0.1. Note that the functions $f_S(s)$ and $g_S(s)$ are even symmetric with respective to $s = 0$, it is easy to know that the relative error η is also even symmetric. Hence, only the curves with $s \geq 0$ are plotted.

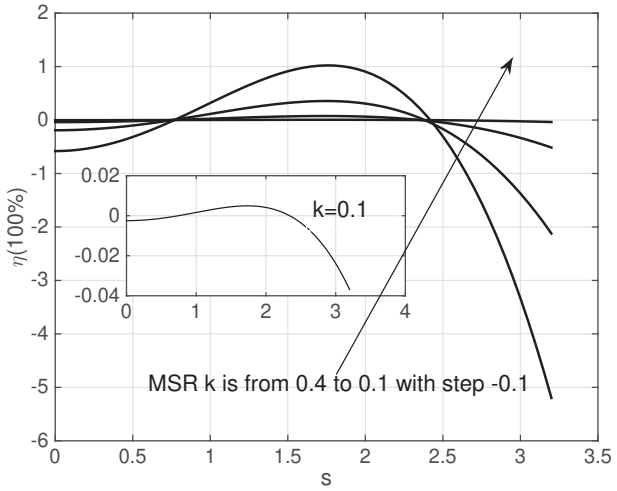


Fig. 4. Relative error between the exact PDF and approximate PDF of the random variable S.

According to the "3σ" criterion of the normal distribution, the probability of 99.7% falls within the range of $[-3\sigma_S, 3\sigma_S]$. It can be seen from the figure that the maximum relative error is within 6%, this approximation is acceptable. In particular, the smaller the MSR k, the better the approximation. For examples, the maximum relative error is within 2% when the MSR $k = 0.3$, while the maximum relative error is significantly reduced to 0.4% when $k = 0.2$.

2.3 Receiver

The output after the additive white Gaussian noise channel is

$$R = S + W = MX + W \tag{14}$$

where W is the AWGN with zero mean and σ_W^2 variance, i.e., $W \sim \mathcal{N}(0, \sigma_W^2)$. In this study, the signal-to-noise ratio (SNR) is defined as the ratio of the average power of the transmitted signal to the average power of the additive white Gaussian noise, i.e.,

$$r = \frac{\sigma_S^2}{\sigma_W^2} = \frac{\mu^2 + \sigma^2}{\sigma_W^2} = \frac{(1 + k^2)\sigma^2}{\sigma_W^2} \,. \tag{15}$$

Referring to the proposed covert communication system shown in Fig. 1, let M' be the pseudo-random sequence on the receiver, then the input of the mean estimator is

$$Y = M'R = M'MX + U = \begin{cases} X + U, & \text{sync} \\ U, & \text{out-sync} \end{cases} \tag{16}$$

where $U = M'W$. The PDF of U can be obtained from Theorem 1, given by

$$f_U(u) = \frac{1}{\sqrt{2\pi}\sigma_W} \exp\left(-\frac{u^2}{2\sigma_W^2}\right) \,. \tag{17}$$

It means that $M'W$ and W have the same PDF, that is, $M'W$ is also a normal random variable with zero mean.

In this study, a shortened part of m-sequence or Gold sequence is chosen as pseudo-random sequence. Specifically, according to Theorem 1, the only requirement for a random sequence is that the sequence is equal probability. When the sequence length is long enough, for example, the length n is 100, 200, or 400, this condition is easily satisfied. According to their properties of autocorrelation and cross-correlation, the expression $M'M = 1$ holds when the pseudo-random generator of the receiver is synchronized perfectly with that of the transmitter, while $M'M$ is approximately zero when they are out of sync. Therefore, if the eavesdropper's pseudo-random sequence is out of sync with the transmitter's, then the input of the mean estimator, Y, does not have any information on the modulated signal X. The corresponding result is that the eavesdropper cannot obtain the covert bits. Obviously, from a security perspective, the more frequent the pseudo-random sequence changes, the better the security. For example, both parties in a legal communication have GPS devices. Both parties agree to change the mask of the random generator every second according to the GPS time. In this case, the eavesdropper is more difficult to intercept the covert communication.

3. Performance Analysis

In this section, we will derive the expression of the mean estimator $\hat{\mu}$ as well as its PDF. Combined with the hard decision in (2), bit error rate of the proposed communication system is obtained.

3.1 Mean Estimator and its PDF

Suppose the pseudo-random generator of the receiver has been synchronized with the transmitter, the sequence after the additive white gaussian noise channel is $\{r_i, i = 1, \cdots, n\}$, and then multiplies it with the synchronous bipolarity pseudo-random sequence $\{m_i, i = 1, \cdots, n\}$ to obtain the $\{y_i = m_i r_i, i = 1, \cdots, n\}$ sequence. The sample mean estimator is used to estimate the mean of the samples $\{y_i\}$, given by

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} m_i r_i = \frac{1}{n}\sum_{i=1}^{n} (x_i + m_i w_i) \,. \tag{18}$$

Accordingly, the mean of the estimator is

$$\mathbb{E}[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[x_i + m_i w_i\right] = (-1)^b \mu \,, \tag{19}$$

and its variance is

$$\begin{aligned} \sigma_{\hat{\mu}}^2 &= \mathbb{E}\left[\hat{\mu}^2\right] - \mathbb{E}^2\left[\hat{\mu}\right] = \frac{\sigma^2 + \sigma_W^2}{n} \\ &= \frac{\sigma^2}{n}\left(1 + \frac{1 + k^2}{r}\right) \,. \end{aligned} \tag{20}$$

Note that both the random variables x_i and $m_i w_i$ follow a normal distribution, specifically, $x_i \sim \mathcal{N}((-1)^b \mu, \sigma^2)$ and $m_i w_i \sim \mathcal{N}(0, \sigma_W^2)$. Therefore, the estimator follows a normal distribution with its PDF given by

$$f(\hat{\mu}) = \frac{1}{\sqrt{2\pi}\sigma_{\hat{\mu}}} \exp\left(-\frac{\left(\hat{\mu} - (-1)^b \mu\right)^2}{2\sigma_{\hat{\mu}}^2}\right) \,. \tag{21}$$

3.2 BER

The derivation of the theoretical bit error rate in AWGN channel is as follows. The PDF of the mean estimation is given in (21), and then combined with the hard decision in (2), the error bit rate (BER) of the proposed communication system is

$$\begin{aligned} \rho &= \frac{1}{2}\mathrm{Erfc}\left(\frac{\mu}{\sqrt{\frac{2\left(\sigma^2 + \sigma_W^2\right)^2}{n}}}\right) \\ &= \frac{1}{2}\mathrm{Erfc}\left(\sqrt{\frac{n}{2}}\frac{k}{\sqrt{1 + \frac{1+k^2}{r}}}\right) \end{aligned} \tag{22}$$

where the complementary error function is define as $\mathrm{Erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} \mathrm{e}^{-t^2} \mathrm{dt}$.

It can be seen from (22) that the bit error rate ρ can be reduced by increasing the number of bit samples n and the mean of the noise sequence μ (or MSR k), as well as increasing the signal-to-noise ratio r. However, the increasing of sampling number n lead to the decrease of bit transmission rate. Increasing the mean value of the noise generator means increasing the transmitting power of the covert signal, thus reducing the concealment of the system.

It can be seen from Fig. 5 that in case of low SNR, the BER is relatively high, that is, the Gaussian noise of the channel has a great influence on the estimator. On the other hand, when the SNR is relatively large, e.g., $r = 25$ dB, there is an error floor, that is, no matter how high the SNR is, the BER can no longer be further reduced. This is because when the SNR approaches infinity, it is the ideal channel, so the error floor is the bit error rate of the ideal channel, which is equal to $\frac{1}{2}\mathrm{Erfc}\left(\sqrt{\frac{\pi}{2}}k\right)$. Moreover, when the SNR is less than 0, that is, the transmitted signal power is lower than the channel noise power, the requirement of a given BER can also be achieved as long as the number of samples per bit period n is large enough (the transmission rate of covert bits is reduced as a side effect). For example, if the system requires a BER of 10^{-3}, the SNR can be as low as -4 dB when the sample length $n = 300$ and MSR $k = 0.3$.

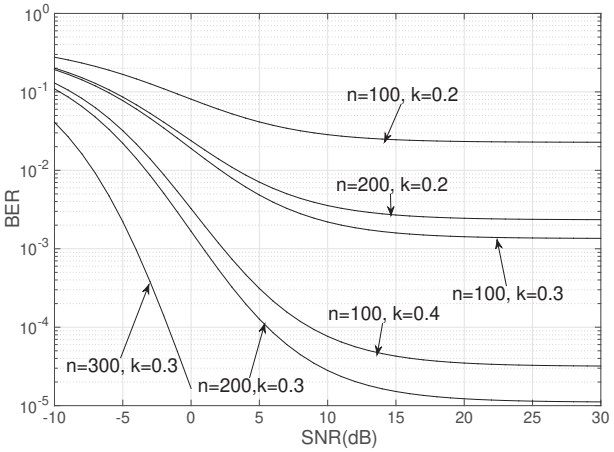


Fig. 5. Theoretical BER under different bit period samples n and MSR k in AWGN channel.

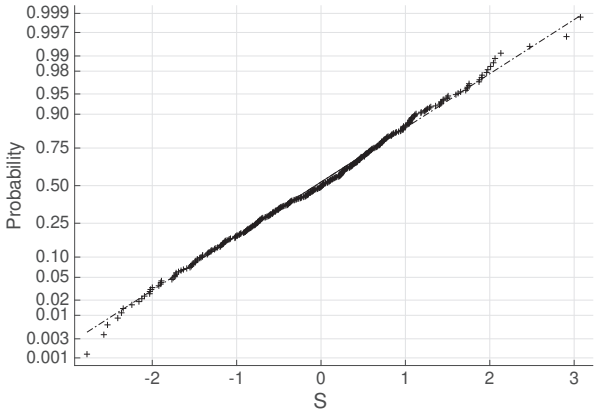


Fig. 6. Normal probability plot of the random variable S.

4. Monte Carlo Evaluation

To verify whether the transmitted signal S is consistent with the normal distribution, simulation is used to examine the normal probability plot of the signal S, and the result using the MATLAB's normplot function is shown in Fig. 6. The plot has the sample data displayed with the plot symbol '+'. Superimposed on the plot is a line joining the first and third quartiles of the sample data (a robust linear fit of the sample order statistics). The purpose of introducing the normplot function in the study is to evaluate how close the random variable is to the normal distribution. If the sample data are normal, the plot will be linear. Other distributions will introduce curvature in the plot. Figure 6 shows that the transmitted signal is completely normally distributed. Therefore, in a covert binary bit period, although the output sequence of the noise generator is a normal distribution of non-zero mean, it is approximately a normal distribution of zero mean after multiplying with bipolar pseudo random sequence. In this way, the transmitted signal containing covert information is statistically consistent with the ambient noise, which is modeled as a normal distribution with zero mean. Therefore, it is difficult for an eavesdropper to distinguish whether the received signal is a transmitted signal containing covert bits or ambient noise, which makes the proposed covert communication system highly concealed.

Next, the sample autocorrelation function (ACF) is used to illustrate the aperiodic nature of the transmitted signal, as shown in Fig. 7, where the number of the samples $n = 100$ in each bit period. The ACF is used to detect the correlation between samples of the transmitted signal $\{s_i, i = 1, 2, \cdots\}$.

According to [20], the formula for the autocorrelation for lag i is

$$r_i = \frac{c_i}{c_0} \tag{23}$$

where $c_i = \frac{1}{n}\sum_{l=1}^{n-i}(s_l - \overline{s})(s_{l+i} - \overline{s})$ and $\overline{s}$ is the mean of the sequence $\{s_i, i = 1, 2, \cdots\}$. In Fig. 7, the largest lag is 4 times the sample length, and the periodicity of the transmitted signal is still not observed. Therefore, the message is camouflaged in the time domain by avoiding intruder to establish correlation.

Next, we focus on the value of MSR, which is a key parameter. The following example indicates that the transmitted signal no longer follows the normal distribution when the value of MSR is large (e.g. $k = 1$). As shown in Fig. 8(a), the PDF curve is almost flat within the range of $[-0.5, 0.5]$. This indicates that when the MSR $k = 1$, the transmitted signal no longer obeys the normal distribution. Furthermore, simulation is used to examine the normal probability plot of the received signal, and the result using the MATLAB's normplot function is shown in Fig. 8(b).The received signal

deviates significantly from the normal distribution. Therefore, on the receiving end, it is possible to distinguish whether the received signal is ambient noise or the presence of the transmitted signal by checking whether the received signal is normally distributed.

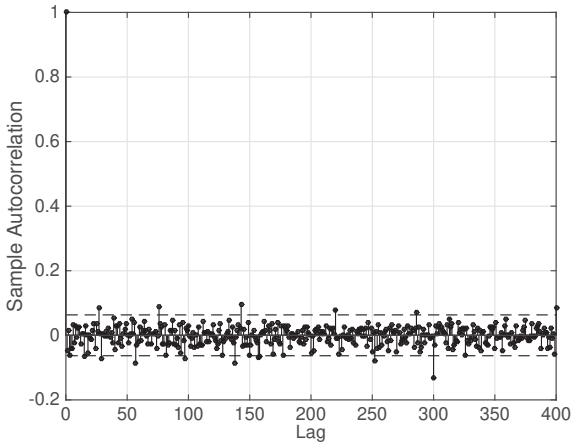
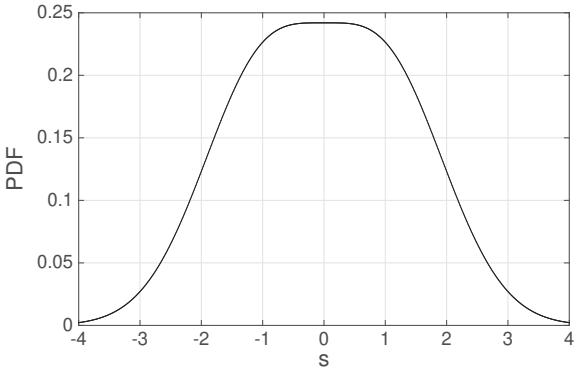
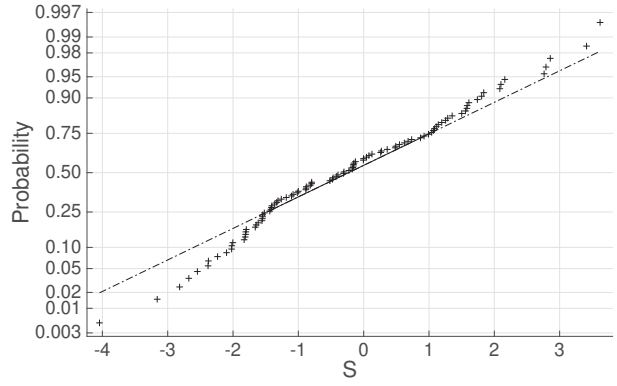


Fig. 7. Sample autocorrelation of the transmitted sequence, where the sample length in each bit period is 100.



(a) Probability density function



(b) Normal probability plot

Fig. 8. The PDF and normal probability plot of the received signal with length $n = 400$, MSR $k = 1$ and SNR $r = 10$ dB.

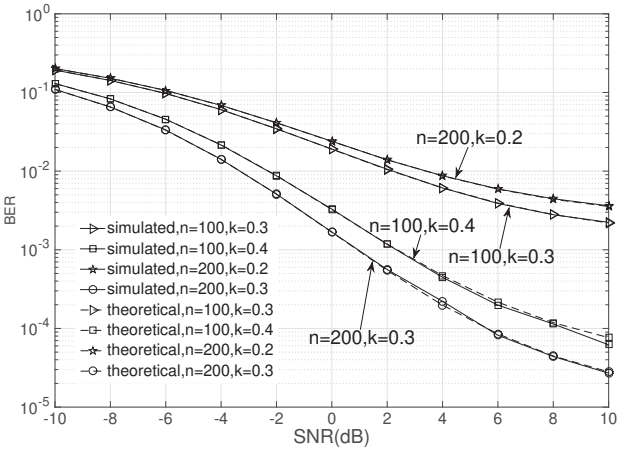


Fig. 9. BER results in AWGN channel with different n and k, where the solid line presents the simulated BER and the dash line presents the theoretical BER.

Finally, the BER results in AWGN channel are shown in Fig. 9 with respect to different MSR k and the sampling length per each bit period n, where the solid line presents the simulated BER and the dash line presents the theoretical BER. Firstly, the simulation results are very consistent with the theoretical derivation. Secondly, it can be seen that in case of the same number of samples n, increasing MSR k has a significant effect on reducing BER. Furthermore, the larger the channel SNR is, the greater the BER reduction is. For example, with a BER of 2×10^{-3} and $n = 100$, there is approximately 9 dB gain when the MSR k increases from 0.3 to 0.4, whereas the power of the transmitted signal increased by $(1 + 0.4^2)/(1 + 0.3^2) = 1.06$ times. Thirdly, if the MSR k is the same, the BER can also be significantly reduced when the number of samples n increases, and the consequence is that the transmission rate of the covert bits also decreases proportionally. Similarly, the larger the channel SNR is, the greater the BER reduction is. Finally, no matter how high the SNR is, the BER can no longer be further reduced, i.e., there is an error floor.

5. Conclusion

In this paper, we put forward a covert communication system using normal distribution with non-zero mean. The non-zero mean normal stochastic sequence modulated by the covert bit is multiplied with the bipolar equiprobable pseudo-random sequence to make the resultant sequence symmetrical about 0. Since the transmitted signal is statistically very similar to ambient noise, it is difficult for an eavesdroppers to distinguish it to achieve the purpose of covert communication. The proposed system works in the physical layer with the advantages of simple structure, strong concealment, good BER performance and very suitable for low-cost, resource-limited and low-rate transmission devices.

Acknowledgments

The authors thank to anonymous reviewers for their time and effort spent in evaluating our manuscript and providing us with constructive comments to improve the quality of the manuscript. This work is supported in part by the National Natural Science Foundation of China under Grant no. 61871348, and in part by Peng Cheng Laboratory under Grant no. PCL2018KP002, and in part by Shenzhen Science and Technology Program under Grant no. JSGG20180508151852303.

References

[1] BANGERTER, B., TALWAR, S., AREFI, K., et al. Networks and devices for the 5G area. *IEEE Communications Magazine*, 2014, vol. 52, no. 2, p. 90–96. DOI: 10.1109/MCOM.2014.6736748

[2] ZHAO, J., NI, S., YANG, L., et al. Multiband cooperation for 5G HetNets: A promising network paradigm. *IEEE Vehicular Technology Magazine*, 2019, vol. 14, no. 4, p. 85–93. DOI: 10.1109/MVT.2019.2935793

[3] NI, S., ZHAO, J., GONG, Y. Optimal pilot design in massive MIMO systems based on channel estimation. *IET Communications*, 2017, vol. 11, no. 7, p. 975–984. DOI: 10.1049/iet-com.2016.0889

[4] LIU, X., JIA, M., NA, Z., et al. Multi-modal cooperative spectrum sensing based on Dempster-Shafer fusion in 5G-based cognitive radio. *IEEE Access*, 2018, vol. 6, p. 199–208. DOI: 10.1109/ACCESS.2017.2761910

[5] ZHAO, J., LI, Q., GONG, Y., et al. Computation offloading and resource allocation for cloud assisted mobile edge computing in vehicular networks. *IEEE Transactions on Vehicular Technology*, 2019, vol. 68, no. 8, p. 7944–7956. DOI: 10.1109/TVT.2019.2917890

[6] ZANDER, S., ARMITAGE, G., BRANCH, P. A survey of covert channels and countermeasures in computer network protocols. *IEEE Communications Surveys & Tutorials*, 2007, vol. 9, no. 3, p. 44–57. DOI: 10.1109/COMST.2007.4317620

[7] SALBERG, A., HANSSEN, A. Secure digital communications by means of stochastic process shift keying. In *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove (USA), 1999, vol. 2, p. 1523–1527. DOI: 10.1109/ACSSC.1999.832004

[8] SALBERG, A., HANSSEN, A. A novel modulation method for secure digital communications. In *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing*. Pocono Manor (USA), 2000, p. 650–654. DOI: 10.1109/SSAP.2000.870206

[9] CHOPRA, A., EVANS, B. Joint statistics of radio frequency interference in multiantenna receivers. *IEEE Transactions on Signal Processing*, 2012, vol. 60, no. 7, p. 3588–3603. DOI: 10.1109/TSP.2012.2192431

[10] YANG, X., PETROPULU, A. Co-channel interference modeling and analysis in a Poisson field of interferers in wireless communications. *IEEE Transactions on Signal Processing*, 2003, vol. 51, no. 1, p. 64–76. DOI: 10.1109/TSP.2002.806591

[11] WIN, M., PINTO, P., SHEPP, L. A mathematical theory of network interference and its applications. *Proceedings of the IEEE*, 2009, vol. 97, no. 2, p. 205–230. DOI: 10.1109/JPROC.2008.2008764

[12] GULATI, K., EVANS, B., ANDREWS, J., et al. Statistics of co-channel interference in a field of Poisson and Poisson-Poisson clustered interferers. *IEEE Transactions on Signal Processing*, 2010, vol. 58, no. 12, p. 6207–6222. DOI: 10.1109/TSP.2010.2072922

[13] NASSAR, M., GULATI, K., SUJEETH, A., et. al. Mitigating near-field interference in laptop embedded wireless transceivers. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas (USA), 2008, p. 1405–1408. DOI: 10.1109/ICASSP.2008.4517882

[14] ABDULLAH, W., CHUAH, T., ABIDIN, A., et al. Measurement and verification of the impact of electromagnetic interference from household appliances on digital subscriber loop systems. *IET Science, Measurement & Technology*, 2009, vol. 3, no. 6, p. 384–394. DOI: 10.1049/iet-smt.2009.0002

[15] CEK, M., SAVACI, F. Stable non-Gaussian noise parameter modulation in digital communication. *Electronics Letters*, 2009, vol. 45, no. 24, p. 1256–1257. DOI: 10.1049/el.2009.2280

[16] XU, Z., WANG, K., GONG, Y., et al. Structure and performance analysis of an SαS-based digital modulation system. *IET Communications*, 2016, vol. 10, no. 11, p. 1329–1339. DOI: 10.1049/iet-com.2015.0761

[17] CEK, M. Covert communication using skewed α-stable distributions. *Electronics Letters*, 2015, vol. 51, no. 1, p. 116–116. DOI: 10.1049/el.2014.3323

[18] XU, Z., GONG, Y., WANG, K., et al. Covert digital communication systems based on joint normal distribution. *IET Communications*, 2017, vol. 11, no. 8, p. 1282–1290. DOI: 10.1049/iet-com.2016.1333

[19] SHLENS, J. *Notes on Kullback-Leibler Divergence and Likelihood Theory*. Google Research. 2014, 4 pages. [Online]. Available at: https://arxiv.org/pdf/1404.2000.pdf

[20] BOX, G., JENKINS, G., REINSEL, G. *Time Series Analysis: Forecasting and Control*. 3rd ed. NJ (USA): Prentice Hall, 1994. ISBN: 9781118745113

About the Authors . . .

Zhijiang XU was born in 1973. He received his Ph.D. degree in Information and Communication Engineering in 2005, from Zhejiang University, China. He held an appointment as associate professor in the College of Information Engineering at Zhejiang University of Technology, China, from 2007 to 2019. Since 2019, he has joined Zhejiang Institute of Mechanical & Electrical engineering as an associate professor in the School of Automation. His research interests include digital communications over fading channels, channel modeling, coding and digital synchronization, etc.

Jingyu HUA was born in Zhejiang province, China in 1978. He received the B.S. and M.S. degrees in Electronic Engineering from the South China University of Technology, Guangzhou, China, in 1999 and 2002. Then in 2006, he received the Ph.D. degree in Electronic Engineering from Southeast University, Nanjing, China. Since 2006, he had joined Zhejiang University of Technology as an assistant professor in the Electronic Engineering Department, and promoted as full professor in 2012. From 2019, he is with Zhejiang Gongshang University as a distinguish professor. He is the author of more than 200 articles and more than 20

inventions. His research interests include the area of parameter estimation, channel modeling, wireless localization and digital filtering in wireless communications. He is currently an associate editor for the IEEE Transactions on Instrumentation and Measurements.

Wenbing JIN (corresponding author) was born in Zhejiang province, China in 1966. He received the B.S. degree in Physics & Electronic Technology and M.S degree in Control Engineering from Zhejiang University, China. Since 2001, he had joined Zhejiang Institute of Mechanical & Electrical Engineering as full professor in the School of Automation. He is the author of more than 30 articles and the editor-in-chief of the two books of the National Twelfth Five-Year Plan. He has received honors from the fourth National College Teaching Master, the first National "Ten Thousand People Program" support object, etc. His research interests include IoT, decentralized control system, etc.