

A Novel Blind Multiuser Detection Algorithm Based on Support Vector Machines

Yan Zhu, Limin Xiao, Shidong Zhou and Jing Wang

Department of Electronic Engineering

Tsinghua University

Beijing, China 100084

Telephone: (8610) 6278-1398

Fax: (8610) 6277-0317

Email: mite98@mails.tsinghua.edu.cn

Abstract—¹ In this paper, we propose a novel method based on support vector machine (SVM) for blind multiuser detection. Compared with existing methods, this algorithm has the feature of shorter iterative length and higher robustness. The simulation results show that in high SNR situation, our method has better BER performance than the normal minimum mean square error(MMSE) algorithm such as Kalman [1].

I. INTRODUCTION

In 1984 Verdú addressed the multiuser detection(MUD) [2] in order to cancel the near-far effect. His works ignites lots of research in this area. Honig proposed the blind MUD based on minima output energy(MOE) criterion [3], which is equivalent to minimum mean square error(MMSE) when the desired user's power is known. In following years, blind MUD has become a worldwide research topic and a number of methods for blind MUD has been proposed, such as least-mean-square(LMS) algorithm [3], recursive least-square(RLS) algorithm [4], [5] and Kalman algorithm [1]. These algorithms have already shown good performance and research about blind MUD algorithm is still going on in recent years.

Support vector machine(SVM) was first introduced in statistical learning theory [6]. It is based on the structural risk minimization(SRM), therefore, it has outstanding performance and robustness when there are only a few samples. Due to these traits, SVM has triggered many research interests. Now it has been widely used in image signal processing, for examples, intelligent classifications and pattern recognition. Since SVM has high complexity, plenty of recent research work has been done to find fast algorithm to offer computational reduction in SVM, such as [7]–[9]. We think future research will make SVM more practical in various areas.

Recently, SVM has been introduced to signal processing in communications, such as equalization [10] and nonlinear channel estimation [11]. In both [10] and [11], they view

equalization of BPSK as a processing that classifies the distorted signals into two classes($\{-1, 1\}$). Simulation results demonstrate SVM's outstanding performance in both linear and nonlinear channels.

There is also some research on support vector machines for MUD. The simulation results of [12] show that the SVM-based MUD can closely match the performance of the optimal Bayesian one-shot detector. Zhang has discussed his SVM-based multiuser detector in [13]. His algorithm can achieve better performance than MMSE detector, while its training samples are less than the MMSE counterpart. In our paper, we propose a blind MUD algorithm based on SVM, which do not need training samples. Our detector can work well when the amount of received data is small. Furthermore, the performance is excellent especially in high SNR.

Here is the outline of the rest parts: In Section II a brief review of the system model and blind MUD will be presented. After that, blind MUD based on SVM will be derived in Section III and some convergency issues are also discussed in this section. Simulation results are shown in the Section IV. Then, in Section V we get the conclusion.

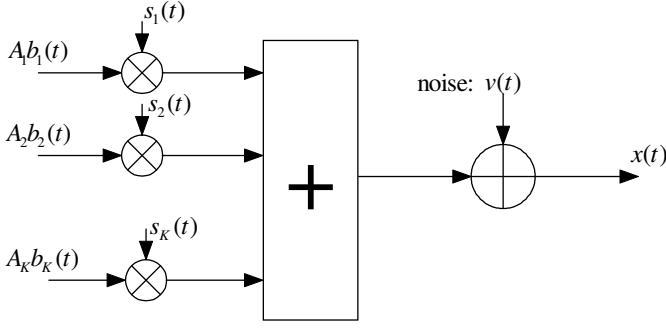
II. SYSTEM MODEL AND BLIND MUD

Consider an antipodal K -user synchronous DS-CDMA system signaling through an additive white Gaussian noise(AWGN) channel. By passing through a chip-matched filter, followed by a chip-rate sampler, the discrete-time output of receiver during one symbol interval can be modeled as (See Fig. II

$$x(n) = \sum_{k=1}^K A_k b_k s_k(n) + v(n) \quad n = 0, 1, \dots, N-1 \quad (1)$$

where

¹This project is sponsored by National 863 Science Foundation, No. 2001AA123012 and China Future Project.



- $v(n)$ ambient channel noise and σ is its variance
- K number of users
- A_k received amplitude of the k th user
- $b_k \in \{-1, 1\}$ is the bit transmitted by k th user
- s_k signature waveform of k th user.
- N processing gain

It is assumed that $s_k(n)$ is supported only on interval $[0, \dots, N-1]$ and unit energy, i.e. $\sum_{n=0}^{N-1} s_k(n) = 1$. Let T_c denote the chip interval, thus, symbol interval $T_s = NT_c$.

Defining

$$\mathbf{x} \triangleq [x(0), x(1), \dots, x(N-1)]^T \quad (2)$$

$$\mathbf{s}_k \triangleq [s_k(0), s_k(1), \dots, s_k(N-1)]^T \quad (3)$$

$$\mathbf{v} \triangleq [v(0), v(1), \dots, v(N-1)]^T \quad (4)$$

we may rewrite (1) as vector form

$$\mathbf{x} = A_1 b_1 \mathbf{s}_1 + \sum_{k=2}^K A_k b_k \mathbf{s}_k + \sigma \mathbf{v} \quad (5)$$

For convenience, in this paper we always assume user-1 is the desired one. Thus, blind multiuser detector is to extract b_1 from \mathbf{x} only using \mathbf{s}_1 . The criterion MOE is commonly used by blind multiuser detector [3]. The detector aims at minimizing:

$$\text{MOE}(\mathbf{w}) \triangleq E \{ (\langle \mathbf{w}, \mathbf{x} \rangle)^2 \} = E \{ (\mathbf{w}^T \mathbf{x})^2 \} \quad (6)$$

subject to

$$\mathbf{w}^T \mathbf{s}_1 = 1 \quad (7)$$

where \mathbf{w} is the detector coefficients for the first user. For Kalman algorithm, we decompose \mathbf{w} as $\mathbf{w} = \mathbf{s}_1 - \mathbf{C}_{1, null} \mathbf{c}_1$, where \mathbf{c}_1 is the adaptive part of \mathbf{w} and the N -by- $N-1$ matrix $\mathbf{C}_{1, null}$ span the null space of \mathbf{s}_1 . For more details, please see [1], [3], [4].

Although Kalman method has good performance, it needs large number of received data and the beginning of these data will be 'wasted'. Thus it is not suitable when number of received data is not large or the data frame has a small size.

Our method based on SVM can overcome such shortcoming.

III. METHOD BASED ON SVM

In this section, we will first derive our SVM-based blind MUD algorithm. Since the convergency issues are very important, we will discuss it in a separated subsection.

A. SVM Multiuser Detector

In DS/CDMA system, BPSK is in common use. BPSK is a constant-modulus modulation ($\{-1, +1\}$). Suppose \mathbf{w} are the multiuser detector's coefficients and the input signal is \mathbf{x} , then in BPSK we have $(\mathbf{w}^T \mathbf{x})^2 = 1$. We assume there are l input CDMA symbols $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l$ and the signature wave of the desired user is \mathbf{s}_1 .

According to the structural risk minimization (SRM) principle [6], to estimate the multiuser detector \mathbf{w} , one has to minimize

$$J(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^l \zeta_i (|1 - (\mathbf{w}^T \mathbf{x}_i)|) \quad (8)$$

where constant $C \geq 0$ and $\zeta_i(\cdot)$ is the loss function. Directly minimizing (8) is difficult, because the term $(\mathbf{w}^T \mathbf{x}_i)^2$ makes the problem into a nonlinear one. Therefore, we use the tricky in [14] to replace $(\mathbf{w}^T \mathbf{x}_i)^2$ with $y_i \mathbf{w}^T \mathbf{x}_i$, where y_i is the output of multiuser detector. Thus we can rewrite the whole question as following:

$$\min_{\mathbf{w}, \eta, \eta^*} R(\mathbf{w}, \eta, \eta^*) = \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{i=1}^l (\zeta_i(\eta_i) + \zeta_i^*(\eta_i^*)) \quad (9)$$

subject to

$$y_i f(\mathbf{x}_i) - 1 \leq \eta_i + \epsilon_i \quad (10)$$

$$1 - y_i f(\mathbf{x}_i) \leq \eta_i^* + \epsilon_i^* \quad \forall i \in [1 \dots l] \quad (11)$$

$$\eta_i, \eta_i^* \geq 0 \quad (12)$$

$$\mathbf{w}^T \mathbf{s}_1 = 1 \quad (12)$$

where

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} \quad (13)$$

Using Lagrange multiplier, we have

$$\begin{aligned} L(\mathbf{w}, \alpha, \alpha^*, \eta, \eta^*, \gamma, \gamma^*) &= \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{i=1}^l (\zeta_i(\eta_i) + \zeta_i^*(\eta_i^*)) \\ &\quad - \beta (\mathbf{w}^T \mathbf{s}_1 - 1) - \sum_{i=1}^l \alpha_i (\eta_i + \epsilon_i + 1 - y_i f(\mathbf{x}_i)) \\ &\quad - \sum_{i=1}^l \alpha_i^* (\eta_i^* + \epsilon_i^* - 1 + y_i f(\mathbf{x}_i)) - \sum_{i=1}^l \eta_i \gamma_i + \eta_i^* \gamma_i^* \end{aligned} \quad (14)$$

Let $\frac{\partial L}{\partial \mathbf{w}}$, $\frac{\partial L}{\partial \eta_i}$ and $\frac{\partial L}{\partial \eta_i^*}$ equal to 0 respectively and notice the constraint (12), we can solve β out. Thus we get the expression of \mathbf{w} without β . Finally we must maximize:

$$W = -\frac{1}{2}\|\mathbf{w}^*\|^2 - \frac{1}{2}(1 - \mathbf{w}^{*\text{T}}\mathbf{s}_1)^2 - \sum_{i=1}^l(\alpha_i\epsilon_i + \alpha_i^*\epsilon_i^*) + \sum_{i=1}^l(\alpha_i^* - \alpha_i) + C \sum_{i=1}^l \left(\zeta_i(\eta_i) - \eta_i \frac{d\zeta_i(\eta_i)}{d\eta_i} + \zeta_i^*(\eta_i^*) - \eta_i^* \frac{d\zeta_i^*(\eta_i^*)}{d\eta_i^*} \right) \quad (15)$$

where

$$\mathbf{w} = \mathbf{w}^* + (1 - \mathbf{w}^{*\text{T}}\mathbf{s}_1)\mathbf{s}_1 \quad (16)$$

and

$$\mathbf{w}^* = \sum_{i=1}^l(\alpha_i^* - \alpha_i)y_i\mathbf{x}_i \quad (17)$$

If we let $\zeta_i(x) = \zeta_i^*(x) = x$ and $\epsilon_i = \epsilon_i^* = \epsilon$, (15) becomes an ϵ -insensitive SVR(support vector regression). If we let $\zeta_i(x) = \zeta_i^*(x) = \frac{1}{2}x^2$ and $\epsilon_i = \epsilon_i^* = 0$, we get the quadratic SVR. In the following parts we only discuss the first one. Therefore we can simplify W :

$$W(\alpha, \alpha^*) = -\frac{1}{2}\|\mathbf{w}^*\|^2 - \frac{1}{2}(1 - \mathbf{w}^{*\text{T}}\mathbf{s}_1)^2 + \sum_{i=1}^l(\alpha_i^* - \alpha) - \epsilon \sum_{i=1}^l(\alpha_i^* + \alpha_i) \quad (18)$$

Substitute (17) into (18) and get the form for computation.

$$\begin{aligned} \max_{\alpha, \alpha^*} W(\alpha, \alpha^*) &= \sum_{i=1}^l(\alpha_i^* - \alpha_i)y_i\mathbf{x}_i\mathbf{s}_1 + \sum_{i=1}^l(\alpha_i^* - \alpha) \\ &- \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)y_i y_j (\mathbf{x}_i^{\text{T}}\mathbf{x}_j + \mathbf{x}_i^{\text{T}}\mathbf{s}_1\mathbf{x}_j^{\text{T}}\mathbf{s}_1) \\ &- \epsilon \sum_{i=1}^l(\alpha_i^* + \alpha_i) - \frac{1}{2} \end{aligned} \quad (19)$$

subject to

$$\alpha_i, \alpha_i^* \in [0, C] \forall i = 1 \dots l \quad (20)$$

(19) is a quadratic programming (QP) problem, which is widely researched in the theory of optimization, such as [15], [16].

Let \mathbf{w}_0 denote the initial value of \mathbf{w} , then we can use \mathbf{w}_0 and (13) to get y_i . After that, we have the "training" samples (\mathbf{x}_i, y_i) . Using these samples and (19)(20), α and α^* can be calculated. Thus we get a new \mathbf{w} from (16). Repeat these operations until it converges. Our algorithm is a block-in-block-out one, i.e. we must let l CDMA symbols into our multiuser detector and after processing, we output a

block of symbols.

Procedure above will probably occur a limit cycle oscillation, i.e. the outputs of multiuser detector are oscillating from iteration to iteration. It is caused by the discontinuity of multiuser detector coefficients. Therefore, we introduce a paramant λ , which satisfies $\lambda > 0$ and $1 - \lambda \ll 1$. Thus, from k -th iteration to $k + 1$ -th iteration, we use following formula to smooth the coefficients \mathbf{w} :

$$\mathbf{w}_{k+1} = \lambda\mathbf{w}_k + (1 - \lambda)\mathbf{w}_{QP} \quad (21)$$

where \mathbf{w}_{QP} are the coefficients of multiuser detector which are solved from the QP problem (19).

B. Convergency Issues

In this section, we consider the convergency of the iterative processing. With the iteration going on, $|y_i|$ will approach 1. So we define [14]:

$$AME(k) \triangleq \frac{1}{l} \sum_{i=1}^l (|y_{i,k}|^2 - 1) \quad (22)$$

where $y_{i,k}$ denotes the output of the multiuser detector for the i -th input CDMA symbols at the k -th iteration. The indicator of convergency can be defined as:

$$AME(k+1) - AME(k) < 5 \times 10^{-3} \quad (23)$$

On the left end of (23), we do not use the absolute value of $AME(k+1) - AME(k)$. The reason is that there is a certain probability for the SVM over-learning. That is to say, $AME(k+1)$ may be larger than $AME(k)$ after a certain iteration.

If this over-learning occurs in the first iteration, it will considerably degrade the performance. To avoid this situation, the initial value of \mathbf{w} should be carefully chosen. A straightforward idea is to let \mathbf{w}_0 be \mathbf{s}_1 , but it does not work well in our simulation. We use following method to find the initial value for \mathbf{w} . Let $\mathbf{c} = [c_1, c_2, \dots, c_l]^{\text{T}}$, where $c_i = (\alpha_i^* - \alpha_i)y_i$. And we also denote $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l]$ and $\mathbf{y} = [y_1, y_2, \dots, y_l]$. Notice $\mathbf{w}^* = \mathbf{X}\mathbf{c}$ and (17), thus, to initialize \mathbf{w} is equivalent to initialize \mathbf{c} . There is an obvious fact:

$$\mathbf{w}^{\text{T}}\mathbf{X} = \mathbf{y} \quad (24)$$

we assign \mathbf{y} in (24) to $\text{sgn}(\mathbf{s}_1^{\text{T}}\mathbf{X})(\triangleq \mathbf{y}_0)$ and use \mathbf{c} to represent \mathbf{w} in (24):

$$\mathbf{c}^{\text{T}}(\mathbf{X}^{\text{T}}\mathbf{X} - \mathbf{X}^{\text{T}}\mathbf{s}_1\mathbf{s}_1^{\text{T}}\mathbf{X}) + \mathbf{s}_1^{\text{T}}\mathbf{X} = \mathbf{y}_0 \quad (25)$$

Solving this linear system, we can get \mathbf{c} , therefore the initial value for \mathbf{w} that we want.²

We summarize the algorithm as following

- 1) Get l CDMA symbols $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l$;

²For practical computation, an regularizing term must be added to make the matrix $\mathbf{X}^{\text{T}}\mathbf{X} - \mathbf{X}^{\text{T}}\mathbf{s}_1\mathbf{s}_1^{\text{T}}\mathbf{X}$ nonsingular.

- 2) Calculate the initial value of \mathbf{w} by (25), therefore, we get the initial value of \mathbf{y} ;
- 3) Using $(\mathbf{x}_i, y_i) \quad i = 1 \dots l$ and (19) to compute \mathbf{w}_{QP} ;
- 4) Get the \mathbf{w}_k using (21), therefore, \mathbf{y} ;
- 5) Calculate AME using (22);
- 6) If AME satisfies (23) then output \mathbf{y} , otherwise, repeat from 3).

IV. SIMULATION RESULTS AND DISCUSSIONS

In this section, we give the simulation results of our new algorithm. We consider altogether 8 users in the synchronous CDMA system with the processing gain N equaling to 31. User 1 is assumed as the desired one. We let $A_1 = 1$ and the SNRs of the other 7 users are all 30dB. The parameter λ in (21) is assigned to 0.9. The parameter ϵ has minor effect on the performance of our SVM detector, because our SVM is an ϵ -insensitive one [6]. In our simulation, we use the optimal one, given by [17]:

$$\epsilon = \tau \sqrt{\sigma_n^2 \frac{\ln l}{l}} \quad (26)$$

where σ_n^2 is the variance of noise, which can be estimated. l is the iteration length. τ is an empirically selected constant that we have chosen as $\tau = 3$.

A. BER v.s. Iterative Length

Kalman method is in common use and better than RLS [1]. We will compare it with our algorithm. Usually, the normal SVM methods have better performance in small samples size (i.e. a few CDMA symbols received) than other algorithm. Our blind multiuser detector also have this feature. In the first simulation, user 1's SNR is fixed at 10 dB. We compare the BER performance of our algorithm with Kalman detector under the same iterative length. As shown in fig. 1, the x -axis is the iterative length (for our algorithm, the iterative length is l) and y -axis is represented the BER performance. The curves of SVM method is far below the Kalman's. When the iterative length grows to 150, the SVM's BER performance is about 1.1×10^{-2} , while Kalman has not converged yet.

B. BER v.s. SNR

We simulate the BER performance of our method over various SNRs of user 1. Because Kalman method gets completely convergent at least after 500 points [1], we set Kalman's iterative length to be 2000 CDMA symbols, while SVM's iterative length is only 100. Fig 2 shows that Kalman is better in low SNR. The reason is that there is large probability for SVM being at over-learning status after the first iteration. Better initial value for iteration or longer iterative length will improve its performance in low SNR. However, even using 100 iterative length, we can see SVM perform much better than Kalman at high SNR. When SNR is greater than 9dB, our algorithm is better than Kalman's. At 15dB our method exceeds Kalman's observably. Its BER is

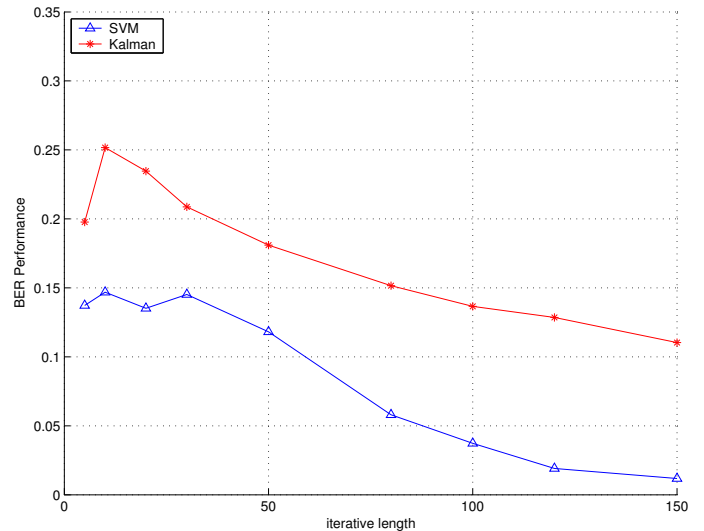


Fig. 1. iterative length vs. BER when the user 1's SNR is fixed

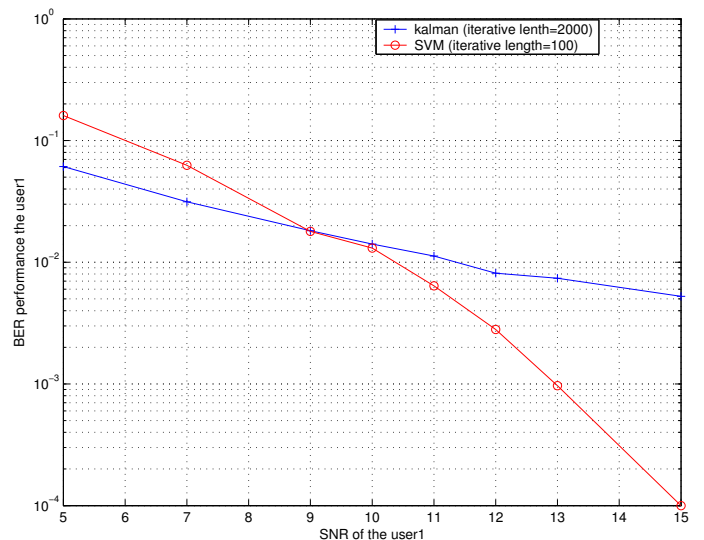


Fig. 2. user 1's SNR vs BER performance

about 1.0×10^{-4} , while Kalman in BER is around 5.3×10^{-3} .

V. CONCLUSION

In this paper, we have derived a new multiuser detector based on SVM. This detector can work well when the received data has small size. Comparisons with other method in BER performance have been done in this paper. The simulation results show that it is a promising method. In following research, we will work on the choice of initial value, the proof of convergency and how to reduce the complexity.

REFERENCES

- [1] X.-D. Zhang and W. Wei, "Blind adaptive multiuser detection based on kalman filtering," *IEEE Trans. Signal Processing*, vol. 50, pp. 87-95, Jan. 2002.

- [2] S. Verdú, "Optimum multiuser signal detection," Ph.D. dissertation, Univ. of Illinois at Urbana-Champaign, 1984.
- [3] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, 1995.
- [4] H. V. Poor and X. Wang, "Code-aided interference suppression for ds/cdma communications— part i: Interference suppression capability," *IEEE Trans. Commun.*, vol. 45, pp. 1101–1111, Sept. 1997.
- [5] ———, "Code-aided interference suppression for ds/cdma communications— part ii: Parallel blind adaptive implementations," *IEEE Trans. Commun.*, vol. 45, pp. 1112–1122, Sept. 1997.
- [6] V. Vapnik, *Statistical Learning Theory*. John Wiley & Sons, Inc, 1998.
- [7] G. Mak, "The implementation of support vector machines using the sequential minimal optimization," Master's thesis, School of Computer Science McGill University, Montreal, Canada, 1999.
- [8] A. A.-R. Angel Navia-Vázquez, Fernando Pérez-Cruz and A. R. Figueiras-Vidal, "Weighted least squares training of support vector classifiers leading to compact and adaptive schemes," *IEEE Trans. Neural Networks*, vol. 12, pp. 1047–1059, Sept. 2001.
- [9] A. N.-V. F. Pérez-Cruz, P. Alzrcón-Diana, and A. Artés-Rodríguez, "An irwls procedure for svr," in *Proc. of EUSIPCO'00*, 2000.
- [10] S. Chen, S. R. Gunn, and C. J. Harris, "The relevance vector machine technique for channel equalization application," *IEEE Trans. Neural Networks*, vol. 12, pp. 1529–1532, June 2001.
- [11] D. J. Sebald and J. A. Bucklew, "Support vector machine techniques for nonlinear equalization," *IEEE Trans. Signal Processing*, vol. 48, pp. 3217–3226, Nov. 2000.
- [12] S. Chen, A. Saminagan, and L. Hanzo, "Adaptive multiuser receiver using a support vector machine technique," presented at the Proc. of IEEE VTC2001 Spring, 2001.
- [13] X. Zhang and B. Zheng, *Communication Signal Processing*. China: National Defence Industry Press, 2002, (in Chinese).
- [14] I. Santamaría, C. Pantaleón, L. Viela, and J. I. nez, "Blind equalization of constant modulus signals using support vector machines," *IEEE Trans. Signal Processing*, to be published.
- [15] Y. Yuan and W. Sun, *Optimization Theory and Its Methods*. Beijing: Science Publisher Of China, 1998, (in Chinese).
- [16] S. R. Gunn. (1998) Matlab support vector machine toolbox. [Online]. Available: <http://www.isis.ecs.soton.ac.uk/isystems/kernel/>
- [17] V. Cherkassky and Y. Ma, "Selection of meta-parameters for support vector regression," in *Proc. of ICANN'2002*, Berlin, Jan. 2002, pp. 687–693.