Tracking Performance of Semi-Supervised Large Margin Classifiers in Automatic Modulation Classification

Hamidreza Hosseinzadeh*
Department of Electrical and Computer Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran
hr.hosseinzadeh@srbiau.ac.ir
Farbod Razzazi
Department of Electrical and Computer Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran
razzazi@srbiau.ac.ir
Afrooz Haghbin
Department of Electrical and Computer Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran
haghbin@srbiau.ac.ir

Received: 12/Oct/2013 Revised: 23/Jun/2014 Accepted: 30/Jun/2014

Abstract

Automatic modulation classification (AMC) in detected signals is an intermediate step between signal detection and demodulation, and is also an essential task for an intelligent receiver in various civil and military applications. In this paper, we propose a semi-supervised Large margin AMC and evaluate it on tracking the received signal to noise ratio (SNR) changes to classify most popular single carrier modulations in non-stationary environments. To achieve this objective, two structures for self-training of large margin classifiers were developed in additive white Gaussian noise (AWGN) channels with priori unknown SNR. A suitable combination of the higher order statistics (HOS) and instantaneous characteristics of digital modulation are selected as effective features. We investigated the robustness of the proposed classifiers with respect to different SNRs of the received signals via simulation results and we have shown that adding unlabeled input samples to the training set, improve the tracking capacity of the presented system to robust against environmental SNR changes. The performance of the automatic modulation classifier is presented in the form of k-fold cross-validation test, classification accuracy and confusion matrix methods. Simulation results show that the proposed approach is capable to classify the modulation class in unknown variable noise environment at even low SNRs.

Keywords: Automatic Modulation Classification; AMC; Tracking Performance Evaluation; Passive-Aggressive Classifier; Self Training; Semi-Supervised Learning.

1. Introduction

Automatic modulation classification (AMC) is the process of identification of the modulation type of a signal in a general non-cooperative environment. It has received significant attention for over two decades now. However, nowadays, there is a vast variety of applications that it is essential to detect and demodulate the signal without given priori information about the received signal [1-3]. In such cases, unlike regular receivers in which the primitive information of the received signals, such as carrier frequencies, frequency bandwidth, bit rate, and modulation type, is available, there is no information available for the received signal, and the receiver is blind. In this case, an automatic modulation classifier as an intermediate step between signal interception and information recovery, helps the receiver employ the correct decoder. The AMC can be used in a wide range of applications including electronic warfare applications, intelligent services systems, spectrum monitoring, signal surveillance, interferer identification and cognitive radio applications [4].

Generally, digital signal type classification algorithms may be divided into two major categories: decision theory based approaches (DTBAs) and feature matching based approaches (FMBAs) [5-6]. DTBA algorithms are based on the received signal likelihood function and hypothesis testing arguments to formulate the classification problem. In contrast, FMBA algorithms usually extract the features form the received signal and determine the membership of signal to each class. The calculation complexity of FMBA algorithms is lower than DTBA algorithms, and they are easy to implement. Our proposed tracking classifier is categorized as FMBA algorithm.

From the published works in AMC classifiers [7-8], it appears clear that unfortunately, most of the previous studies rely on the availability of large labeled datasets to analyze the classifier. However, it is not applicable in most non stationary environments applications. Thus, it is required to deal with both labeled and unlabeled data to train the classifier, simultaneously. In this approach, the classifier is first trained with labeled data and then used to predict the labels of the unlabeled data. A subset of the unlabeled data, together with their predicted labels, is then selected to augment to the labeled data. This training model is called self-training [9-12].

In this paper, we propose two large margin architectures to track the signal-to-noise ratio (SNR)

* Corresponding Author

changes of input signals, and then classify the modulated signals. The architectures are developed based on passive-aggressive online algorithm [13] to train the classifier and determine the modulated signal type. Our preliminary result on tracking performance of the proposed classifiers was reported in [14].

The remainder of the paper is organized as follows: in section 2 we provide a description of selected feature extraction and margin based algorithms as the basis of the proposed architectures. Brief information on the development of passive-aggressive algorithm is also provided in this section. In Section 3, we present the proposed architectures for classifier tracking. The simulation results are presented in section 4. Finally, the paper is concluded in section 5 with a summary of our findings.

2. Problem Description

According to the increasing use of digital systems and usage of digital signals in software radio, the research was focused on digital signal types. Considering the changes in message parameters, there are four general digital signal types, M-ASK, M-PSK, M-FSK and M-QAM [15]. The modulation techniques of digital input signals, which are considered in this paper, are FSK2, FSK4, ASK2, ASK4, PSK2, PSK4, PSK8, QAM16, QAM32, and QAM64.

2.1 Feature Extraction

Feature extraction is a crucial part of a pattern recognition system where its aim is to reveal the distinctive properties of an object to be recognized. In this paper, the effective features are considered as a combination of higher order statistics and instantaneous characteristics of digital signal types. The rest of this subsection is devoted to describe these features briefly.

2.1.1 Instantaneous Feature

Instantaneous features are suitable for signals which contain instantaneous phase or instantaneous frequency [16]. In this work, the instantaneous features for classification were selected from the proposed features by Azzouz and Nandi [17-18]. These features were derived from the instantaneous properties of the received signals. Therefore, these features are called as instantaneous features. The instantaneous key features which were used for the proposed tracking algorithm were derived from the instantaneous amplitude $a(t)$, and the instantaneous frequency $f(t)$, of the signal under consideration.

The first feature is the maximum value of the power spectral density of the normalized-centered instantaneous amplitude of the intercepted signal which is formulated as follows:

$$\gamma_{max} = max\left(\frac{|DFT(a_{cn}(i))|^2}{N_s}\right) \tag{1}$$

Where N_s is the number of the sample in the range and $a_{cn}(i)$ is value of centralized normalized instantaneous amplitude that is defined by

$$a_{cn}(i) \overset{\Delta}{=} a_n(i) - 1\,, \quad a_n(i) \overset{\Delta}{=} \frac{a(i)}{m_a} \tag{2}$$

and m_a is the average value of instantaneous amplitude over one frame, i.e.

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

This feature is designed to discriminative between constant envelopes (CE) signals (e.g., FSK and PSK) and non-CE signals (e.g., ASK).

The second feature is the standard deviation of absolute value of normalized-centered instantaneous frequency over non-weak segments of the intercepted signal which is calculated as:

$$\sigma_{af} \overset{\Delta}{=} \sqrt{\frac{1}{L}\left[\sum_{a_{cn}(i)>a_t} f_{cn}^2(i)\right] - \left[\frac{1}{L}\sum_{a_{cn}(i)>a_t} |f_{cn}(i)|\right]^2} \tag{4}$$

where $f_{cn}(i)$ is the centralized normalized instantaneous frequency and it is defined by

$$f_{cn}(i) = \frac{f_c(i)}{r_b},\ \ f_c(i) - m_f,\ m_f = \frac{1}{N}\sum_{i=1}^{N} f(i) \tag{5}$$

where r_b is the bit rate, and a_t is a preset threshold for detecting non-weak samples because instantaneous frequency is very noise sensitive. In this paper, the threshold for detection of non-weak samples is chosen as $a_t = 0.95$ [17]. This feature is designed to discriminative between FSK signals.

2.1.2 Higher Order Statistics (HOS)

The first set of employed statistical features is moments. A moment of a random variable may be defined as:

$$M_{p,q} = E[s^{p-q}(s^*)^q] \tag{6}$$

Where *p* is called the moment order and s* stands for the complex conjugation of s.

The second set of employed statistical features is cumulants which is the most widely used feature in this area. The symbolism for *p*th order cumulants is similar to that of the *p*th order moment.

$$C_{p,q} = Cum[s, \ldots, s, s^*, \ldots, s^*] \tag{7}$$

The mentioned expression have *(p-q)* terms of *s*, in addition to *q* terms of *s**. Cumulants may be expressed in term of moments as

$$Cum[s_1, \ldots, s_n] = \sum_{\forall v}(-1)^{q-1}(q-1)!\,E\left[\prod_{j\in v_1} s_j\right] \ldots E\left[\prod_{j\in v_q} s_j\right] \tag{8}$$

where the summation index is over all partition $v = (v_1, \ldots, v_q)$ for the set of indices $(1, \ldots, n)$, and q is the number of elements in a given partition.

Based on fisher discriminant analysis (FDA) [19], we selected a proper set of higher order moment and cumulants as below. FDA represents the capability of the selected features for separation of two predefined classes and is defined by

$$d_{ij} = \frac{(\mu_i - \mu_j)^2}{{\sigma_1}^2 + {\sigma_2}^2} \quad i \neq j \tag{9}$$

where μ and σ are mean and variance of these two classes. The important selected statistical features are $M_{41}, M_{61}, M_{84}, C_{40}, C_{61}, C_{63}, C_{80}, C_{82}, C_{84}$. Unfortunately, these characteristics are noise dependent. Therefore, a classifier tracking strategy should be devised to decrease the effect of this dependency, as far as possible.

2.2 Large-Margin Classifier

Suppose to have a binary classification problem and a training set of l labeled samples $\{x_i, y_i\}_{i=1}^{l}$ and u unlabeled samples $\{X_i\}_{i=1}^{u}$ where $x_i \in R^d$ is an input vector describing i^{th} samples and $y_i \in \{-1,1\}$ is its labels. We want to learn a discriminative function f in online and assign the correct label to an unseen new test samples.

2.2.1 Passive-Aggressive Classifier

Passive-Aggressive (PA) algorithm [13] is a maximum margin based learning algorithm, which has been mainly used in online learning. Online PA algorithm, on one hand, modifies the current classifier $w_i . x + b$ in order to correctly classify the current example x_i, by updating the weight vector from w_i to w_{i+1}; and on the other hand, the new classifier $w_{i+1} . x + b$ should be as close as possible to the current classifier $w_i . x + b$. Our idea for tracking the environmental conditions was based on PA algorithm, and we pursued both above ideas at the same time. The vector w_1 is initialized to $(0, \ldots, 0)$. In the time i, the new weight vector w_{i+1} was determined by solving the following optimization problem,

$$\begin{aligned} w_{i+1} &= argmin_w \frac{1}{2} \|w - w_i\|^2 \\ s.t.\ & (w_{pi}.x_i - w_{yi}.x_i) \geq 1 \end{aligned} \tag{10}$$

where w_{pi}, w_{yi} are the weight vectors that is produced by using predicted labels and true labels respectively. With a Lagrange multiplier, τ, the stationary points of the Lagrangian can be computed. Then, we can reformulate the optimization problem as follows:

$$w_{i+1} = w_i + \tau_i x_i; \quad \tau_i = \frac{1 - (w_{yi}.x_i - w.x_i)}{2\|x_i\|^2} \tag{11}$$

In order to ensure convergence of the algorithm, Cramer et al. [13] have presented a scheme that employ robust update strategy. This scheme has been obtained by introducing non-negative slack variable into the optimization problem. It is derived with soft-margin classifiers [20]. This makes the optimization problem (10) to the following formulation:

$$\begin{aligned} w_{i+1} &= \min_w \frac{1}{2} \|w - w_i\|^2 + C\xi \\ s.t.\ & \left(w_{pi}.x_i - w_{yi}.x_i\right) \geq 1 - \xi,\ \xi \geq 0 \end{aligned} \tag{12}$$

where ξ is slack variable, and c is a penalty parameter.

In real-world problems, to improve the discriminative power of f, the training data are usually mapped into a higher-dimensional feature space via a non-linear mapping $\emptyset(x)$, induced by a kernel function $k(x, x') = \emptyset(x).\emptyset(x')$. In the case of multiclass classification, the prediction output is a vector in $\mathbb{R}^k$ where each element in the vector corresponds to score assigned to the respective label.

These score calculations, have been devised in [21]. The prediction of PA algorithm is set to be the label with the highest score.

3. Proposed Architectures for Tracking Performance Evaluation

In this section, we propose two large margin structures, using the mentioned passive-aggressive algorithm and evaluate the tracking performance and classification accuracy of them. In this study, the SNR of the signals is a priori unknown. The tasks of the proposed tracking algorithms are adaptation to the environment's SNR in addition to detecting the modulation type. In the tracking performance evaluation, it is assumed that the SNR changes are much slower than signal changes. To evaluate the system, we made a synthetic signal which the SNR of its temporal segments were decreased gradually. In the rest of this section, we released two systems for tracking performance evaluation.

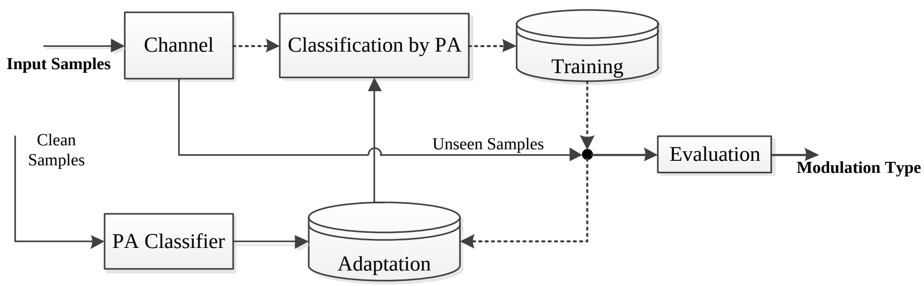

Fig. 1. The overall block diagram of batch learning

3.1 Batch Learning Architecture for Tracking

In this section, we introduce a semi-supervised batch learning architecture. This architecture is semi-supervised since using a reasonable set of recognized samples; it is capable to classify the input samples correctly. The system is first adapted to the unlabeled training samples to have a new adapted trained classification system. In the classification phase, this new classifier is used to classify unknown samples.

The overall block diagram of the batch learning system is presented in Fig. 1. Following the main signal flow in the architecture, we trained a general classifier using the union of gathered features of the clean samples ($SNR = \infty$). The clean samples are generated noise-free. The unlabeled input signals are randomly divided into s equal size subsets. Subsequently, PA classifiers are used as test labels and then adapted to the new subset by $s - 1$ times. In other words, during the each iteration, one of the s subsets is introduced to this system to classify.

Then, using the predicted samples of this subset, the classifier was adapted.

3.2 Online Learning Architectures for Tracking

In this section, we propose self-training classifier architecture so that the classification result of each sample is determined and evaluated as the sample enters the system. The evaluated samples are then re-used as new training samples. In this study, it is assumed that the label of input samples is not determined.

The block diagram of online learning system is shown in Fig. 2. As depicted, the classifier is first trained with the clean labeled samples. The obtained result is used to classify a new unlabeled input sample that have been entered, and classify them. Then, the unlabeled input samples that are predicted one by one is collected to augment into the labeled samples gradually. Then, the classifier is adapted using this new labeled sample and the result is used to classify the next unlabeled samples every

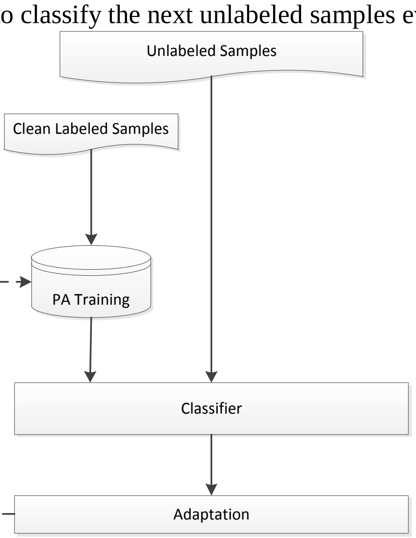

Fig. 2. The overall block diagram of online learning

time. This procedure is repeated until the last unlabeled sample is entered. The technique is presented in Fig. 2.

4. Experimental Results

This section presents simulation results on the proposed architectures. The radial basis function (RBF) of the form $k(x_i, x_j) = e^{-\gamma\|x_i - x_j\|^2}$ is employed as the kernel function of PA classifier. In practice, the standard method to determine the kernel and misclassification penalty parameters is through the grid search method [22]. Therefore, a grid search technique was used to find the optimal values of these parameters.

The carrier frequency (f_c) was assumed to be 11 MHz, the sampling rate (f_s) was 86 kHz, and symbol rate (r_s) was 10 kHz. We assumed that the carrier frequency had been correctly estimated before or it had been known. Therefore, we only considered complex baseband signals. In addition, it was assumed that the simulated signals were bandwidth limited. The Gaussian noise was added according to 0 dB, 3 dB, 4 dB, 6 dB, 8 dB, 9 dB, 12 dB and 20 dB SNRs. Each signal used in this study was generated using MATLAB. For every signal we generated 300 realizations which created randomly for every trial to ensure results are independent of the considered input samples. Therefore, the number of samples in a segment is 3000. To evaluate the system, we made a synthetic signal with 8 different segments which the SNR of its temporal segments were decreased gradually form 20 dB to 0 dB. This results in 24000 samples in the synthetic signal. We have adopted 3-fold cross-validation procedure on labeled and unlabeled dataset.

We compared the performance of the algorithms on the basis of the classification accuracy. Classification accuracy for experiment is taken as the ratio of the number of samples correctly classified to the total number of samples.

Table 1. Classification rate of proposed batch learning architecture in different SNRs (%)

Modulation classes	SNR			
	0 dB	4 dB	8 dB	12 dB
FSK2	54.2	58.1	81.2	100
FSK4	76.1	93.5	100	100
ASK2	66.7	91.9	100	100
ASK4	58.4	74.0	91.0	100
PSK2	80.3	100	100	100
PSK4	87.1	95.8	100	100
PSK8	81.1	87.8	100	100
QAM16	44.0	61.9	75.5	96.6
QAM32	84.1	81.1	100	100
QAM64	81.0	99.2	100	100
Mean	71.30	84.33	94.78	99.66

Table 2. Classification rate of proposed online learning architecture in different SNRs (%)

Modulation classes	SNR			
	0 dB	4 dB	8 dB	12 dB
FSK2	95.0	100	100	100
FSK4	99.8	100	100	100
ASK2	80.8	94.8	100	100
ASK4	69.6	87.0	100	100
PSK2	92.0	100	100	100
PSK4	71.3	89.1	100	100
PSK8	70.3	88.0	100	100
QAM16	99.8	99.9	98.9	100
QAM32	70.1	87.6	100	100
QAM64	66.5	83.1	100	100
Mean	81.52	92.95	99.89	100

Table 3. Confusion Matrix of proposed batch learning algorithm in SNR=4 dB (%)

True modulations	Predicted modulations									
	FSK2	FSK4	ASK2	ASK4	PSK2	PSK4	PSK8	QAM16	QAM32	QAM64
FSK2	58.1	41.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
FSK4	0.0	93.5	0.0	0.0	0.0	0.0	0.0	0.0	6.5	0.0
ASK2	0.0	0.0	91.9	6.6	0.0	1.5	0.0	0.0	0.0	0.0
ASK4	0.3	0.0	15.5	74.1	0.0	10.1	0.0	0.0	0.0	0.0
PSK2	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0
PSK4	0.0	0.0	0.0	0.0	0.0	95.8	0.0	4.2	0.0	0.0
PSK8	0.0	0.0	0.0	0.0	0.0	0.0	87.8	0.0	12.2	0.0
QAM16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	61.9	0.0	38.1
QAM32	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	81.2	18.8
QAM64	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	99.3

Table 4 Confusion Matrix of proposed online learning algorithm in SNR=4 dB (%)

True modulations	Predicted modulations									
	FSK2	FSK4	ASK2	ASK4	PSK2	PSK4	PSK8	QAM16	QAM32	QAM64
FSK2	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
FSK4	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ASK2	0.0	0.0	94.8	5.2	0.0	0.0	0.0	0.0	0.0	0.0
ASK4	0.2	0.0	12.7	87.1	0.0	0.0	0.0	0.0	0.0	0.0
PSK2	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0
PSK4	0.0	0.0	0.0	0.0	0.0	89.2	0.0	10.8	0.0	0.0
PSK8	0.0	0.0	0.0	0.0	0.0	0.0	87.9	6.1	6.0	0.0
QAM16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.0	0.1
QAM32	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.3	87.7	0.0
QAM64	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16.8	0.0	83.2

Classification accuracy assessments of different classes have been provided by the confusion matrix and accuracy matrix in percentage.

4.1 Tracking Performance Evaluation of Proposed Architectures

In this section, we have evaluated the performance of proposed architectures with different SNR values via simulation. Tables 1-2 show the classification rates of the proposed classifier in 0, 4, 8 and 12 dB SNRs.

From the mentioned results in Tables 1-2, it can be deduced that the performance of classifier in different SNR were generally good. This fact is because of the adaptation ability of the proposed classifiers. However, the performance is slightly degraded in lower SNRs. It should be noted that the previous Offline algorithms have not been able to learn by 300 samples.

As a sample, the confusion matrix is presented at SNR=4 dB to analyze in the confusion of different classes. These results are presented in Tables 3-4.

As it is observed in Table 4, the recognition accuracy of online batch learning architecture for all signals except QAM64 was very good even at low SNRs. It can be seen that there is a tendency for QAM64 modulation to be mostly confused with QAM16 modulation. Because the constellation shape of these classes are very similar.

In addition, the tracking performance of online learning classifier is compared to accuracy in non-variable stationary environment in the same SNR. Numerical results of the accuracy in stationary environment were reported in [23]. This comparison is shown in Fig. 3.

According to Fig. 3, it can be observed that the classification accuracies which were obtained from the proposed tracking methods are close to classification accuracy in stationary environment.

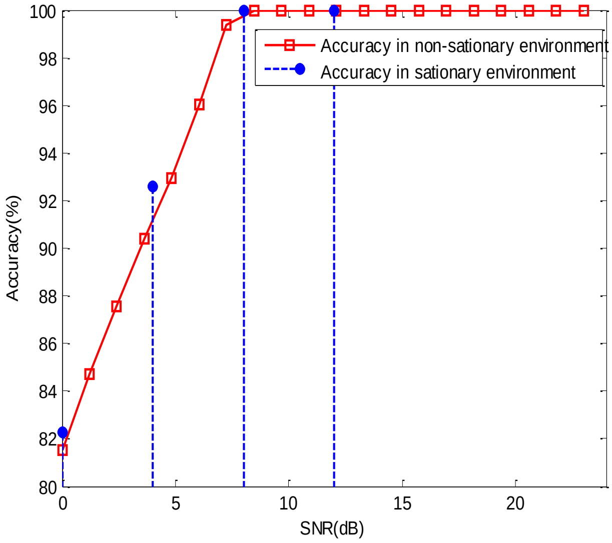

Fig. 3 The curve of accuracy tracking relative to SNR

Table 5. Tracking Performance comparison (%)

SNR	Tracking evaluation of proposed online learning architecture	Tracking evaluation of proposed batch learning architecture	Supervised PA classifier with priori labeled data	Supervised SVM classifier with priori labeled data
0 dB	81.82	80.06	85.06	84.18
4 dB	92.93	88.27	97.21	97.36
8 dB	99.84	99.84	99.84	99.84
12 dB	100	100	100	100
Noise-free	100	100	100	100

4.2 Tracking Performance Comparison

This section, demonstrates the effectiveness of the proposed methods. Therefore, the performance of the proposed tracking architectures were compared to the tracking algorithm that was trained by PA algorithm in the matched SNR (SNR aware mode) with priori labeled data. In addition, this comparison was repeated for SVM algorithm. The results are indicated in Table 5.

As it is observed in Table 5, all of the evaluated algorithms have similar performance in SNR$\geq$ 8. Simulation results show that performance of proposed algorithms deteriorate with decreasing SNR. The tracking performance of online learning classifier is 98% of supervised PA classifier with priori labeled data. Therefore, the proposed algorithm has generally good performance and its accuracy is close to the analysis in the matched SNR.

5. Conclusions

Automatic modulations classification plays a significant role in civil and military applications. In this paper, a new method are developed for classifier tracking in non-stationary environments using a semi-supervised self-trained large margin classifier for classification of most popular single carrier modulations, i.e., FSK2, FSK4, ASK2, ASK4, PSK2, PSK4, PSK8, QAM16, QAM32, and QAM64 which are commonly used in a cognitive radio. Towards this objective, two features are employed that include a selected set of the instantaneous characteristics and higher order statistics of received signal. Simulation results show that the approach is capable to classify the modulation class in unknown variable noise environment at even low SNRs. In addition, the evaluation of the tracking performance shows that the proposed architectures have a good ability to adapt to the environment. The tracking performance evaluation for high frequency band, fading and multipath phenomena in addition to multi-carrier modulation schemes will be investigated as our future work.

References

[1] J. J. Popoola and R. van Olst, "A Novel Modulation-Sensing method," IEEE Vehicular Technology, Vol. 6, No. 3, Sep. 2011, pp. 60-69.

[2] B. Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection," IEEE Circuits and Systems, Vol. 2, No. 2, Second Quarter 2009, pp. 27-45

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," IEEE J. Select. Areas Commun., Vol. 23, No. 2, 2005, pp. 201-220.

[4] J. J. Popoola and R. van Olst, "The performance evaluation of a spectrum sensing implementation using an automatic modulation classification detection method with a Universal Software Radio Peripheral," Expert Systems with Applications, Vol. 40, No. 3, 2013, pp. 2165-2173.

[5] O. Dobre, A. Abdi, Y. Bar-Ness and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," IET Commun., Vol. 1, No. 2, Apr. 2007, pp. 137-156.

[6] H. C. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," IEEE Trans. Wireless Commun., Vol. 7, No. 8, 2008, pp. 3089-3105.

[7] A. Ebrahimzadeh, and R. Ghazalian, "Blind digital modulation classification in software radio using the optimized classifier and feature subset selection," Engineering Applications of Artificial Intelligence, Vol. 24, No. 1, 2011, pp. 50-59.

[8] Z. Zhu, M. W. Aslam, A. K. Nandi, "Genetic algorithm optimized distribution sampling test for M-QAM modulation classification," Signal Process., Vol. 94, 2014, pp. 264-277.

[9] M. M. Adankon and M. Cheriet, "Help-Training for semi-supervised support vector machines," Pattern Recognition, Vol. 44, No. 1, Sep. 2011, pp. 2220-2230.

[10] X. Zhu, "Semi-supervised learning literature survey", Technical Report 1530, Computer Sciences, University of Wisconsin-Madison, 2008.

[11] X. Zhu and A. B. Goldberg, Introduction to Semi-Supervised Learning, Morgan and Claypool Publishers, 2009.

[12] O. Chapelle, B. Scholkopf, A. Zien (Eds.), Semi-Supervised Learning, MIT Press: Cambridge, MA, 2006.

[13] K. Crammer, O. Dekel, J. Keshet, S. Shalev-shwartz, and Y. Singer, "Online passive-aggressive algorithms," Journal of Machine Learning Research, Vol. 7, 2006, pp. 551-585.

[14] H. Hosseinzadeh, F. Razzazi, A. Haghbin, "Tracking Performance of Online Large Margin Semi-supervised classifiers in Automatic Modulation Classification." In: Proceedings of International Symposium on Telecommunications (IST'2012), 2012, pp. 387-392.

[15] J. G. Proakis, Digital Communications, McGraw-Hill: New York, 2001.

[16] J. Lopatka and P. Macrej, "Automatic modulation classification using statistical moments and a fuzzy classifier," in Proc. Of International Conference on Signal Processing (ICSP'00), Aug. 2000, Vol. 3, pp. 1500-1506.

[17] E. E. Azzouz and A. K. Nandi, "Automatic identification of digital modulation types," Signal Processing, Vol. 47, No. 1, Nov. 1995, pp.55-69.

[18] E. E. Azzouz and A. K. Nandi, "Algorithms for automatic modulation recognition of communication signals," IEEE Trans. Commun., Vol. 46, No. 4, 1998, pp.431-436.

[19] R. Duda, P. Hart, and D. Stock, Pattern Classification, Wiley, 2000.

[20] V. Vapnik, Statistical Learning Theory, John Wiley & Sons, Inc.: New York, 1998.

[21] K. Crammer and Y. Singer, "A new family of online algorithms for category ranking," Journal of Machine Learning Research, Vol. 3, 2003a, pp. 1025–1058.

[22] C. C. Burges, "A tutorial on support vector machines for pattern recognition," In Proceedings of Int. Conference on Data Mining and Knowledge Discovery, Vol. 2, No. 2, 1998, pp. 121-167.

[23] H. Hosseinzadeh, F. Razzazi, A. Haghbin, "Online Large Margin Semi-supervised Algorithm for Automatic Classification of Digital Modulations." In: Proceedings of 21st Iranian Conference on Electrical Engineering (ICEE'2013), 2013, pp. 1-6.

Hamidreza Hosseinzadeh received the B.Sc. degree in Electrical Engineering from Shahr-e-Rey branch, Islamic Azad University, Tehran, Iran, in 2009 and the M.Sc. degree in Electrical Engineering from Science and Research Branch, Islamic Azad University, Tehran, Iran, in 2012. He is currently a Ph.D. candidate at department of Electrical and Computer Engineering in Islamic Azad University, Science and Research Branch, Tehran, Iran. His research areas of interest include signal processing, pattern recognition, and communication systems.

Farbod Razzazi received his B.Sc. and M.Sc. in Electrical Engineering from Sharif University of Technology in 1994 and 1996, respectively. He received his Ph.D. from Amirkabir University of Technology (Tehran Polytechnic) in 2003 in Electrical Engineering. He is currently an Assistant Professor in Department of Electrical Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran. His current research interests are pattern recognition methods and their applications in speech recognition systems.

Afrooz Haghbin received her B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 2001. She received her M.Sc. degree from Tehran University and her PhD degree from Tarbiat Modares University, Tehran, Iran, all in electrical engineering in 2004 and 2009, respectively. She is currently with the electrical and computer department of Science and Research Branch, Islamic Azad University, Tehran, Iran, as assistant professor. Her research interests include MIMO wireless communications, channel coding, pre-coding, multicarrier modulation and estimation theory.