Automatic Modulation Classification Based on Statistical Features and Support Vector Machine

Weifeng Zhang

Science and Technology on Communication Information Security Control Laboratory, Jiaxing Zhejiang, China
Email: zwf.zhang@gmail.com

Abstract

Automatic classification of the modulation format of a detected signal plays an important role in cognitive radio system and several military and civilian applications. Without previous knowledge of the received data, automatic modulation classification (AMC) becomes a difficult task. This paper discusses AMC of digital schemes which are commonly used in today's communication systems. Several statistical features are extracted to represent signals and Support Vector Machines (SVMs) are then applied to classify the unknown modulation schemes. Experiment results demonstrate that the proposed AMC algorithm is a practical algorithm and it has high robustness in a wide range of SNR.

1. Introduction

As an intermediate step between signal detection and demodulation, Automatic modulation classification (AMC) provides valuable insight into signal's structure, origin and properties. Furthermore, automatic modulation classification has become a key step in cognitive radio [1], in which AMC is used to make intelligent receivers can recognize the modulation type without having any prior information of the detected signals. Obviously, it is a difficult task to identify modulation type with little prior knowledge of the detected signals. In the past decades, researchers in the communications and signal processing community have proposed many AMC algorithms, which can be generally grouped into two categories, the likelihood-based (LB) methods [2, 3, 4], and the feature-based (FB) methods [5-8]. Within the LB framework, AMC is a multiple composite hypothesis-testing problem. It computes the likelihood function of the received signal and the decision is made by comparing the likelihood ratio against a threshold. In the FB approach, on the other hand, several features are usually extracted and the signals are represented by these features first. Then a decision is made based on their observed values. The validity of FB-AMC mainly depends on the signal features extracted and classifier designed.

In this paper we proposed an AMC algorithm which falls under the category of feature-based methods. Most of the FB-AMC algorithms to date do not work over a wide range of signal-to-noise (SNR). Especially when the SNR is low, the modulation types can not be recognized correctly. In our AMC approach, several statistical features are extracted after preprocessing the detected signal. Then statistical machine learning algorithm – support vector machine [9, 10], is adopted as the classifier to make decision based on these extracted features. Simulation results show that the proposed algorithm is effective to recognize modulation types including 2FSK, 4FSK, 2ASK, 4ASK, 2PSK,4PSK on a wide range of SNRs.

2. Signal Model and Feature Extraction

The received signal can be expressed as the following model:

$$\begin{aligned} s(t) &= s(t) + n(t) \\ &= a(t)\cos[\omega_C t + \varphi(t)] + n(t) \end{aligned} \tag{1}$$

where $s(t)$ represents the modulated signal, and $n(t)$ is zero-mean additive white Gaussian noise. In the proposed AMC algorithm, 6 statistical features, which are similar to the features used in literature [6], are extracted to describe modulated signals.

• Maximum value of the spectral power density of the normalized-centered instantaneous amplitude:

$$\gamma_{\max} = \frac{\max |FFT[a_{cn}(i)]|^2}{N_s} \tag{2}$$

where $a_{cn}(i)$ is the value of the normalized-centered instantaneous amplitude at time $t = i / f_s, (i = 1,2,...,N_s)$, defined by

978-1-4673-5225-3/14/$31.00 ©2014 IEEE

$$a_{cn}(i)=\frac{a(i)}{\frac{1}{N_s}\sum_{i=1}^{N_s}a(i)}-1 \tag{3}$$

• Standard deviation of the absolute value of the centered non-linear component of the instantaneous phase

$$\sigma_{ap}=\sqrt{\frac{1}{C}(\sum_{a_n(i)>a_t}\varphi_{NL}^2(i))-(\frac{1}{C}\sum_{a_n(i)>a_t}|\varphi_{NL}(i)|)^2} \tag{4}$$

where $\varphi_{NL}(i)$ is the value of the centered non-linear component of the instantaneous phase at time instants $t=i/f$, which can be calculated as follows:

$$\varphi_{NL}(i)=\varphi(i)-\frac{1}{N_s}\sum_{i=1}^{Ns}\varphi(i) \tag{5}$$

C means the number of samples in $\{\varphi_{NL}(i)\}$ for which $a_n(i)>a_t$ and a_t is a threshold for $\{a(i)\}$ below which the estimation of the instantaneous phase is very sensitive to the noise.

• Standard deviation of the centered non-linear component of the direct instantaneous phase

$$\sigma_{dp}=\sqrt{\frac{1}{C}(\sum_{a_n(i)>a_t}\varphi_{NL}^2(i))-(\frac{1}{C}\sum_{a_n(i)>a_t}\varphi_{NL}(i))^2} \tag{6}$$

• Standard deviation of the absolute value of the normalize-centred instantaneous amplitude

$$\sigma_{aa}=\sqrt{\frac{1}{N_S}(\sum_{i=1}^{N_s}a_{cn}^2(i))-(\frac{1}{N_S}\sum_{i=1}^{N_s}|a_{cn}(i)|)^2} \tag{7}$$

• Standard deviation of the normalized- centered instantaneous amplitude in the non-weak intervals

$$\sigma_{a}=\sqrt{\frac{1}{C}(\sum_{a_n(i)>a_t}a_{cn}^2(i))-(\frac{1}{C}\sum_{a_n(i)>a_t}a_{cn}(i))^2} \tag{8}$$

• The kurtosis of the normalized instantaneous amplitude

$$\mu_{42}^{a}=\frac{E\{a_{cn}^4(i)\}}{\{E[a_{cn}^2(i)\}^2} \tag{9}$$

All signal features are selected because of its low complexity and low dependency on SNRs.

3. Modulation Classification using SVM

Multi-class Support Vector Machine (SVM) [9, 10] is used as the classifier in our AMC algorithm. To understand multi-class SVMs, we first consider a two class SVM. Assume that the training feature-label pairs are:

$$(\mathbf{x}_i,l_i) \qquad i=1,...,N \tag{10}$$

where $\mathbf{x}_i$ is the training feature vector of the ith training sample belong class 1 or class 2, and associated labels $l_i=1$ be for class 1 and $l_i=-1$ be for class 2, respectively. SVMs require the solution of the following optimization problem:

$$\begin{gathered}\min_{w,b,\zeta}\frac{1}{2}\mathbf{w}^T\mathbf{w}+C\sum_{i=1}^{N}\zeta_i \\ subject\ \ to\ \ l_i(\mathbf{w}^T\varphi(\mathbf{x}_i)+b)\geq 1-\zeta_i,\quad \zeta_i\geq 0,\ \ i=1,...,N\end{gathered} \tag{11}$$

where $\mathbf{w}$ and b are hyper-plane parameters, ζ_i is positive slack variables, $C>0$ is the penalty parameter, and φ maps training vectors into a higher dimensional SVM feature space. The kernel function, which is defined as $\kappa(\mathbf{x}_i,\mathbf{x}_j)=\varphi(\mathbf{x}_i)^T\varphi(\mathbf{x}_j)$, represents the correlation of objects in the SVM feature space. Kernel function used in our proposed algorithm is the popular radial basis function kernel which is defined as follows: $\kappa(\mathbf{x}_i,\mathbf{x}_j)=\exp(-\gamma\|\mathbf{x}_i-\mathbf{x}_j\|^2),\quad \gamma>0$. The kernel parameter γ and penalty parameter C should be carefully chosen by grid-searching and cross-validation.

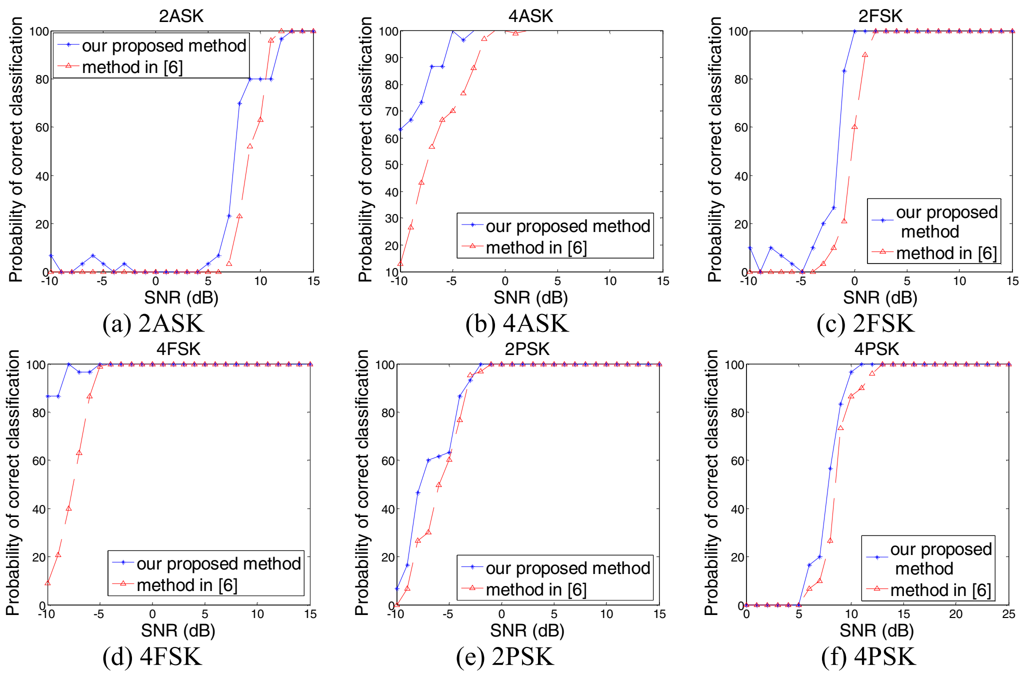
In this paper, multi-class SVMs are needed to classify multiple kinds of modulation types. For multi-class classification, there are two approaches. The first one is "one-versus-all" method, which constructs K classifiers for K classes and each classifier separates a single class between all other classes. The second one is called "one-versus-one" which provides more accurate results since the boundary region structure is often simpler to describe functionally and is used in this paper. For K classes problem, $C_K = K(K-1)/2$ binary classifiers are constructed, each separating one class from another ignoring all the other classes. And the C_K binary SVM classifiers are combined by majority voting. In this paper, signals are represented as a 6 dimensional vector by using the method descript in section II. And $C_K = 15$ binary SVM classifiers are build in total for six classes of modulation types including 2ASK, 4ASK, 2FSK, 4FSK, 2PSK and 4PSK common used in digital communication system.

4 Experiment Results

In this section experiments were done for verifying the validity of the proposed AMC algorithm. The proposed classifier was trained using software simulated signals that were generated in the MATLAB environment. The SNR value of simulated signals ranged from -10dB to 15dB. In 1dB increments, 100 segments, whose duration is 1 second, are generated for each modulation type. Thus, a total of 2600 signal segments were obtained to train the multi-class SVM descript in section III.

The classifier was trained over a wide range of SNRs (from -10dB to 15dB) simultaneously during the training stage. This training method make the classifier be able to handle signals of different SNRs without re-training unless the SNR is out of the classifier's training range. To optimize the performance of SVM classifiers, the penalty C and the kernel parameter γ must be tuned. This is done by a grid cross-validation search with the entire training data set.

To evaluate the performance of the proposed AMC algorithm, the trained classifier is used to test simulated testing signals. 50 segments of simulated signal segments are generated for each modulation type in 1dB increments whose SNR covered from -10dB to 15dB. The duration of these testing signal segments are also 1 second. The performance measurement used in this paper is the popular probability of correct classification $P_c^{(i|i)}$ which denotes the probability of classifier considers the signal's modulation type is i while the actual modulation type of incoming signal is i. To compare the performance of our proposed algorithm with the previous work's result, we also implement the method in literature [6].



(a) 2ASK (b) 4ASK (c) 2FSK

(d) 4FSK (e) 2PSK (f) 4PSK

Fig. 1 Performance of the proposed AMC algorithm

The detailed test results are reported in the above figures. These figures show the classification results at different SNRs. As expected, the probability of correct classification of each modulation type increased along with the SNR rising. And it is worth mentioning that the performance of our algorithm in this paper makes considerable progress than the method in literature [6], especially when the SNR is lower than 5dB. The classifier described in [6] is a decision tree classifier whose threshold at each node is fixed. Thus it has three unavoidable issues. The first is how to decide the best decision threshold values for the features. And the second one is it is very difficult to find an optimal decision threshold in a large range of SNRs and it has little probability to find some features that are independent on SNR. In contrast, the SVM can adjust the decision thresholds adaptively to the training data. The well-trained SVM can handle the classification problem in a wide range of SNR.

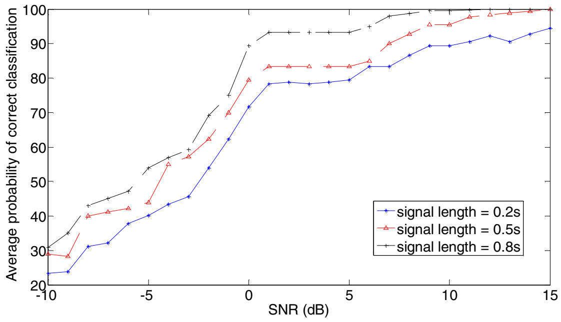


Fig. 2 Average probability of correct classification for different length of signals over a wide range of SNR

The dependence of the classification success rate on the data length was also studied. To evaluate the performance on shorter signals, three groups of testing signals were generated. Each group has 50 testing signal segments, whose duration is 0.2 second, 0.5 second and 0.8 second respectively. It is known that the use of a short signal length will in general degrade the classification performance. However, the experimental results, as seen in figure 2, showed that the AMC algorithm in this paper still get a acceptably high probability of correct classification even when the signal length is only 0.2 second. It achieved an average classification rate above 81% for a poor SNR condition (SNR=0dB).

5. Conclusion

In this paper, we proposed a new FB-AMC algorithm. Statistical features are extracted to represent signals and multi-class SVMs are used as the classifiers to identify modulation types. Experimental results demonstrate the AMC method in this paper is a practical approach, which is quite robust to signal length over a wide range of SNR.

6. References

1. J. Mitola et al., "Cognitive radio: Making software radios more personal", *IEEE Pres. Commun*., vol. 6, no. 4, Aug. 1999, pp. 13018.

2. W. Wei and J. M. Mendel, "Maximum-likelihood classification for digital amplitude-phase modulations," *IEEE Trans. Commun.*, vo,. 48, 2000, pp. 189-193.

3. C. Long, K. Chugg, and A. Polydoros, "Further results in likelihood classification of QAM signals," in Proc. IEEE MILCOM , 1994, pp. 57-61.

4. A. E. E1-Mahdy and N. M. Namazi, "Classification of multiple M-ary frequency-shift keying signals over a Rayleigh fading channel," *IEEE Trans. Commun*., vol. 50, 2002, pp. 967-974.

5. O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "The classification of joint analog and digital modulations," in Proc. IEEE MILCOM, 2005, pp. 1-6.

6. E. E. Azzouz and A. K. Nandi, Automatic Modulation Recognition of Communication Signals. Kluwer Academic, 1996.

7. S. S. Soliman and S. Z. Hsue, "Signal classification using statistical moments," *IEEE Trans. Commun*, vol. 40, 1992, pp. 908-916.

8. A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Trans. Commun*., vol. 48, 2000, pp. 416-429.

9. V. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, NY, 1995.

10. C. Cortes and V. Vapnik. *Support-vector network. Machine Learning*, 1995, pp. 273-297.