

Direct Symbol Decoding using GA-SVM in Chaotic Baseband Wireless Communication System

Hui-Ping Yin^{a,1}, Hai-Peng Ren^{a,*}

^a*Shaanxi Key Laboratory of Complex System Control and Intelligent Information Processing, Xi'an University of Technology, Xi'an, China*

Abstract

To retrieve the information from the serious distorted received signal is the key challenge of communication signal processing. The chaotic baseband communication promises theoretically to eliminate the inter-symbol interference (ISI), however, it needs complicated calculation, if it is not impossible. In this paper, a genetic algorithm support vector machine (GA-SVM) based symbol detection method is proposed for chaotic baseband wireless communication system (CBWCS), by this way, treating the problem from a different viewpoint, the symbol decoding process is converted to be a binary classification through GA-SVM model. A trained GA-SVM model is used to decode the symbols directly at the receiver, so as to improve the bit error rate (BER) performance of the CBWCS and simplify the symbol detection process by removing the channel identification and the threshold calculation process as compared to that using the calculated threshold to decode symbol in the traditional methods. The simulation results show that the proposed method has better BER performance in both the static and time-varying wireless channels. The experimental results, based on the wireless open-access research platform, indicate that the BER of the proposed GA-SVM based symbol detection approach is superior to the other counterparts under a practical wireless multipath channel.

Keywords: GA-SVM, ISI, Pattern classification, CBWCS, Symbol decoding

*Corresponding author

Email address: renhaipeng@xaut.edu.cn (Hai-Peng Ren)

¹Hui-Ping Yin (e-mail:huipingyin@qq.com)

1. Introduction

Filter design and coherent/noncoherent decoding are the key points in communication signal processing [1]. Researchers from communication field like the analytical decoding methodologies with rigorous theoretical base. However, that might lead to more and more complicated algorithm in complex channel communication [2, 3, 4]. By turning around, we treat the binary information decoding in the communication systems as a pattern recognition problem. It can be simply dealt with using the machine learning (ML) approach, which requires no direct analytical support, but works well due to its deepen insight treatment of the signal. In this paper, we illustrate this point using the symbol decoding problem in chaotic baseband wireless communication system [5] as a paradigm.

Owing to its intrinsic properties, such as high sensitivity upon initial conditions, wideband, orthogonality, easy to generate, chaos has gained momentum in providing effective solutions to solve significant practical problems in the communication fields [6, 7, 8, 9]. Since early 1990s, chaos has been used in communication field [10, 11, 12], mainly focusing on the theoretical analysis in ideal channel with white noise in most of the early research. Until Ref [13] reported that chaos was successfully used in a commercial wired fiber-optic channel and achieved higher bit transmission rate in 2005, the research on chaos-based communication has turned to practical communication channel. Due to limited bandwidth, multipath propagation, doppler shift, and complex noise, the wireless channel seriously distorts the signal transmitted in it. In terms of this issue, Ref [14] proved the Lyapunov spectrum invariance property of chaotic signal for the first time after transmitting through the wireless channel, which shows chaos can be used to transmit information through wireless channel without information loss. At the same time, more properties of chaos benefiting communication applications have been found, the properties include that chaotic signal has a very simple corresponding matched filter to maximize the signal to noise ratio

[15, 16], and the chaos property can be used to theoretically remove inter-symbol interference (ISI) caused by multipath propagation completely [2, 5]. However, this theoretical ISI free decoding method needs three parts of information including 1) channel information, which can be identified using the probe signal; 2) the past decoded information bits, which are available at the current time for the threshold calculation; 3) the future information bits to be transmitted, which are not available at the current time. To avoid using the future bits, Ref [2] proposed a suboptimal decoding threshold, experimental tests in [5] showed the CBWCS achieved better BER performance using the suboptimal threshold with lower computation cost as compared to the conventional wireless communication system with complex channel equalization. Although the results in [5] are attracting enough, it leaves margin for further improvement. The first attempt was predicting the future chaotic baseband waveform in [3], which was used to decode future information bits. Due to the long term prediction difficulty of the chaotic signal, only one future bit is derived and used together with the past information bits to calculate more accurate threshold. Both simulation and experimental results verified the improved BER performance using the predicted future bit information. The methods in both [3] and [5] needed to identify the channel information for decoding. To avoid this operation, the threshold was predicted directly in [4] using the echo state network (ESN). This method not only avoids using the identified channel information in decoding but also avoids the error caused by the inaccurate identified channel information, therefore, simulation and experimental results are not surprising to show the superior BER performance with respect to the methods in [3] and [5]. All above attempts to improve the BER performance is started from the conventional communication engineering thought, that is, to decode information using single feature, decoding threshold. By considering this problem from pattern recognition viewpoint, the decoding process is a kind of two-category classification problem. Therefore, machine learning (ML) becomes a powerful tool to solve such problems. Following this novel insight, a genetic algorithm (GA) optimized support vector machine (SVM) is proposed to decode the information bit directly in CBWCS

in this paper.

As a promising classification method, machine learning has been attracting much research interest from various fields due to numerous successful applications to solve practical problems [17, 18, 19, 20, 21, 22, 23]. In the field of communication and signal processing, different types of machine learning methods, including unsupervised learning, reinforcement learning, and supervised learning, find significant applications in many cases [24, 25, 26]. SVM is one of the machine learning methods that has state-of-the-art capability for classification, and it has been applied in various fields [27, 28, 29]. The idea behind the SVM is to map the input vectors into a high-dimensional feature space in which they become linearly separable. The mapping from the input vector space to the feature space is non-linear which can be implemented in a convenient and efficient way through the kernel functions (i.e., the so-called kernel trick). A SVM classifier is in principle a binary classifier, this means that it can predict or classify input data belonging to two distinct classes. SVM has been employed to design link adaptation methods for wireless communication networks [30, 31, 32]. There are extensive studies on not only improving the accuracy of SVMs [33], but also applying SVM to various classification problems such as signal detection in visible light communications [34]. However, its usage in communication systems has not been well explored.

In this paper, a SVM is used into CBWCS to decode the information bit directly according to the waveform received. The SVM is utilized for the classification task [35, 36], and its parameters are optimized by the genetic algorithm (GA) to explore the best classification ability of SVM. The global optimizing GA helps to decrease the SVM structure risk and rent the SVM ability to use small training data set in order to achieve desired generalization performance. This point is very important for practical communication system application because larger set of training data means the channel bandwidth is occupied more.

The main contributions of this work are summarized as follows: first, we identify the information symbol decoding process for chaotic baseband wireless

communication as a binary classification problem for the first time; second, we suggest an appropriate learning approach, including the training data with the corresponding class labels and GA based SVM parameters optimization; third, we test the proposed method in both simulation and experiment to show its superiority with respect to the comparison methods.

Numerical simulations and experiments based on software defined radio system are extensively performed in order to evaluate the proposed GA-SVM based symbol detection method. The proposed method achieves better performance with simplified decoding process as compared to the decoding methods using the calculated threshold [2, 3, 5].

The rest of the paper is organized as follows. In Section 2, we introduce the GA-SVM based symbol detection approach in chaotic baseband wireless communication system. In Section 3, the performance of the GA-SVM based symbol detection method is compared with the other symbol detection methods using decoding threshold in simulation. In Section 4, the proposed method is evaluated in experiment based on wireless open-access research platform (WARP) system. Finally, we conclude this paper in Section 5.

2. GA-SVM based symbol decoding

2.1. Working principle of CBWCS

The block diagram of CBWCS is given in figure 1. From figure 1, we know that the information bits (symbols) to be transmitted are sent to the chaotic shape-forming filter (CSF) [37] to generate the chaotic baseband signal $x(t)$, and the baseband signal is fed into the up-carrier for channel modulation, then, the chaotic signal is transmitted through wireless channel. At the receiver, the signal is captured by antenna, and fed into down-carrier to remove the carrier frequency. The down-carrier output $r(t)$, is fed into the corresponding matched filter (MF) in order to get the maximum signal to noise ratio (SNR) at the predefined sampling point from the filter output signal $y(t)$. The sampling

point with maximum SNR is compared with the calculated threshold to decode the symbol (bit) transmitted.

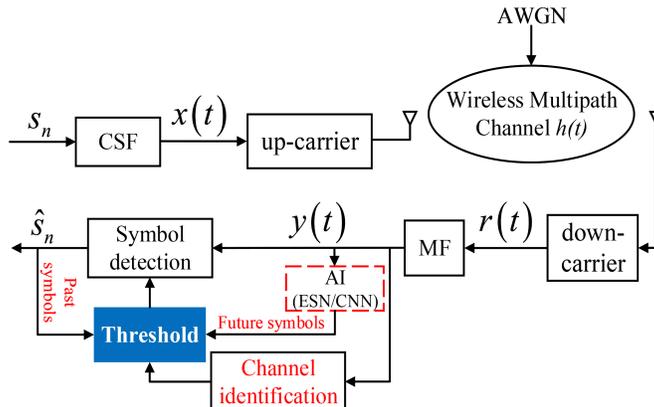


Figure 1: Block diagram of the threshold-based symbol detection in CBWCS.

The threshold used in the CBWCS is one of the key point from conventional communication engineering viewpoint. Zero could be used as threshold, but it has the worst performance, because this means the inter-symbol interference (ISI) from past and future symbols is not considered [3]. Suboptimal threshold is proposed in [2], which is calculated by using the decoded past symbols and the channel parameters identified from the probe stream [5]. To further consider the future symbol effect, an echo state network (ESN) is proposed in [3] to predict the future output waveform of the matched filter, from which the future one symbol is predicted, finally, the decoded one future symbol is used together with the past decoded symbols and the identified channel information in order to obtain more accurate decoding threshold, in this way, the BER performance is improved. In order to avoid the waveform prediction difficulty in [3], reference [4] proposes to predict the decoding threshold directly, instead of predicting waveform then decoding a future symbol from the waveform, which simplifies the procedure and achieves better BER performance, because the effect of imperfect waveform prediction is removed. Here, one point to be noticed is that not only ESN could be used for the symbol prediction but also other AI methods, like

convolutional neural network (CNN) can also play such a role [38] as shown in figure 1.

Now we conclude the above mentioned methods. First of all, all these methods have the same start point, that is, to calculate the threshold derived in [2], which is then used to decode symbol. Second, because of this start point, complicated computations are required, including 1) the channel parameters identification, 2) the future waveform prediction and/or 3) the future symbol prediction, and 4) the threshold calculation. All these procedures bring the error contribution to the final BER.

In fact, could we jump out of the start point above, from another viewpoint, we can treat the decoding process as a pattern recognition problem, that is, classification of two categories from the filter output signal. By this way, the complicated channel parameters identification is removed, the complicated future symbol prediction methods are avoided, therefore, the BER contribution of these procedures is eliminated, which promises to decrease BER.

In the following section, we will give the framework of CBWCS based on direct GA-SVM symbol decoding.

2.2. Symbol decoding based on GA-SVM in CBWCS

Inspired by the pioneering works based on learning-based method in wireless communication, we focus on the symbol detection in chaotic baseband wireless communication system using machine learning approach. The symbol detection in wireless communication is converted to be a binary classification problem. Compared to the symbol detection method using decoding threshold [2, 3, 5], the GA-SVM model provides relatively simplified process.

After we change the start point from the decoding threshold to the final decoding task. We design a simplified framework of CBWCS based on GA-SVM symbol decoding as shown in figure 2, where the decoding part is significantly simplified as the block marked by green with respect to the corresponding part in figure 1.

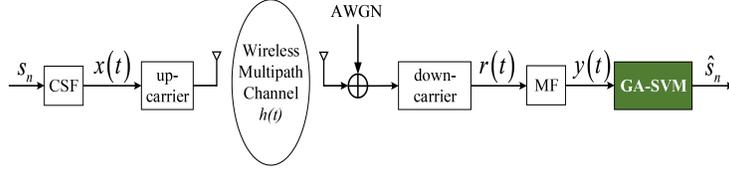


Figure 2: Block diagram of the symbol decoding based on GA-SVM in CBWCS.

From figure 2, we know that one piece of matched filter output is used as input of the classifier, and the classifier output is the decoding symbol.

From the pattern recognition viewpoint, machine learning has achieved lots of advancement [39, 40, 41, 42, 43, 44]. The power of machine learning depends on the training data set and the proper classifier structure. SVM is considered as the most powerful classifier [28, 45, 46] by employing the high dimensional feature representation provided if the proper structure is selected.

In the following sub-sections, we tackle the problems of training data set and the SVM design with minimum structure risk.

2.3. Training data set

To design training data set, we first analyze the feature of the baseband signal in CBWCS [5]. The baseband signal is generated by a chaotic shape-forming filter (CSF) [37] as shown in figure 2, given by $x(t) = \sum_{n=-\infty}^{\infty} s_n \cdot p(t - n/f)$, where the basis function $p(t)$ is given by

$$p(t) = \begin{cases} (1 - e^{-\beta/f}) e^{\beta t} \left(\cos \omega t - \frac{\beta}{\omega} \sin \omega t \right), & (t < 0) \\ 1 - e^{-\beta(t-1/f)} \left(\cos \omega t - \frac{\beta}{\omega} \sin \omega t \right), & (0 \leq t < 1/f) \\ 0, & (t \geq 1/f) \end{cases}, \quad (1)$$

where ω and β are parameters of the basis function satisfying $\omega = 2\pi f$, $0 < \beta \leq f \cdot \ln 2$, and f is the base frequency.

Figure 3 shows the basis function plot for $\beta = \ln 2$ and $f = 1$.

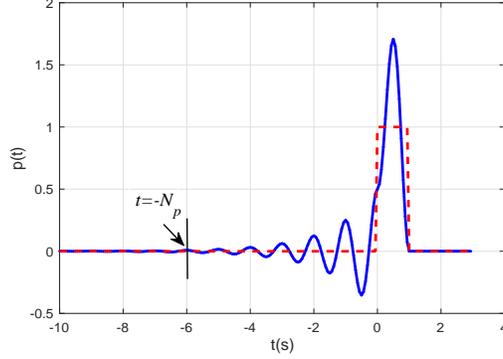


Figure 3: The basis function $p(t)$ for $\beta = \ln 2$ and $f = 1$.

From figure 3, we know that $x(t)$ can be approximated by

$$x(t) = \sum_{n=\lfloor t \rfloor}^{\lfloor t \rfloor + N_p} s_n \cdot p(t - n), \quad (2)$$

where $\lfloor t \rfloor$ indicates the largest integer less than or equal to t , and N_p is a positive integer. From figure 3, we can see that when $N_p = 6$, it satisfies $p(t) \approx 0$ for $t < -N_p$, thus, the value of parameter N_p is set to be 6 in this paper. A statistical average model of wireless channel [47] can be defined

$$h(t) = \sum_{l=0}^{L-1} \alpha_l \delta(t - \tau_l), \quad (3)$$

where α_l and τ_l are the attenuation and propagation delay corresponding to path l from the transmitter to the receiver, and $\delta(\cdot)$ is the Dirac delta function. Assuming the delay τ_l ($l = 0, 1, \dots, L - 1$) satisfies $0 = \tau_0 < \tau_1 < \dots < \tau_{L-1}$, the channel fading α_l can be modeled as a negative exponential decay $\alpha_l = e^{-\gamma \tau_l}$, where γ is the damping coefficient [14].

At the receiver of the CBWCS, the received signal is given by

$$\begin{aligned} r(t) &= h(t) * x(t) + w(t) \\ &= \sum_{l=0}^{L-1} \alpha_l x(t - \tau_l) + w(t) \\ &= \sum_{l=0}^{L-1} \alpha_l \sum_{n=-\infty}^{\infty} s_n p(t - \tau_l - n/f) + w(t), \end{aligned} \quad (4)$$

where ' $*$ ' denotes convolution and $w(t)$ is an Additive White Gaussian Noise (AWGN).

The filter output is given by

$$\begin{aligned}
y(t) &= g(t) * r(t) \\
&= \int_{\tau=-\infty}^{\infty} p(-\tau)r(t-\tau) d\tau \\
&= \sum_{l=0}^{L-1} \alpha_L \sum_{n=-\infty}^{\infty} s_n \left[\int_{\tau=-\infty}^{\infty} p(\tau) p\left(\tau-t+\tau_l+\frac{n}{f}\right) d\tau \right] \\
&\quad + \int_{\tau=-\infty}^{\infty} p(\tau-t) w(\tau) d\tau,
\end{aligned} \tag{5}$$

where $g(t) = p(-t)$ is the impulse response of the matched filter (MF) [15].

From Eqs. (4) and (5), we learn that the received chaotic baseband signal of the current symbol is affected by both the past symbols and the future ones through the multipath wireless channel. This fact makes it possible for us to use the received signal (the past, the current, the future) to decode the current symbol and at the least the past and future N_p symbols affect the decoding of the current symbol.

In the following, we analyze the influence of different combinations of past and future symbols on the waveform corresponding to the current symbol as shown in figures 4 and 5.

We assume the current symbol $s_n = -1$, and the corresponding chaotic waveforms are represented by the blue solid line and red dashed line respectively with different combinations of past and future symbols, where $S_p = [s_{n-6}, \dots, s_n]$ represents the past and current symbols set, $S_f = [s_n, \dots, s_{n+6}]$ represents the current and future symbols set. We can see that only a little difference is caused by different s_{n-6} in figure 4(a) and by different s_{n-5} in figure 4(b); figure 4(c)-(f) show the differences caused by different s_{n-4} , s_{n-3} , s_{n-2} , s_{n-1} respectively; figure 5(a)-(f) show the differences caused by different s_{n+1} , s_{n+2} , s_{n+3} , s_{n+4} , s_{n+5} , s_{n+6} respectively.

From figures 4 and 5, we know that the waveform of the current symbol is influenced by both the past and future symbols, the closer to the current symbol

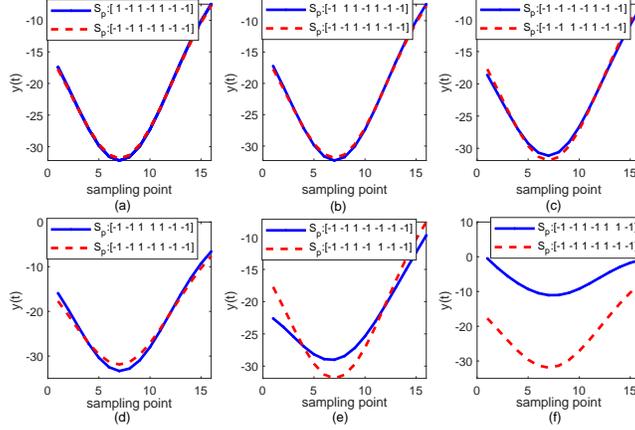


Figure 4: The current symbol waveform difference caused by past symbols.

the bit difference locates, the bigger effect on the waveform corresponding to the current symbol, 4 past and 4 future symbols make main contributions to the current symbol waveform.

To decode the current symbol, the filter output signal corresponding to four past symbols, the current symbol, four future symbols is needed to train the SVM model. 4 past symbols, 1 current symbol, 4 future symbols could generate 2^9 possible combinations, each of which consists of waveform corresponding to 9 bits. So we have the waveform corresponding to 4608 symbols as training data source. But 4608 bits as training data will cost too much time, the efficiency is decreased accordingly. By lessen the restriction, we select the signal corresponding to 3 past symbols, 1 current symbol, and 3 future symbols as the training data set, i.e., 896 bits, thus the amount of training data is greatly reduced. In what follows, we will compare different data sets to see the performance.

In the proposed method, the transmission data is sent frame by frame, the frame structure is shown in figure 6.

The data frame configuration is compatible with other previous methods and the conventional methods, which also helps to cope with the time variant channel state [5]. one data frame consists of two parts, i.e., the probe data and

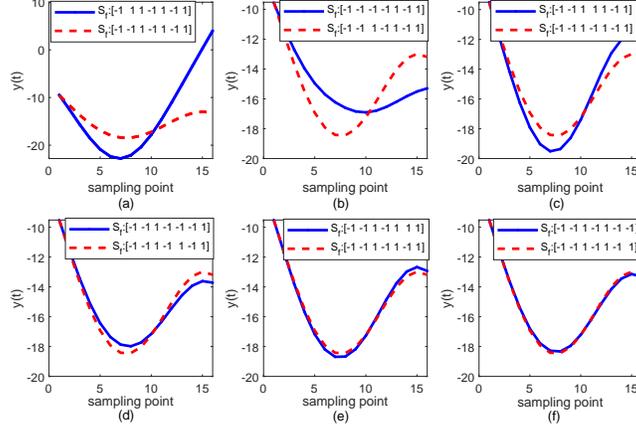


Figure 5: The current symbol waveform difference caused by future symbols.



Figure 6: The data frame in the wireless communication system.

information data. The probe data is also called training data, which is used for frame synchronization, clock synchronization, and the training data of AI method [3, 4] as well, which is known to the receiver. Whereas, the information data is the data to be transmitted. The frame data including the probe data and information data is passed through the chaotic shape-forming filter to obtain the baseband signal, then, the baseband signal is passed through the up-carrier as shown in figure 2, which is transmitted through the transmission antenna. After the received signal is captured by the receiver antenna, the signal is passed through the down-carrier, and matched filter, then the matched filter output corresponding to a frame is divided into two parts, the probe data part is used for frame synchronization, clock synchronization, and SVM training, it is also used for ESN training in the comparison methods. Afterwards, the information data is decoded using the trained SVM and different comparison algorithms. This point guarantees the fair comparison of the different methods. In this paper, the probe data length can be either 4608 or 896, which will be investigated in

the following parts.

2.4. GA-SVM

A SVM classifier is in principle a binary classifier. This means that it can predict or classify input data belonging to two distinct classes. SVMs have proved to be a very efficient machine learning technique [29].

The idea behind the SVMs is to map the input vectors into a high-dimensional feature space in which the features become linearly separable. The mapping from the input vector space to the feature space is non-linear which can be done in a convenient and efficient way through kernel functions $K(\mathbf{u}_i, \mathbf{u}_j)$. Given two vectors as \mathbf{u}_i and \mathbf{u}_j from input space, the output result in the dot product of their mapping in the high-dimensional feature space is

$$K(\mathbf{u}_i, \mathbf{u}_j) = (\Phi(\mathbf{u}_i) \cdot \Phi(\mathbf{u}_j)), \quad (6)$$

where Φ is an explicit mapping from input space to features space [48]. Details about the SVM theory can be found in many excellent tutorials and textbooks [27, 35].

Depending on the applications, different types of kernel functions can be used. The radial basis function (RBF) kernel, $K(\mathbf{u}_i, \mathbf{u}_j) = \exp(-g\|\mathbf{u}_i - \mathbf{u}_j\|^2)$, is used in our work.

We assume that the training data set is $D = \{(\mathbf{u}_1, v_1), (\mathbf{u}_2, v_2), \dots, (\mathbf{u}_m, v_m)\}$, where \mathbf{u} indicates the input vector, the corresponding class label is $v \in \{-1, +1\}$, and sub-index m is the number of training data. Two classes can be separated by a hyperplane $\mathbf{W}^T \Phi(\mathbf{u}) + b = 0$, and it can be converted to be a constrained minimization problem given by

$$\begin{aligned} \min_{\mathbf{W}, b} \quad & \frac{1}{2} \|\mathbf{W}\|^2 + C \sum_{i=1}^m \xi_i \\ \text{s.t.} \quad & v_i (\mathbf{W}^T \Phi(\mathbf{u}_i) + b) \geq 1 - \xi_i, i = 1, 2, 3, \dots, m, \end{aligned} \quad (7)$$

where C is the penalty term, ξ is a slack variable.

The penalty parameter C and the kernel function parameter g affect the classification accuracy of SVM, in most cases, they are chosen by experience. In this paper, the genetic algorithm (GA) is used to search for better combinations of the parameters. It can obtain the optimal solution after a series of iterative computations [49] using the training data. After the GA optimization, a best combination of the parameters fit for the training data is obtained, we refer it to as GA-SVM in this paper.

The optimal values for \mathbf{W} and b can be found by solving the constrained minimization problem in Eq. (7) using the Lagrangian multiplier α_i ($i = 1, \dots, m$) [50]. Thus, the optimal discriminant function is given by

$$f(\mathbf{u}) = \text{sgn} \left(\sum_{i=1}^m \alpha_i v_i K(\mathbf{u}_i, \mathbf{u}) + b \right). \quad (8)$$

In the training data, those \mathbf{u}_i with nonzero α_i are the “support vectors”, if we assume that N_s is the number of the support vectors, Eq. (8) can be rewritten by

$$f(\mathbf{u}) = \text{sgn} \left(\sum_{s=1}^{N_s} \alpha_s v_s K(\mathbf{u}_s, \mathbf{u}) + b \right), \quad (9)$$

where \mathbf{u}_s denotes the support vector, α_s is the nonzero value of the corresponding Lagrangian multiplier.

The block diagram of GA-SVM can be shown by figure 7.

As shown in figure 7, the input vector of the model is $\mathbf{u} = (u^1, u^2, \dots, u^{112})^T$, where u^1, u^2, \dots, u^{112} are the matched filter output signal samples corresponding to 3 past symbols, one current symbol and 3 future symbols, where the over-sampling rate $N_r = 16$, and the output of the model, v , is equal to the current symbol information, i.e., the class label.

Then the trained model using the probe data can be used to decode the symbol from the information data waveform in the CBWCS. For decoding the symbol s_n , the input vector, \mathbf{u}_n , is fed into the trained GA-SVM model, the corresponding output of the model is the current symbol information, i.e., $\hat{s}_n = f(\mathbf{u}_n)$, which is shown as figure 8.

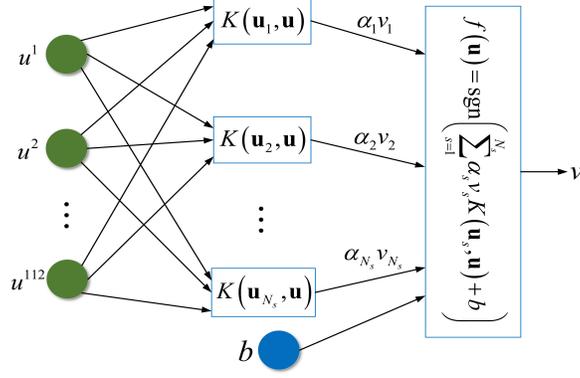


Figure 7: The block diagram of GA-SVM.

3. Simulation results

In this section, we present the simulation results for the GA-SVM based symbol decoding in the chaotic baseband wireless communication system.

To eliminate the effect of inter-symbol interference (ISI), an optimal decoding threshold θ_n is used for detecting the symbol s_n [2], which can be given by

$$\theta_n = I = I_{past} + I_{future}, \quad (10)$$

where $I_{past} = \sum_{l=0}^{L-1} \sum_{i=-\infty}^{-1} s_{n+i} I_{l,i}$, $I_{future} = \sum_{l=0}^{L-1} \sum_{i=1}^{\infty} s_{n+i} I_{l,i}$, and $I_{l,i} = \alpha_l \int_{\tau=-\infty}^{\infty} p(\tau) p\left(\tau + \tau_l + \frac{i}{f}\right) d\tau$ can be calculated by

$$I_{l,i} = \begin{cases} \alpha_l (C - D^{-1}) (A \cos(\omega\tau_l) + B \sin(\omega\tau_l)), \\ if \left(|\tau_l + \frac{i}{f}| \geq \frac{1}{f}\right) \\ \alpha_l \{A (C - D) \cos(\omega\tau_l) + B (C + D) \sin(\omega\tau_l) + 1 - |\tau_l f + i|\}, \\ if \left(0 \leq |\tau_l + \frac{i}{f}| < \frac{1}{f}\right), \end{cases} \quad (11)$$

where $A = \frac{(\omega^2 - 3\beta^2)f}{4\beta(\omega^2 + \beta^2)}$, $B = \frac{(3\omega^2 - \beta^2)f}{4\omega(\omega^2 + \beta^2)}$, and $C = e^{-\beta|\tau_l + \frac{i}{f}|} (2 - e^{-\frac{\beta}{f}})$, $D = e^{\beta|\tau_l + \frac{i}{f}|} e^{-\frac{\beta}{f}}$.

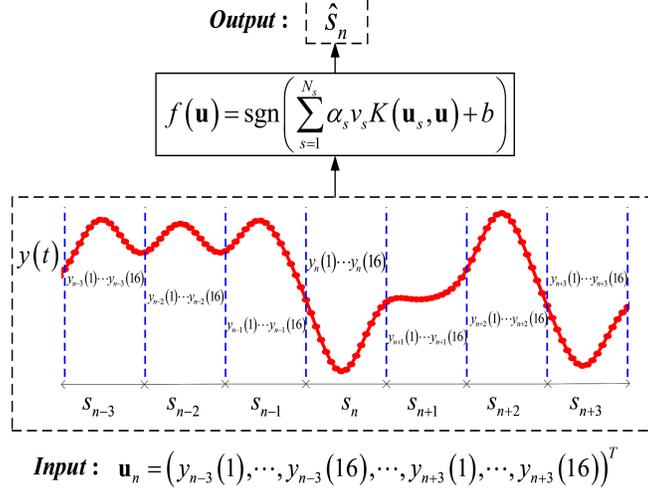


Figure 8: The symbol decoding schematic diagram using GA-SVM.

The BER using the optimal threshold $\theta_n = I$ can be calculated by

$$p(\text{error} | \theta_n = I) = \frac{1}{2} \text{erfc} \left(\frac{P}{\sqrt{2\sigma_w^2}} \right), \quad (12)$$

where $\text{erfc}(\cdot)$ is the complementary error function, $P^2/(2\sigma_w^2)$ is the filtered signal to noise ratio (SNR), and $P = \sum_{l=0}^{L-1} I_{l,0}$ is the sum of the multipath power for the corresponding symbol, L is the number of multipaths, σ_w^2 is the variance of the Gaussian noise w .

The theory bit error rate curves in the simulation results are calculated using Eq. (12) with the corresponding multipath parameters.

3.1. The BER performance of different training data sets

The GA-SVM is trained using 896 bits, 4608 bits, corresponding to the combination of 7 bits and 9 bits, separately, as aforementioned.

In figure 9, we investigate the impact of training data sets of different size on the BER of CBWCS by using 896 bits, 4608 bits versus different SNR values. It can be seen that as the number of the training data increases, the BER of the GA-SVM based symbol decoder decreases.

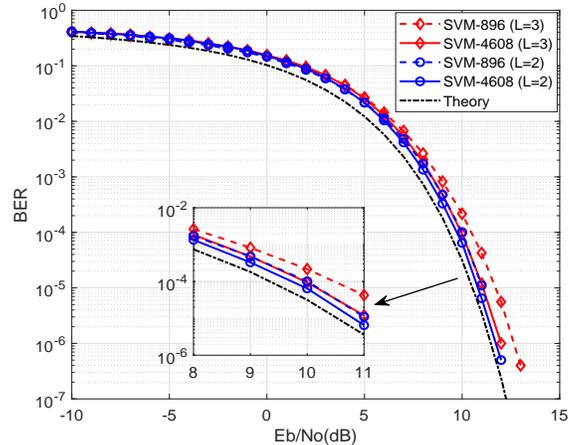


Figure 9: The BER results using training data sets with different size under different multipaths. The solid lines represent the BER results using 4608 bits, the dashed lines represent the BER results using 896 bits. The blue lines with circle markers are the BER curves under 2 paths, and the red lines with diamond markers are the BER curves under 3 paths. Black dash-dotted line is the theory BER corresponding to $L = 2$.

The training data set will grow exponentially considering one more past symbol and one more future symbol correspondingly. The BER comparison results in figure 9 show that the BER performance using 4608 training bits is not improved too much than the performance using 896 training bits, but the number of the training bits is increased greatly, the training complexity is increased correspondingly. By comprehensively considering the performance and cost, 896 bits are used in this paper.

3.2. BER performance of different methods

The simulations under different multipath cases using different decoding methods are performed. Figures 10 and 11 show the BER performance comparison for different decoding methods under 2 and 3 paths, respectively. In figures 10 and 11, the γ is set to be 0.6, and the time delay parameters for figure 10 are $\tau_0 = 0$, $\tau_1 = 1$, for figure 11, $\tau_0 = 0$, $\tau_1 = 1$, $\tau_2 = 2$.

The results using different decoding methods for two-path channel ($L = 2$)

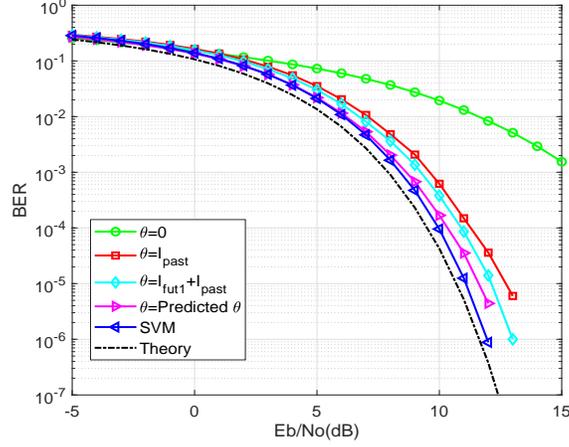


Figure 10: BER results using different decoding methods under 2 paths. Black dash-dotted line is the theory BER corresponding to $L = 2$, blue solid line with left triangle markers is the BER curve using GA-SVM, magenta solid line with right triangle markers is the BER curve using the decoding threshold predicted by ESN, cyan solid line with diamond markers is the BER curve using the decoding threshold $\theta = I_{fut1} + I_{past}$, red solid line with square markers is the BER curve using the decoding threshold $\theta = I_{past}$, green solid line with circle markers is BER curve using zero threshold.

and three-path channel ($L = 3$) in figures 10 and 11 show that the CBWCS has the worst BER by using the decoding threshold without considering the intersymbol interference (ISI) from past and future symbols, i.e., $\theta = 0$. Using the decoding threshold $\theta = I_{past}$ in [2] which considered the ISI caused only by past symbols, the BER is lower. Using the decoding threshold $\theta = I_{past} + I_{fut1}$ which considered the ISI caused by the past symbols and one future symbol predicted by echo state network (ESN) in [3], the BER is much lower. Using the decoding threshold predicted directly by ESN in [4], the BER performance is further improved as compared to the other methods above. Using the direct decoding based on GA-SVM proposed in this paper, the BER is the lowest among them. With the number of multipath increasing, the distance from theoretical BER to BER of the proposed method increases. But the proposed method is still the best among all methods.

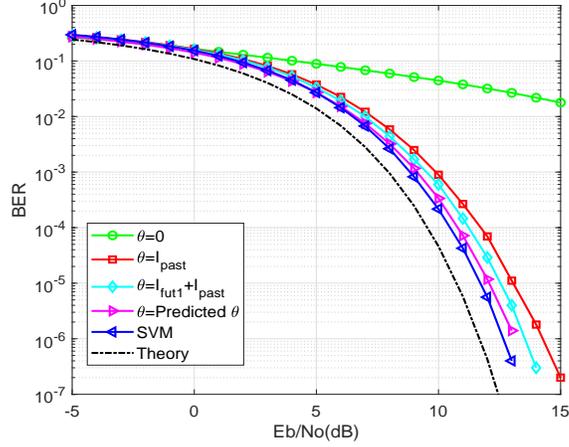


Figure 11: BER results using different decoding methods under 3 paths. Black dash-dotted line is the theory BER corresponding to $L = 3$, blue solid line with left triangle markers is the BER curve using GA-SVM, magenta solid line with right triangle markers is the BER curve using the decoding threshold predicted by ESN, cyan solid line with diamond markers is the BER curve using the decoding threshold $\theta = I_{fut1} + I_{past}$, red solid line with square markers is the BER curve using the decoding threshold $\theta = I_{past}$, green solid line with circle markers is BER curve using zero threshold.

In the simulation cases in figures 10 and 11, the channel parameters are fixed. In fact, the practical channel parameters are time variant. We assume that the channel parameters are unchanged within one data frame, but be variant from one frame to the next. The channel parameter, γ obeys the uniform distribution in the range of $[0.3, 0.9]$. Time delay are $\tau_0 = 0, \tau_1 = 1$ for two-path channel, and $\tau_0 = 0, \tau_1 = 1, \tau_2 = 2$ for three-path channel. Figures 12 and 13 give the BER comparison results under such quasi-time-varying channel corresponding to $L = 2$ and $L = 3$, separately.

In our time-varying channel simulation, the frame structure is 896 probe data and 1152 information bits. The probe bits are used for GA-SVM model training, and for the other methods, the probe bits are used not only for AI training, but also for channel parameters estimation.

From figures 12 and 13, the comparison results using the proposed method

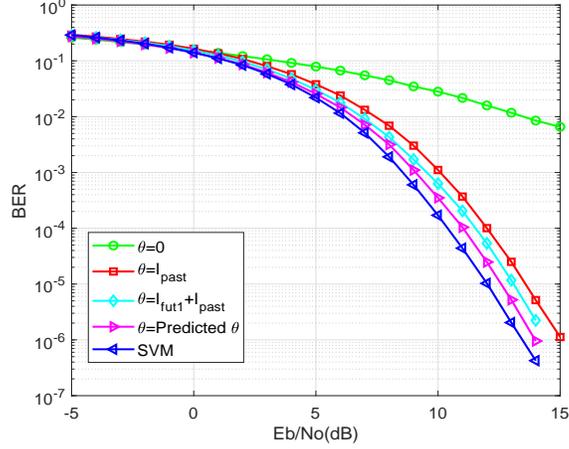


Figure 12: BER results with different decoding methods under 2 paths in a time-varying wireless channel. Blue solid line with left triangle markers is the BER curve using GA-SVM, magenta solid line with right triangle markers is the BER curve using the decoding threshold predicted by ESN, cyan solid line with diamond markers is the BER curve using the decoding threshold $\theta = I_{fut1} + I_{past}$, red solid line with square markers is the BER curve using the decoding threshold $\theta = I_{past}$, green solid line with circle markers is BER curve using zero threshold.

and other methods show that the proposed method has even better result as compared to the static channel case. The conclusion is the same as that can be drawn from figures 10 and 11, that is, the proposed method has the best performance under the time variant channel.

4. Experimental results

To test the proposed method, the experiment is carried out using wireless open-access research platform version 3 (WARP V3) designed by Rice University. For digital signal processing, the Xilinx Virtex6 LX240T FPGA is used, dual-channel and 2.4GHz/5GHz dual-band transceiver are supported by two MAX2829 RF chips, the maximum transmission power is 20dBm by using the dual-band power amplifier, the 12-bit low power analog/digital converter AD9963 is used to provide two ADC channels with sampling rates of 100 MSPS

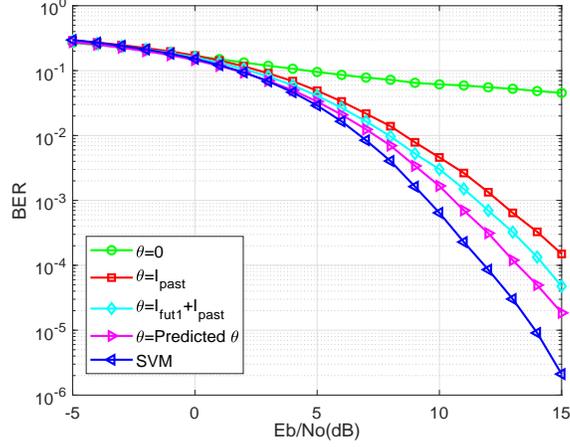


Figure 13: BER results with different decoding methods under 3 paths in a time-varying wireless channel. Blue solid line with left triangle markers is the BER curve using GA-SVM, magenta solid line with right triangle markers is the BER curve using the decoding threshold predicted by ESN, cyan solid line with diamond markers is the BER curve using the decoding threshold $\theta = I_{fut1} + I_{past}$, red solid line with square markers is the BER curve using the decoding threshold $\theta = I_{past}$, green solid line with circle markers is BER curve using zero threshold.

and two DAC channels with sample rates of 170 MSPS, and two 10/100/1000 Ethernet interfaces (Marvell 88E1121R) are used to realize the high-speed signal exchange with the Personal Computer (PC). The photo of the WARP system is shown in figure 14.

Each WARP node has two radio antennas, namely RFA and RFB. In our test, only the RFA is used, which is operating at the 5GHz carrier frequency with 20MHz bandwidth.

The BER performance of the GA-SVM based symbol decoding method under the practical channel case is tested. The test is performed at the university campus, and the photo of the test scenario is shown in figure 15, the distance between TX and RX is about 30 meters, there are many trees between them.

As mentioned before, the information is transmitted frame by frame, there are 2048 bits in one frame, which contains 896 probe data for GA-SVM model



Figure 14: The WARP system photo.



Figure 15: The test scenario at the campus.

training, also used for clock synchronization, 1152 bits are information data.

Four different methods are used in the scenario for comparison, the experimental BERs versus transmission power are shown in figure 16, we can see that the method based on GA-SVM get much better performance than the other comparison methods, which is consistent with the simulation results. The channel estimation in practice is so complicated that it is hard to obtain accurate channel state information, so the error contributes to the threshold calculation in these methods using channel information. In the GA-SVM learning process, the effects of the channel identification error and prediction error are excluded, that is one reason why the performance of GA-SVM is the best among the comparison methods. The most important reason to achieve the best performance is that GA-SVM with the minimum structure risk maps the waveform pattern of different symbols and the channel information into the trained GA-SVM struc-

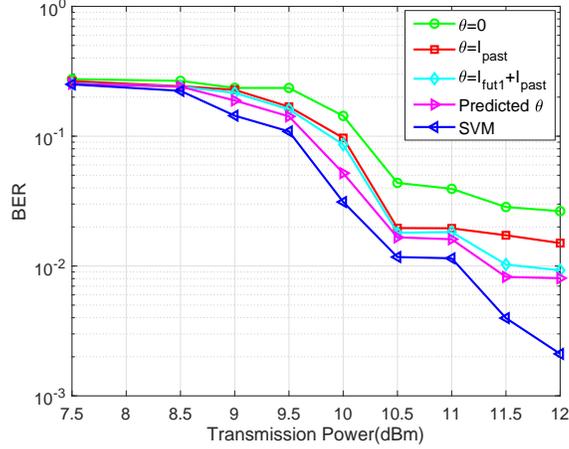


Figure 16: The experimental BER curves using different decoding methods. Blue solid line with left triangle markers is the BER curve using GA-SVM, magenta solid line with right triangle markers is the BER curve using the decoding threshold predicted by ESN, cyan solid line with diamond markers is the BER curve using the decoding threshold $\theta = I_{fut1} + I_{past}$, red solid line with square markers is the BER curve using the decoding threshold $\theta = I_{past}$, green solid line with circle markers is BER curve using zero threshold.

ture to the best possible extent, thanks to the power of machine learning.

The point to be noticed here is that the BER curves in figure 16 seem “abnormal”, because, with the transmission power increasing, the BER does not decrease monotonically as intuitively expected. In fact, the underlying assumption of this expectation is that the signal to noise ratio (SNR) is increasing with the increasing of the transmission power, however, practically, it is not such situation, we control the transmission power exactly in experiment, but we cannot guarantee the noise level be the same at different transmission moments. Therefore, in the practical environment, when we increase the transmission power with not very large amount, the SNR might not increase, even in some extreme cases, the SNR is decreased because of the noise level is increased larger relatively. Therefore, we observe the “abnormal”, but very normal BER curves as given in figure 16. For the same reason, in simulation, we can control the noise level very well (exactly), even we can repeat “noise”. But in practical experiment,

we cannot calculate the SNR or Eb/No as done in the simulation, therefore, we cannot compare the experimental results with the simulation results in the same term. However, this obstacle does not affect the conclusion that we draw from the curves, i.e., under the same SNR (at the same transmission moment), our proposed method is superior to the comparison methods, although the BER for large transmission power might be not lower than that for low transmission power because the SNR might not be increased in this case.

Discussion: The real-time implementation of GA-SVM could be accomplished by employing the field programmable gate array (FPGA) resource in the experimental WARP set up. There are also many references [51, 52, 53, 54] dealing with the problem, which could be referred. Because the main purpose of this paper is to validate the effectiveness and the superiority of the proposed algorithm as compared to other methods. Hardware implementation is not discussed here any more.

5. Conclusion

A novel idea to treat the symbol decoding task in the CBWCS as a pattern classification problem is proposed in this paper. Under the guidance of this idea, GA-SVM is trained using the probe data in the transmission data frame, and then used as a symbol decoder to decode the information. This approach avoids the channel information identification or future symbol prediction as well as the BER increasing due to these processes, by this way, the proposed method simplified the procedure and improved the BER performance. The simulation and experimental results prove the feasibility and superiority of the proposed method. Although the implementation complexity of GA-SVM is considerable, it can be solved using FPGA hardware implementation.

References

References

- [1] B. Sklar, Digital communications: fundamentals and applications, Publishing House of Electronics Industry, 2006. doi:10.1038/hr.2009.214.
- [2] J. L. Yao, C. Li, H. P. Ren, C. Grebogi, Chaos-based wireless communication resisting multipath effects, Phys. Rev. E 96 (3) (2017) 032226. doi:10.1103/PhysRevE.96.032226.
- [3] H. P. Ren, H. P. Yin, C. Bai, J. L. Yao, Performance improvement of chaotic baseband wireless communication using echo state network, IEEE Trans. Commun. 68 (10) (2020) 6525–6536. doi:10.1109/TCOMM.2020.3007757.
- [4] H. P. Yin, C. Bai, H. P. Ren, Echo state network based symbol detection in chaotic baseband wireless communication, Submitted to Digit. Commun. Netw.
- [5] J. L. Yao, Y. Z. Sun, H. P. Ren, C. Grebogi, Experimental wireless communication using chaotic baseband waveform, IEEE Trans. Veh. Technol. 68 (1) (2019) 578–591. doi:10.1142/9789811201523_0007.
- [6] C. Williams, Chaotic communications over radio channels, IEEE Trans. Circuits Syst. I: Fundamental Theory and Applications 48 (12) (2001) 1394–1404. doi:10.1109/TCSI.2001.972846.
- [7] E. M. Bollt, Review of chaos communication by feedback control of symbolic dynamics, Int. J. Bifurcation Chaos 13 (02) (2003) 269–285. doi:10.1142/s0218127403006546.
- [8] P. P. Chen, Y. Fang, K. X. Su, G. R. Chen, Design of a capacity-approaching chaos-based multiaccess transmission system, IEEE Trans. Veh. Technol. 66 (12) (2017) 10806–10816. doi:10.1109/TVT.2017.2723608.

- [9] C. Bai, H. P. Ren, G. Kolumbán, Double-sub-stream m-ary differential chaos shift keying wireless communication system using chaotic shape-forming filter, *IEEE Trans. Circuits Syst. I-Regul. Pap.* 67 (10) (2020) 3574–3587. doi:10.1109/TCSI.2020.2993674.
- [10] S. Hayes, C. Grebogi, E. Ott, Communicating with chaos, *Phys. Rev. Lett.* 70 (20) (1993) 3031–3034. doi:10.1103/physrevlett.70.3031.
- [11] E. Rosa Jr, S. Hayes, C. Grebogi, Noise filtering in communication with chaos, *Phys. Rev. Lett.* 78 (7) (1997) 1247–1250. doi:10.1103/PhysRevLett.78.1247.
- [12] N. J. Corron, D. W. Hahs, A new approach to communications using chaotic signals, *IEEE Trans. Circuits Syst. I* 44 (5) (1997) 373–382. doi:10.1109/81.572333.
- [13] A. Argyris, D. Syvridis, L. Larger, V. Annovazzi Lodi, P. Colet, I. Fischer, J. Garcia-Ojalvo, C. R. Mirasso, L. Pesquera, K. A. Shore, Chaos-based communications at high bit rates using commercial fibre-optic links, *Nature* 438 (7066) (2005) 343–346. doi:10.1038/nature04275.
- [14] H. P. Ren, M. S. Baptista, C. Grebogi, Wireless communication with chaos, *Phys. Rev. Lett.* 110 (18) (2013) 184101. doi:10.1103/PhysRevLett.110.184101.
- [15] N. J. Corron, J. N. Blakely, M. T. Stahl, A matched filter for chaos, *Chaos* 20 (2) (2010) 023123. doi:10.1063/1.3432557.
- [16] N. J. Corron, J. N. Blakely, Chaos in optimal communication waveforms, *Proc. R. Soc. A-Math. Phys. Eng. Sci.* 471 (2180) (2015) 20150222. doi:10.1098/rspa.2015.0222.
- [17] M. J. Er, Y. Zhou, *Theory and Novel Applications of Machine Learning*, BoD–Books on Demand, 2009. doi:10.5772/56681.

- [18] J. Xu, L. Q. Song, J. Y. Xu, G. J. Pottie, M. Van der Schaar, Personalized active learning for activity classification using wireless wearable sensors, *IEEE J. Sel. Top. Signal Process.* 10 (5) (2016) 865–876. doi:10.1109/JSTSP.2016.2553648.
- [19] X. Chen, F. Kopsaftopoulos, Q. Wu, H. Ren, F. K. Chang, Flight state identification of a self-sensing wing via an improved feature selection method and machine learning approaches, *Sensors* 18 (5) (2018) 1379. doi:10.3390/s18051379.
- [20] S. Hosseinyalamdary, Deep Kalman filter: Simultaneous multi-sensor integration and modelling; A GNSS/IMU case study, *Sensors* 18 (5) (2018) 1316. doi:10.3390/s18051316.
- [21] W. Woo, B. Gao, A. Bouridane, B. Ling, C. Chin, Unsupervised learning for monaural source separation using maximization–minimization algorithm with time–frequency deconvolution, *Sensors* 18 (5) (2018) 1371. doi:10.3390/s18051371.
- [22] G. C. Yang, J. Yang, W. H. Sheng, F. Junior, S. Li, Convolutional neural network-based embarrassing situation detection under camera for social robot in smart homes, *Sensors* 18 (5) (2018) 1530. doi:10.3390/s18051530.
- [23] M. Coutinho, R. de Oliveira Albuquerque, F. Borges, L. García Villalba, T. H. Kim, Learning perfectly secure cryptography to protect communications with adversarial neural cryptography, *Sensors* 18 (5) (2018) 1306. doi:10.3390/s18051306.
- [24] C. X. Jiang, H. J. Zhang, Y. Ren, Z. Han, K. C. Chen, L. Hanzo, Machine learning paradigms for next-generation wireless networks, *IEEE Wirel. Commun.* 24 (2) (2016) 98–105. doi:10.1109/MWC.2016.1500356WC.
- [25] T. T. Nguyen, G. J. Armitage, A survey of techniques for internet traffic

- classification using machine learning, *IEEE Commun. Surv. Tutor.* 10 (1-4) (2008) 56–76. doi:10.1109/SURV.2008.080406.
- [26] S. Z. Bi, R. Zhang, Z. Ding, S. G. Cui, Wireless communications in the era of big data, *IEEE Commun. Mag.* 53 (10) (2015) 190–199. doi:10.1109/MCOM.2015.7295483.
- [27] N. Guenther, M. Schonlau, Support vector machines, *Stata J.* 16 (4) (2016) 917–937. doi:10.1177/1536867X1601600407.
- [28] B. Ghaddar, J. Naoum Sawaya, High dimensional data classification and feature selection using support vector machines, *Eur. J. Oper. Res.* 265 (3) (2018) 993–1004. doi:10.1016/j.ejor.2017.08.040.
- [29] Y. Q. Ma, G. D. Guo, Support vector machines applications, Springer International Publishing, 2014. doi:10.1007/978-3-319-02300-7.
- [30] R. C. Daniels, R. W. Heath, An online learning framework for link adaptation in wireless networks, in: *Proc. 2009 Inf. Theory Appl. Workshop, ITA, 2009*, pp. 138–140. doi:10.1109/ITA.2009.5044935.
- [31] S. Yun, C. Caramanis, Multiclass support vector machines for adaptation in MIMO-OFDM wireless systems, in: *Proc. 2009 47th Annu. Allerton Conf. Commun., Control, Comput., Allerton, 2009*, pp. 1145–1152. doi:10.1109/ALLERTON.2009.5394863.
- [32] A. Rico Alvarino, R. W. Heath, Learning-based adaptive transmission for limited feedback multiuser MIMO-OFDM, *IEEE Trans. Wirel. Commun.* 13 (7) (2014) 3806–3820. doi:10.1109/TWC.2014.2314104.
- [33] C. Feng, S. Z. Liao, Scalable Gaussian kernel support vector machines with sublinear training time complexity, *Inf. Sci.* 418 (2017) 480–494. doi:10.1016/j.ins.2017.08.033.
- [34] Y. L. Yuan, M. Zhang, P. F. Luo, Z. Ghassemlooy, L. Lang, D. S. Wang, B. Zhang, D. H. Han, SVM-based detection in visible light communications, *Optik* 151 (2017) 55–64. doi:10.1016/j.ijleo.2017.08.089.

- [35] A. Ben Hur, J. Weston, A users guide to support vector machines, in: Data mining techniques for the life sciences, Springer, 2010, pp. 223–239. doi:10.1007/978-1-60327-241-4_13.
- [36] C. C. Chang, C. J. Lin, LIBSVM: A library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (3) (2011) 27. doi:10.1145/1961189.1961199.
- [37] H. P. Ren, C. Bai, C. Grebogi, Chaotic shape-forming filter and corresponding matched filter in wireless communication, Chapter in Advances on Nonlinear Dynamics of Electronic Systems, World Scientific Press, 2019. doi:10.1142/9789811201523_0007.
- [38] H. P. Ren, H. E. Zhao, C. Bai, H. P. Yin, C. Grebogi, Artificial intelligence enhances the performance of chaotic baseband wireless communication, arXiv preprint arXiv:1907.01521.
- [39] C. M. Bishop, Pattern recognition and machine learning, Springer-Verlag New York, Inc., 2006. doi:10.1117/1.2819119.
- [40] L. X. Wang, Y. J. Ren, R. H. Zhang, Algorithm of digital modulation recognition based on support vector machines, in: Proc. 2009 Int. Conf. Mach. Learning Cybernetics, Vol. 2, 2009, pp. 980–983. doi:10.1109/ICMLC.2009.5212366.
- [41] C. Z. Wang, Q. H. Hu, X. Z. Wang, D. G. Chen, Y. H. Qian, Feature selection based on neighborhood discrimination index, IEEE Trans. Neural Netw. Learn. Syst. 29 (7) (2018) 2986–2999. doi:10.1109/TNNLS.2017.2710422.
- [42] L. Li, C. Y. Wang, W. Li, J. B. Chen, Hyperspectral image classification by AdaBoost weighted composite kernel extreme learning machines, Neurocomputing 275 (2018) 1725–1733. doi:10.1016/j.neucom.2017.09.004.

- [43] S. Chun, I. Hwang, W. Son, J. H. Chang, W. Park, Recognition, classification, and prediction of the tactile sense, *Nanoscale* 10 (22) (2018) 10545–10553. doi:10.1039/C8NR00595H.
- [44] D. S. Bulgarevich, S. Tsukamoto, T. Kasuya, M. Demura, M. Watanabe, Pattern recognition with machine learning on optical microscopy images of typical metallurgical microstructures, *Sci Rep* 8 (1) (2018) 1–8. doi:10.1038/s41598-018-20438-6.
- [45] J. H. Ruan, H. Jiang, X. Y. Li, Y. Shi, F. T. Chan, W. Rao, A granular GA-SVM predictor for big data in agricultural cyber-physical systems, *IEEE Trans. Ind. Inform.* 15 (12) (2019) 6510–6521. doi:10.1109/TII.2019.2914158.
- [46] Y. Q. Guo, X. P. Jia, D. Paull, Effective sequential classifier training for SVM-based multitemporal remote sensing image classification, *IEEE Trans. Image Process.* 27 (6) (2018) 3036–3048. doi:10.1109/TIP.2018.2808767.
- [47] J. Meinilä, P. Kyösti, T. Jämsä, L. Hentilä, Winner II channel models, *Radio Technologies and Concepts for IMT-Advanced* (2009) 39–92. doi:10.1002/9780470748077.ch3.
- [48] T. Hofmann, B. Schölkopf, A. J. Smola, Kernel methods in machine learning, *Annals of Statistics* 36 (3) (2008) 1171–1220. doi:10.1214/009053607000000677.
- [49] C. L. Huang, C. J. Wang, A GA-based feature selection and parameters optimization for support vector machines, *Expert Syst. Appl.* 31 (2) (2006) 231–240. doi:10.1016/j.eswa.2005.09.024.
- [50] N. Cristianini, J. Shawe Taylor, et al., *An introduction to support vector machines and other kernel-based learning methods*, Cambridge Univ. Press, 2000. doi:10.1609/aimag.v22i2.1566.

- [51] D. Anguita, A. Boni, S. Ridella, A digital architecture for support vector machines: theory, algorithm, and FPGA implementation, *IEEE Trans. Neural Netw.* 14 (5) (2003) 993–1009. doi:10.1109/TNN.2003.816033.
- [52] X. P. Pan, H. H. Yang, L. Q. Li, Z. B. Liu, L. Hou, FPGA implementation of SVM decision function based on hardware-friendly kernel, in: *Proc. 2013 Int. Conf. Comput. Inf. Sci., ICCIS, 2013*, pp. 133–136. doi:10.1109/ICCIS.2013.43.
- [53] S. Kumar, J. Manikandan, V. Agrawal, Hardware implementation of support vector machine classifier using reconfigurable architecture, in: *Proc. 2017 Int. Conf. Adv. Comput., Commun. Informatics, ICACCI, 2017*, pp. 45–50. doi:10.1109/ICACCI.2017.8125814.
- [54] F. F. Lopes, J. C. Ferreira, M. A. Fernandes, Parallel implementation on FPGA of support vector machines using stochastic gradient descent, *Electronics* 8 (6) (2019) 631. doi:10.3390/electronics8060631.