Novel Automatic Modulation Classification using Correntropy Coefficient

Aluisio I. R. Fontes, Lucas C. P. Cavalcante and Luiz F. Q. Silveira

***Abstract*— This paper deals with automatic modulation classification (AMC) of communication signals. A new method for the automatic classification using a similarity measure derived from Information Theoretic Learning (ITL), called *correntropy coefficient*, is proposed. Unlike many of the conventional methods, the proposed method does not require any signal pre-processing. Further, the proposed AMC technique uses a simple scheme of evaluating the correntropy coefficients, calculated over templates containing the common features of digitally modulated signals, in the classification task. The performance of the classifier is presented in the form of classification hit-rates under AWGN noisy conditions, with Signal-to-Noise Ratios (SNRs) at the range of -5 dB to 15 dB. Simulation results with binary modulations show classification hit-rates of 83% for -5 dB and 99% for 0 dB.**

***Keywords*— Correntropy Coefficient, Automatic Modulation Classification, Radio Cognitivo.**

I. INTRODUCTION

Modern wireless systems employ adaptive techniques aiming to provide high data throughput while maximizing requirements like coverage, Quality of Service (QoS) and capacity. Link adaptation (Adaptive modulation and coding) and Hybrid Automatic Repeat ReQuest (HARQ) are just some examples of such techniques. A promising strategy to improve data rate in wireless channels is the cognitive radio, an intelligent wireless communication system that adapts to the communication environment to provide a highly reliable communication and an efficient utilization of radio spectrum [6]. Automatic Modulation Classification proves to be essential in such scenarios.

Automatic Modulation Classification is a class of techniques for recognizing the type of digital modulation scheme used to generate a received modulated signal, with little or even no prior knowledge (such as its phase, frequency or amplitude) about the modulated signal itself [11]. Performing AMC is already a hard task, specially because the lack of knowledge about the signal. It becomes even harder when the received signal suffers from interferences, noise and channel fading. AMC techniques currently reported on literature [2]–[5], [7], [8], [10], [16], [17] employ a pre-processing module in order to extract signal features usable for classification, which may, depending on the applied mechanism, make assumptions about the received signal which may not hold (e.g. AWGN being the unique source of noise), or even can demand a high computational cost to be implemented. Besides, in the AMC literature, certain parameters can be assumed to be known before invoking an AMC process [1]. In particular, a perfect symbol timing with no timing offset is assumed to be known in this study.

This paper evaluates the direct use of *correntropy coefficient*, derived from Information Theoretic Learning (ITL), for AMC purposes. Correntropy coefficient supplies a similarity measure between two random signals, implicitly taking into account all even-order moments of the signals [9]. It should be stressed that the larger the statistical similarity between two random signals, the larger will be the respective correntropy coefficient calculated.

In the proposed AMC technique, each digital modulation scheme is characterized by a template containing the common features of digitally modulated signals for such scheme. In the receiver, the correntropy coefficient is calculated between the received signal and each comparative template, and a decision is made in favour of the modulation that produces the biggest correntropy coefficient value.

Numerical results presented in this study indicate that the proposed AMC scheme, based only on calculation of a statistical similarity measure via correntropy coefficient, is so effective in the classification task of binary modulations of digital signals affected by AWGN noise, that the use of an additional pre-processing phase for feature extraction of the received signal is not necessary.

The remainder of this paper is organized as follows. Section II briefly describes the AMC's state-of-the-art, discussing the requirements and limitations of the techniques commonly found on literature. Section III presents some similarity measures derived from the ITL framework, in special the correntropy coefficient. Section IV presents the proposed AMC technique, based on the correntropy

Aluisio I. R. Fontes, Lucas C. P. Cavalcante and Luiz F. Q. Silveira¸ Universidade Federal do Rio Grande do Norte, Natal - RN, Brazil, E-mails: aluisio@dca.com, lucas@dca.ufrn.br, lfelipe@dca.ufrn.br.

coefficient, while Section V presents numerical results obtained for the proposed AMC scheme. Finally, Section VI presents conclusions, considerations and future work perspectives.

II. AUTOMATIC MODULATION CLASSIFICATION (AMC)

Research on AMC can be grouped in two approaches, either *using statistics from the received signal* to define a Maximum Likelihood (ML) function or *extracting signal features*, for performing the classification using different Pattern Recognition techniques [10], [13]. Authors in [2] presents a survey of AMC approaches (updated in [16]) considering ML-based AMC. In any of those approaches, the classifying system must be capable of correctly determining the type of modulation scheme for a given signal sample among a set of N candidate modulation schemes. An ideal AMC must also attend to the following requirements [2]:

- Provide high probability of True Positive (TP) and low probability of False Positive (FP) classifications, requiring for that a short observation interval;
- Perform AMC on signals from different modulation schemes and subject to varying environment conditions, such as different propagation effects and noises;
- Be embeddable, work on real-time, and have low computational complexity.

The AMC process involves essentially two stages [2], namely: (i) a signal pre-processing stage for feature extraction; and (ii) an adequate selection of the modulation scheme based on the selected signal features by the classifier system. Figure 1 shows how those stages interact on AMC.

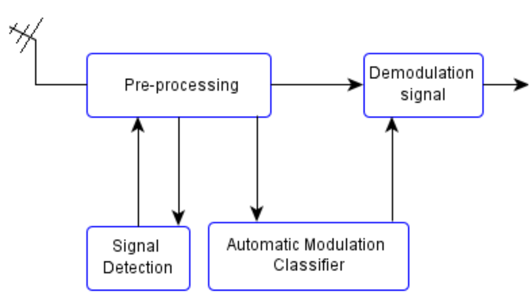


Fig. 1. Automatic Modulation Classification (AMC)

On the *signal pre-processing* stage, the focus is on estimating system parameters, such as *carrier frequency*, *symbol period*, or *signal power*, or even providing noise reduction and channel equalization. However, commonly pre-processing techniques for AMC are not restricted to such tasks and include e.g. signal feature extraction.

In general, many techniques have been proposed on literature for signal feature extraction, allowing improving the performance of the classifiers in the following AMC stage. Authors in [3] propose calculating the standard deviation of the received signal in the pre–processing stage and using it to train an Artificial Neural Network (ANN) for classification purposes. In [8], the *wavelet* transform for feature extraction on QAM, PSK and FSK signal samples is used, while authors in [5] use features obtained from wavelet domain to perform AMC with the use of an ANN. There are also works proposing different AMC techniques based on higher order statistics and cyclostationary features from the modulated signals [4], Principal Component Analysis (PCA) [7], and ANN with Fuzzy Logic [17]. In fact, some of those pre-processing tasks demand a high computational cost, which limits its application, or even make it not practical for real-time systems with current off-the-shelf technology. This work proposes an AMC technique that uses a similarity measure of reasonable computational complexity in association with a very simple classifier.

III. INFORMATION THEORETIC SIMILARITY MEASURES

%The common problem faced by many data processing professionals is how to best Similarity is a key concept to quantify temporal signals [9]. Correntropy is a generalized similarity measure between two arbitrary scalar random variables X and Y defined by

$$v(X,Y) = E_{XY}[k(X,Y)] = \iint k(x,y)p_{X,Y}(x,y)dxdy, \tag{1}$$

where the expected value is over the joint space and $k(\cdot,\cdot)$ is any continuous positive definite kernel. Correntropy is a well-defined function provided $k(x,\, y)$ belongs to L_∞ (i.e. its maximal value is finite) [9].

It can be verified that the specific positive definite kernel $k(x,\, y) = xy$ substituted in Equation (1) yields cross-correlation. In this paper, the correntropy measure is based on the Gaussian kernel, given by Equation (2), which is a symmetric, translation-invariant kernel.

$$k(x,\, y) = K_\sigma(x_i - y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_i - y_i)^2}{2\sigma^2}}. \tag{2}$$

Substituting the Taylor series expansion of the Gaussian kernel function in Equation (1), and assuming that it is valid to interchange the integral with the sum, the correntropy can be expressed by [9]

$$v(X,Y) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma^{2n} n!} E[(X-Y)^{2n}]. \tag{3}$$

Equation (3) states that the correntropy is constituted by a summation of all even moments of the difference variable. Thus correntropy keeps the nice bivariate form of correlation, but is still sensitive to the sum of second- and higher-order moments of the random variables. This is an interesting characteristic, because in many applications this sum may be sufficient to quantify better than correlation the relationships of interest and it is simpler to estimate than the higher-order moments [9]. This makes correntropy extremely sensitive to higher-order moments and allows it to extract more information from the random variables. Besides, as the internal product on $K_\sigma(x_i - y_i)$ also tends to zero, correntropy is seen as a robust similarity measure sensitive to time-varying random processes.

In practice, the joint PDF in Equation (1) is unknown and only a finite number of data $\{(x_i,\, y_i)\}_{i=1}^{N}$ are available, leading to the sample estimator of correntropy [9]

$$V(\mathbf{X},\mathbf{Y}) = \frac{1}{N}\sum_{i=1}^{N} K_\sigma(x_i - y_i). \tag{4}$$

In Equation (4), the Gaussian kernel is responsible for mapping the random vectors in a feature space named *Reproducing Kernel Hilbert Space* (RKHS) [14].

Auto-correntropy is named after the correntropy measured between two random samples from the same random process, while cross-correntropy is named after the correntropy between samples from different random processes [15].

Principe [9] proposes the centered cross-correntropy, a centered correlation function measured on the RKHS whose estimator is defined as per Equation (5). It is used in favor of cross-correntropy as the non-linear mapping performed by the Gaussian kernel does not ensures zero mean even when the original samples are centered.

$$\begin{aligned} U(\mathbf{X},\mathbf{Y}) = \frac{1}{N}\sum_{i=1}^{N} K_\sigma(x_i - y_i) - \\ \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} K_\sigma(x_i - y_j). \end{aligned} \tag{5}$$

Santamaria et al. [12] present a new similarity measure, named *correntropy coefficient* and defined in accordance with Equation (6). It corresponds to the cosine of the angle between two random sample vectors transformed on the RKHS. Since correntropy coefficient implicitly takes into account all even-order moments of the difference vector, this measure is capable of extracting more information than the conventional correlation coefficient. Here, $U(\mathbf{X},\mathbf{Y})$ corresponds to the centered cross-correntropy between the vectors **X** and **Y**, while $U(\mathbf{X},\mathbf{X})$ and $U(\mathbf{Y},\mathbf{Y})$ correspond respectively to the centered auto-correntropy of the vectors **X** and **Y**. It can be seen that the correntropy coefficient assumes zero value when the two random variables are independent, and take values near to 1 (or -1) as more similar (or similar but with opposite values) the vectors are.

$$\eta(\mathbf{X},\mathbf{Y}) = \frac{U(\mathbf{X},\mathbf{Y})}{\sqrt{U(\mathbf{X},\mathbf{X})}\sqrt{U(\mathbf{Y},\mathbf{Y})}}. \tag{6}$$

This study has as hypothesis that correntropy coefficient may be used to characterize the dynamic interdependencies of different digital modulations, and therefore may fit well for AMC purposes. Reasons are twofold. The correntropy coefficient is sensitive to non-linearities and higher order statistical information, which are, in general, present in the modulated signals. Additionally, the received signal can be seen as a sequence of random variables, representing the addition of the modulated signal and the channel effects.

IV. Proposed Classifier

According to the proposed architecture, at the receiver, the input modulated signal is sampled at a time interval and it is stored for classification purposes. The classification occurs in two steps, as illustrated in Figures 2 and 3, and described as below:

1) At first, the proposed classifier calculates the correntropy coefficient between the stored signal and a set of comparative templates of noise-free modulated signals, which are clustered according to a specific modulation scheme. The amount of templates in each cluster associated with a modulation scheme is equal to 2^M, where M is the symbol number of the constellation. Besides, it was verified by numerical results, that the length of each template should be equal to $\frac{2}{T_s}\log_2 M$, where T_s is the sampling rate.

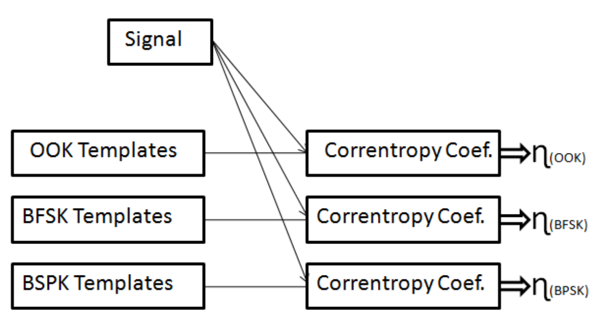


Fig. 2. Step 1 - Proposed Classifier

At 'Correntropy Coef.' block on Figure 2, each template of the set is compared with the input signal, as shown with more details in the most left part of Figure 3. The arguments for the calculation of each η are:

a) the self-correntropy of the input signal;
b) the self-correntropy for either element of the template's set (i.e 00, 01, 10, 11);
c) the correntropy between the input signal and each element described on previous item.

The result is a set of four links between input signals and the four elements (the template series) of every given set of modulation.

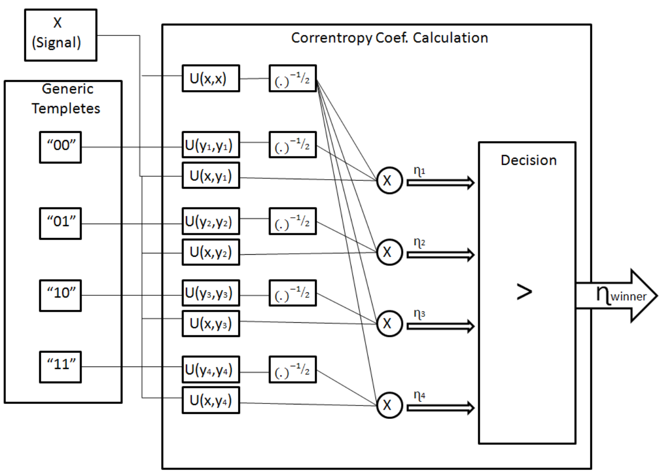


Fig. 3. Detailed schema of Step 1 plus Step 2

2) At the second step, a comparator decides in favor of the biggest value of correntropy coefficient calculated on the previous stage. The algorithm sets then the winner modulation based on the cluster that holds the winner coefficient. It's shown by the most right part of the Figure 3.

Due to the capacity of correntropy coefficient to characterize dynamic interdependencies between two random signal, even when they are corrupted by noise [12], no pre-processing stage is required for the proposed AMC, and so its computational complexity is relatively low. It should be observed that this computational simplicity is one of the main advantages of the proposed AMC scheme.

V. Simulation Experiments And Results

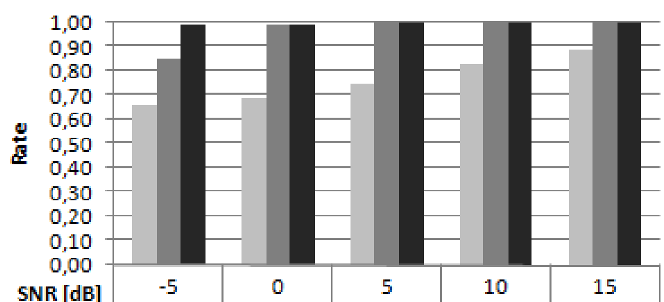
The proposed scheme is applied on the BFSK, BPSK and OOK binary digital modulation techniques.

The performance of the proposed classifier is presented in the form of classification hit-rates under AWGN noisy conditions, with Signal-to-Noise Ratios (SNRs) at the range of -5 dB to 15 dB. The symbol rate is set to 1 KSym/s. The carrier frequency is chosen to be 20 KHz and sampled with a sampling rate 100 KSa/s. The pulse shape was rectangular. A minimum amount of 3000 blocks was generated, for each modulation technique at -5 dB, 0 dB, 5 dB, 10 dB, 15 dB. Each block contains 2621 samples. The proposed AMC has been developed and evaluated by computational simulation via MATLAB®.

An important parameter to be adjusted in the *correntropy coefficient* is the σ (variance) used on the Gaussian kernel. It is a scaling factor that needs to be selected as a function of the data set dynamic range and the number of observed samples. In this work, σ was empirically evaluated and set to $1/\sqrt{2\pi}$.

Analyzing the equations of the proposed AMC, the correntropy coefficient has computational cost of $O(n^2)$, due to the double Gaussian summation in Equation (5).

Figure 4 shows the classification hit-rates of the proposed AMC scheme as a function of the evaluated SNRs. Considering BPSK, which does not have neither frequency variation nor information in the amplitude, the hit-rate is approximately equal to 99% for a SNR of -5 dB, which indicates that the proposed AMC, based only on correntropy measure, can be efficiently used in this classification context.



SNR [dB]	-5	0	5	10	15
BFSK	0,66	0,69	0,75	0,83	0,89
OOK	0,85	0,99	1,00	1,00	1,00
BPSK	0,99	0,99	1,00	1,00	1,00

Fig. 4. Numerical results with proposed AMC.

OOK signals has one interesting characteristic considering its susceptibility for classification purposes. The absence of energy in one OOK symbol could hide its statistical feature that our classifier method is trying to detect. This is the main reason of poor results for low SNR even compared to BPSK. In this situation, the whole OOK signal is statistically similar to AWGN.

From the obtained results, it can be verified that correntropy coefficient's capacity of extracting higher order statistical moments information provides a good signal classification without a high computational cost due to a pre-processing stage.

VI. Conclusion

This paper investigates the use of a AMC technique, based only on *correntropy coefficient*, for classification purposes of digital binary modulations. The proposed technique eliminates the need for any pre-processing stage, and presents good classification hit-rates. As future work, we intend to i) apply the method to higher-order modulation techniques (e.g. 4-QAM, QPSK, 4-ASK, 4-FSK); ii) eliminate the need for any prior knowledge about the received signal's sampling rate; and also iii) explore several channel effects in transmission as fast fading and multi-path in order to verify if the method's effectiveness remains a suitable alternative to traditional techniques.

Acknowledgment

The authors Aluisio Fontes and Lucas Cavalcante are respectively supported by the Committee for the Advancement of University Academic Staff - CAPES and by REUNI.

Referências

[1] A. A. B.-N. Y. Dobre, O.A. and W. Su, "The classification of joint analog and digital modulation." IEEE Military Commun. Conf., Vol. 5, 2005.

[2] O. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: classical approaches and new trends," *Communications, IET*, vol. 1, no. 2, pp. 137 –156, april 2007.

[3] I. E.-R. N. El-Madany, "Cognitive digital modulation classifier using artificial neural networks for NGNs." Alexandria,: 7th International Conference On Wireless And Optical Communications Networks (WOCN), 2010, pp. 1–5.

[4] A. Fehske, J. Gaeddert, and J. Reed, "A new approach to signal classification using spectral correlation and neural networks." Baltimore,: IEEE International Symposium on Dynamic Spectrum Access Networks, 2005, pp. 144–150.

[5] K. Hassan, I. Dayoub, W. Hamouda, and M. Berbineau, "Automatic modulation recognition using wavelet transform and neural networks in wireless system," vol. 2010, p. 13, 2010.

[6] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. ., pp. 201 – 220, february 2005.

[7] F. He, Y. Yin, and L. Zhou, "Principal component analysis of cyclic spectrum features in automatic modulation recognition." San Jose, CA,: IEEE Military Communications Conference (MILICON), 2010, pp. 1737–1742.

[8] Y. H. K.C. Prokopiw, W. Chan, "Modulation identification of digital signals by the wavelet transform," vol. 147, no. 4, pp. 169–176, 2000.

[9] J. C. Principe, *Information Theoretic Learning - Renyiś Entropy and Kernel Perspectives*, 1st ed. Nova York, EUA: Springer, 2010.

[10] M. Rahman, A. Haniz, M. Kim, and J. Takada, "Automatic modulation classification in wireless disaster area emergency network (W-DAEN)," in *Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), 2011 Sixth International ICST Conference on*, june 2011, pp. 226 –230.

[11] B. Ramkumar, "Automatic modulation classification for cognitive radios using cyclic feature detection," vol. 9, no. 2, pp. 27–45, 2009.

[12] I. Santamaria, P. P. Pokharel, and J. C. Principe, "Generalized correlation function: Definition, properties, and application to blind equalization," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2187–2197, 2006.

[13] H.-C. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 8, pp. 3098 –3105, august 2008.

[14] J. W. Xu, P. P. Pokharel, A. R. C. Paiva, and J. C. Principe, "Nonlinear component analysis based on correntropy." Vancouver,: International Joint Conference on Neural Networks, 2006, pp. 1851–1855.

[15] J. X. Xu, H. Bakardjian, A. Cichocki, and J. C. Principe, "A new nonlinear similarity measure for multichannel signals," vol. 21, pp. 222–231, 2008.

[16] J. Xu, W. Su, and M. Zhou, "Likelihood-ratio approaches to automatic modulation classification," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 41, no. 4, pp. 455 –469, july 2011.

[17] W. Zong, E. Lai, and C. Quek, "Digital modulation classification using fuzzy neural networks," vol. 30, pp. 101–116, 2006.