

# Performance Evaluation of Machine Learning based Signal Classification using Statistical and Multiscale Entropy Features

Arnau Mata Llenas, Janne Riihijärvi, Marina Petrova  
Institute for Networked Systems, RWTH Aachen University  
Kackertstrasse 9, D-52072 Aachen, Germany  
email: jar@inets.rwth-aachen.de

**Abstract**—In this paper we study the performance of machine learning based signal classification approaches in realistic wireless environments. We focus in particular on impact of interference from modulated signals and influence of realistic wireless channel conditions on classification performance. For this we use both numerical simulations as well as software defined radio based implementations. We also propose to use additional time series statistics originating from complex systems research as features for the classifier, in addition to classical second and higher order statistics previously employed. Our results show that the extended feature set results in robust classification performance in wide variety of channel conditions, and also when significant modulated interference is present in addition to Gaussian noise.

## I. INTRODUCTION

The radio environment where a typical wireless network operates in is becoming more and more complex. This is due to two major trends, first being the introduction of new radio access technologies, and the second being the ongoing densification of network infrastructures in order to meet exponentially increasing capacity demands. Spectrum sharing and coexistence issues both in co-channel and adjacent channel deployments are thus becoming more and more important for radio resource management in order to avoid excessing interference between wireless networks. Cognitive wireless networking principles have become an increasingly studied approach for solving such complex interference management problems in an automated fashion, in particular without introducing explicit signaling protocols for each combination of interacting wireless technologies [1]–[3]. However, using such cognitive approaches requires accurate estimation of the state of the radio environment, in particular forming an understanding of which kinds of wireless technologies are deployed and are causing interference with the managed network. Recognition of wireless transmitters is straightforward if coherent detection approaches can be used. Unfortunately this requires separate detection routines or even physical circuitry for each potential technology to be recognized, severely limiting the architectural scalability of such a system.

In this paper we study the use of machine learning techniques for signal classification based only on sampling of the *power* of the signal over time. Such approaches are

extremely attractive as they can be implemented on software for almost any radio, and do not require specialized signal processing routines or circuitry to function. Promising earlier results were achieved in [4] where classical higher order time series statistics of the power samples were used together with support vector machines (SVMs) as machine learning tool for signal classification. Based on these results several authors extended this work towards different modulation types as well as exploring alternative machine learning approaches and implementation architectures [5]–[8]. Common themes in this prior work are the exclusive of numerical simulations only especially in terms of the radio channel, and sole reliance on classical time series statistics as source of features. Also the focus of this earlier work has usually been on signal classification that is only contaminated by multipath effects and noise.

In this paper we extend the scope of machine-learning based signal classification to remove all of these limitations. We employ software defined radio boards to experiment with actual channels, as opposed to relying solely on simulations. We also study in detail the influence of weak interfering modulated signals contaminating the signal to be classified, as opposed to assuming a Gaussian noise environment only. Finally, we extend the used feature set to include modern *complexity metrics* as well. These provide information on the non-linear time dynamics of the signal that is not captured by the classical time series statistics. We show that combining different feature types and using SVMs as the machine learning approach can provide robust and efficient signal classifiers that operate well in realistic environments, and also do not suffer significantly from pollution by modulated signals.

The rest of this paper is structured as follows. In Section II we provide a brief introduction to signal classification using support vector machines, as well as discuss the selection of features used for the classifier. In Section III we describe the evaluation scenarios as well as the experiments conducted using software defined radio based implementation. We then present our results in Section IV before drawing our conclusions in Section V.

## II. SVMs FOR SIGNAL CLASSIFICATION AND FEATURE SELECTION

Support vector machines are one of the most widely used machine learning techniques for solving classification problems [9]. They operate by finding an optimum separation of the feature space on which the problem is defined by hyperplanes into distinct regions, each region resulting in one particular classification decision. For example, in the signal classification case the dimensions of the feature space represent measurable characteristics of the finite signal samples based on which we seek to classify the samples into their respective modulation types. An SVM is *trained* using samples with known modulation, based on which the classification boundaries are optimized. Classical SVMs use linear decision boundaries only, but using the so-called *kernel trick* by mapping the data into a higher-dimensional base by means of suitably selected basis functions much more complex decision boundaries become possible [10].

A more formal description of SVMs can be given as follows. Our data consists of  $n$  pairs  $(\mathbf{x}_i, M_i)$ , where  $\mathbf{x}_i$  is a feature vector, and  $M_i$  is the modulation type used for the signal from which  $\mathbf{x}_i$  is extracted. In practical applications  $\mathbf{x}_i$  will include modulation-specific components which form the basis of the classification, but these are to varying extent contaminated by noise, interfering signals, and various propagation effects including multipath fading. In the binary classification case with  $M_i \in \{-1, 1\}$  the optimum linear decision boundary is then found by minimizing for  $\mathbf{w}$  subject to the constraints

$$M_i(\mathbf{w} \cdot \mathbf{x}_i - b) \geq 1, \quad \forall i = 1, \dots, n \quad (1)$$

based on our data. Once  $\mathbf{w}$  is found, future samples can be classified by the rule  $\mathbf{x} \mapsto \text{sign}(\mathbf{w} \cdot \mathbf{x} - b)$ . For obtaining non-linear classification boundaries as discussed above, only change needed is to replace the dot products by a (non-linear) *kernel function*  $K(\mathbf{x}_i, \mathbf{x}_j)$ .

We shall now introduce the different features used in the following for signal classification with SVMs. Our baseline will be formed classical higher order statistical (HOS) features used for much of the earlier work in the literature [4]–[8]. In particular, we shall use the standard deviations of the marginal distributions of amplitude, phase, and frequency, as well as different statistics of the real and imaginary parts of the complex envelope of the signal treated as separate time series. For the latter, we focus on the cumulants of second, third and fourth order for these time series, as well as the different cross-cumulants of up to fourth order (see, for example, [11] for definitions). Cumulants are very powerful features for signal classification since they are by definition additive, and the higher order cumulants vanish for Gaussian processes, in particular receiver noise. Therefore especially the third and fourth order cumulants and cross-cumulants can be used as non-linear characterizations of the signals to be classified.

In addition to classical statistical features, we shall study the performance of *multiscale sample entropy* (MSE) for signal classification. The idea behind MSE is to estimate the entropy

rates of the signal in different timescales, characterizing how predictable the future values of the signal are given a sequence of measured samples. The MSE statistics are based on the concept of the *approximate entropy* by Pincus [12], defined as follows. Let  $u(t)$  denote the time series we seek to characterize (in our case samples of the real and imaginary parts of the signal envelope). For a positive constant  $m \in \mathbb{N}$  we define the vectors  $\mathbf{x}(t) \equiv (u(t), \dots, u(t+m-1))$  each consisting of  $m$  consecutive values from our original time series. For any two of these vectors we define then a distance  $d(\mathbf{x}(s), \mathbf{x}(t))$ , in the following taken to be the  $L_0$  norm of the difference of these vectors. Based on these quantities we then define the empirical probabilities

$$C_s^m(r) \equiv \frac{1}{N-m+1} \#\{t \mid d(\mathbf{x}(s), \mathbf{x}(t)) \leq r\}, \quad (2)$$

where  $N$  is the number of available samples, and  $r \geq 0$  specifies the tolerance of similarity. From these we can form the respective information contents by

$$\Phi^m(r) \equiv \frac{1}{N-m+1} \sum_{s=1}^{N-m+1} \ln C_s^m(r), \quad (3)$$

leading to the definition of *approximate entropy* as the entropy rate

$$\text{ApEn}(m, r) \equiv \lim_{N \rightarrow \infty} (\Phi^m(r) - \Phi^{m+1}(r)). \quad (4)$$

In practice the associated statistics

$$\text{ApEn}(m, r, N) \equiv \Phi^m(r) - \Phi^{m+1}(r) \quad (5)$$

are used instead as the amount of data available is always finite. Note that approximate entropy  $\text{ApEn}(m, r)$  measures the mean information content of the conditional probability for two subsequences of  $u(t)$  that are similar at  $m$  points continue to be similar for  $m+1$  points. Richman and Moorman showed in [13] that  $\text{ApEn}$  is a biased estimator of the entropy rate for small data quantities, and have proposed an alternative, called the *sample entropy*, and typically denoted  $\text{SampEn}(m, r, N)$ . We shall use this bias-corrected estimator in the following.

Varying  $m$  in the above definitions allows characterization of entropy rates at different time scales. Unfortunately large values of  $m$  result in exponentially increasing amount of data being needed for reliable entropy estimates to be formed. Much more statistically stable alternative is the multiscale entropy analysis of Costa *et al.* [14], [15]. This is based on the study of the *coarse-grained* process  $u^{(\tau)}(t)$  defined by

$$u^{(\tau)}(t) \equiv \frac{1}{\tau} \sum_{i=\tau(t-1)+1}^{\tau t} u(i). \quad (6)$$

Computing the sample entropy for  $u^{(\tau)}(t)$  with different values of the *scale factor*  $\tau$  will be called the multiscale sample entropy (MSE) family of statistics in the following.

### III. EVALUATION SCENARIOS

For evaluating the impact of feature selection on classification performance, we have conducted both extensive simulations as well as run emulation experiments using a channel emulator and SDR boards for feature extraction. The set of signals used was chosen to be the same as in [4]. In particular, six digital and two analog modulation schemes were used in the simulations, namely BPSK, QPSK, GMSK, 16QAM, 64QAM, AM, and FM. All of these were known by the classifier in the experiments. Further, 8PSK was used as an additional digital modulation scheme for an unknown signal in selected experiments. Different parameters for the generated signals were chosen to match various present-day technologies such as GSM, WCDMA, LTE, Bluetooth and Wi-Fi. Signal powers were selected to result in SINRs ranging from -5 dB to 30 dB. Depending on the experiments, line-of-sight and Rayleigh channel models were employed in the simulations, together with superimposed AWGN and  $1/f$  noises.

For the channel emulator experiments the feature extraction was implemented using a WARP software defined radio board [16], with signals generated using Agilent E4438C vector signal generator. The parameters of the generated signals were carefully chosen to match the simulations. EB Propsim C8 [17] was used as the channel emulator, configured to use COST 259 and 3GPP TR 25.943 UMTS evaluation models. These allow the emulation of a wide variety of environments, such as different urban, rural and hilly terrain settings. Additionally, the relative speeds of the transmitter and receiver can be emulated, allowing the study of the effects of mobility on classification performance. In addition to the use of the channel emulator, more limited validation experiments were conducted with over-the-air transmissions over the 2.4 GHz ISM band. For these the signal generator was again used to generate signals with desired modulation types, but this time feeding those signals into an arbitrary waveform generator to upconvert them to 2433 MHz central frequency after which they were fed into an antenna. The WARP SDR board connected to a receiving antenna was then placed a short distance apart in a typical office building, and the received samples were recorded using a modified WARPLab reference implementation. These samples were then subsequently downloaded to a computer, where the same SVM codebase was used for classification as for the simulated experiments.

### IV. RESULTS

We begin by presenting results from the simulation studies. Figure 1 illustrates the results from the MSE analysis for the different modulation types and scale factors in a scenario with SNR of 0 dB using the signal amplitude only. We see that the MSE curves for different modulation types are fairly distinct despite the low values of SNR, indicating that the use of entropy based features should result in good classification performance. In order to keep the number of considered features low, MSE values at 3–4 different scale factor values were used in the following results depending on the scenario.

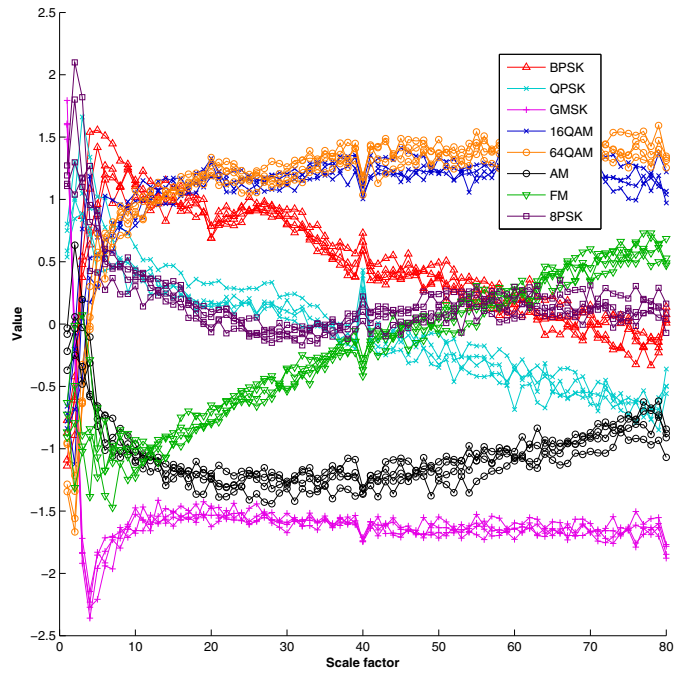


Fig. 1. MSE features of amplitude for different modulation types.

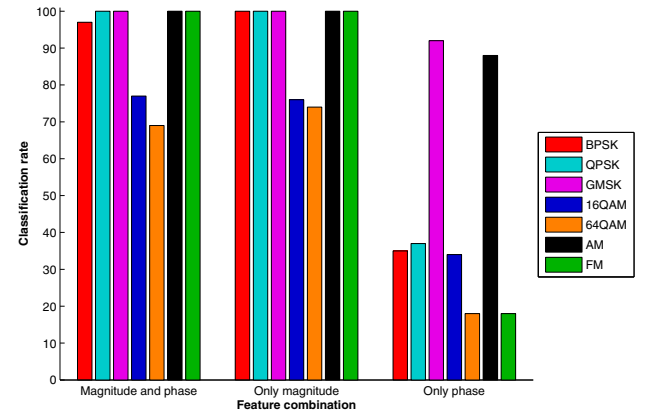


Fig. 2. Magnitude/phase vs. MSE classification rate.

There values covered the ends of the shown MSE curves, as well as 1–2 values from the middle. Figure 2 shows the resulting classification performance for different feature combinations, also assuming and SNR of 0 dB. We see that using phase information alone is largely uninformative, while magnitude only as well as magnitude and phase based MSE analysis results in good classification performance, with the exception of confounding between the 16QAM and 64QAM.

Since the two QAM-modulations are not well classified by MSE features alone, we shall now explore the performance of SVM classifiers using both MSE and classical HOS features. Figure 3 shows the classification results for MSE-only, HOS-only and combined classifiers. We see that the combination of

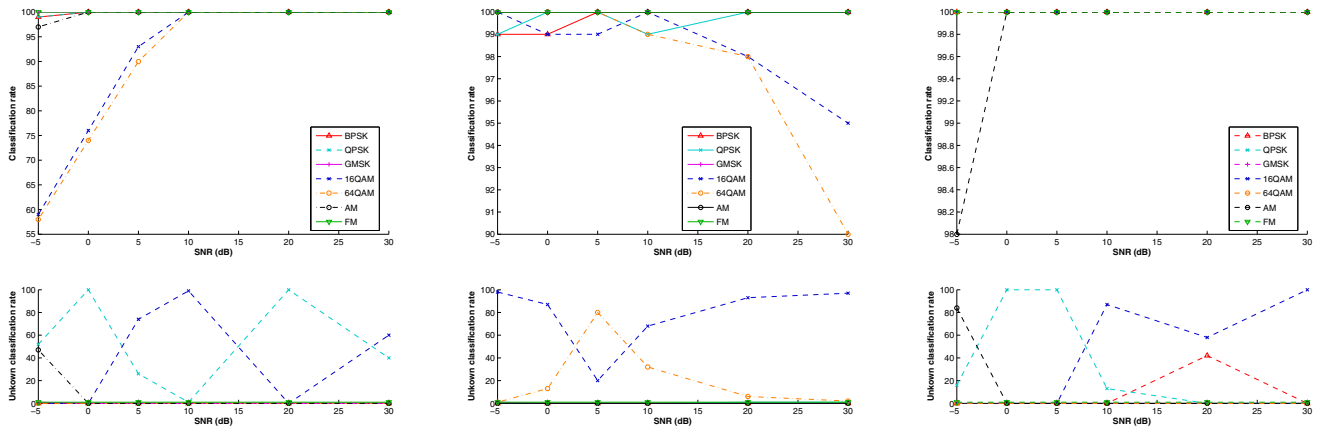


Fig. 3. Classification performance with SVMs combining traditional HOS and MSE figures for different modulation types.

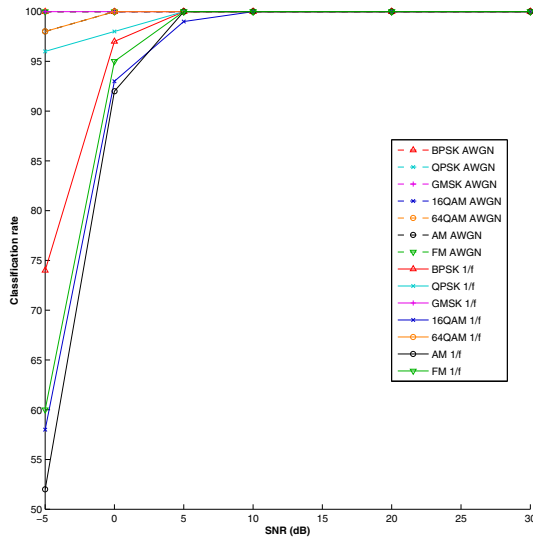


Fig. 4. Influence of noise type

the two clearly has the best performance by a large margin across all SNR values.

All of the previous results were obtained assuming AWGN contamination only. Figure 4 shows the comparison of classification results for the AWGN case against  $1/f$  noise. The latter clearly results in reduced performance in low SNR conditions, even when combined features are used. This underlines the importance of considering different noise types when evaluating the performance of signal classification solutions for low SINR conditions.

We shall now move on to the case in which the signal of interest is not only suffering contamination by noise, but also interference. This is particularly common in DSA and CR based application scenarios due to the coexistence of several signal sources on a given frequency band. Figures 5, 6 and 7 illustrate the resulting classification performance as a function

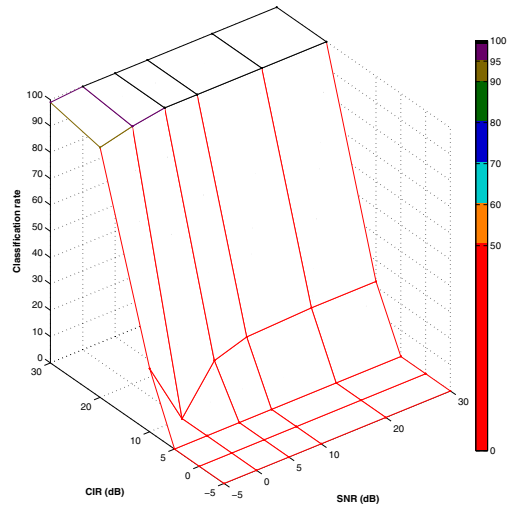


Fig. 5. Confusion plot for QPSK modulation.

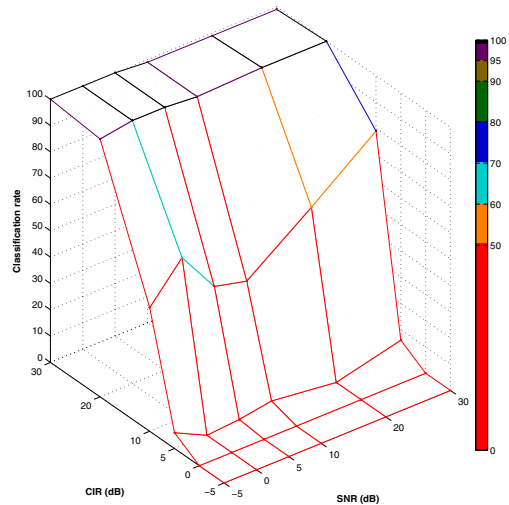


Fig. 6. Confusion plot 16QAM modulation.

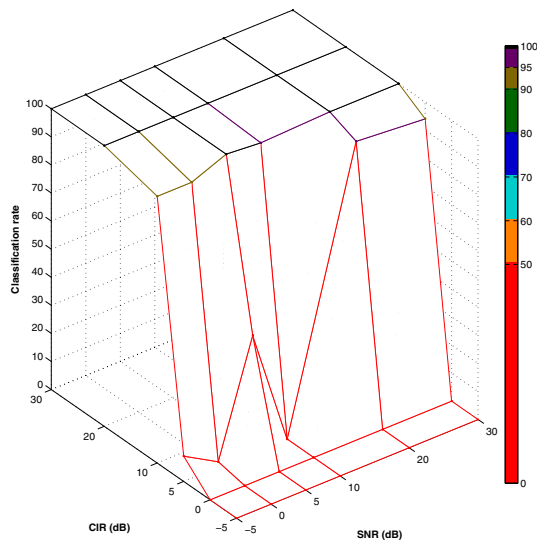


Fig. 7. Confusion plot for 64QAM modulation with AM interference.

of the CIR and SNR values of the scenario for different example combinations between the modulation types of the signal of interest and the interferer. In the first two figures both of these have same modulation type, whereas in the third case the modulation types differ. From the results we see that the most challenging case for the SVM classifier arises when modulation types are the same, in which case 10–20 dB CIR is needed for accurate classification. The classification performance remains good also in the presence of very strong noise. For mixed modulation types accurate classification can on the other hand already be achieved with 5–10 dB CIR.

Let us now move from simulation results to the WARP implementation and channel emulation experiments. Figure 8 shows the classification results for the over the air validation experiments for both MSE and HOS based SVMs. The results show that first of all the implementation in the actual testbed works well, with MSE based classification achieving practically perfect results, with HOS based classifier also performing well. Following these first tests we commenced an extensive performance evaluation of the different classifiers using the PropSim channel emulator. Figure 9 illustrates the impact of the vehicular mobility on the classification results for different feature selections. Again the hybrid features combining both MSE and HOS components perform the best, with MSE-only feature set outperforming the HOS-only feature set. Finally, Figure 10 shows the aggregate classification results for different feature combinations and channel models obtained using the channel emulator. The relative performance of MSE and HOS based classifiers clearly depends on the scenario, but the hybrid classifier in general performs very well. The only exception is the case of urban channel model with slow mobility, in which case the HOS classifier has clear edge in performance.

## V. CONCLUSIONS

In this paper we studied the performance of support vector machine based signal classification in realistic wireless environments. Our focus was in particular on impact of interference from modulated signals and influence of realistic wireless channel conditions on classification performance. We also proposed additional time series statistics originating from complex systems research to be used as features for the classifier, complementing the classical second and higher order statistics previously employed in the literature. Our results based on both extensive numerical simulations as well as experiments using software defined radio boards and state of the art channel emulator show that the extended feature set results in robust classification performance in wide variety of channel conditions, and also when significant modulated interference is present in addition to Gaussian noise.

## ACKNOWLEDGEMENT

We thank RWTH Aachen University and the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) for providing financial support through the UMIC research centre.

## REFERENCES

- [1] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5g wireless communication networks," *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 122–130, 2014.
- [2] S.-Y. Lien, K.-C. Chen, Y.-C. Liang, and Y. Lin, "Cognitive radio resource management for future cellular networks," *Wireless Communications, IEEE*, vol. 21, no. 1, pp. 70–79, 2014.
- [3] J. Van De Beek, T. Cai, S. Grimoud, B. Sayrac, P. Mähönen, J. Nasreddine, and J. Riihijärvi, "How a layered rem architecture brings cognition to today's mobile networks," *Wireless Communications, IEEE*, vol. 19, no. 4, pp. 17–24, 2012.
- [4] M. Petrova, P. Mahonen, and A. Osuna, "Multi-class classification of analog and digital signals in cognitive radios using support vector machines," in *Wireless Communication Systems (ISWCS), 2010 7th International Symposium on*. IEEE, 2010, pp. 986–990.
- [5] R. Kannan and S. Ravi, "Second-order statistical approach for digital modulation scheme classification in cognitive radio using support vector machine and k-nearest neighbor classifier," *Journal of Computer Science*, vol. 9, no. 2, p. 235, 2013.
- [6] C. Cardoso, A. Castro, and A. Klautau, "An efficient fpga ip core for automatic modulation classification," *Embedded Systems Letters, IEEE*, vol. 5, no. 3, pp. 42–45, Sept 2013.
- [7] A. Tsakmalis, S. Chatzinotas, and B. Ottersten, "Modulation and coding classification for adaptive power control in 5g cognitive communications," in *Signal Processing Advances in Wireless Communications (SPAWC), 2014 IEEE 15th International Workshop on*, June 2014, pp. 234–238.
- [8] R. Kannan and S. Ravi, "Classification of digital signals in cognitive radio based on second-order statistical approach," in *Emerging Trends in Electrical Engineering and Energy Management (ICETEEEM), 2012 International Conference on*, Dec 2012, pp. 166–170.
- [9] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.
- [10] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*. Springer series in statistics Springer, Berlin, 2001, vol. 1.
- [11] D. R. Brillinger, *Time series: data analysis and theory*. Siam, 1981, vol. 36.
- [12] S. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 6, pp. 2297–2301, Mar. 1991.

Classified modulation	BPSK	100 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 0.0%
	QPSK	0 0.0%	100 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	31 3.9%	76.3% 23.7%	
	GMSK	0 0.0%	0 0.0%	100 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	
	16QAM	0 0.0%	0 0.0%	0 0.0%	100 12.5%	0 0.0%	0 0.0%	69 8.6%	59.2% 40.8%	
	64QAM	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 12.5%	0 0.0%	0 0.0%	100% 0.0%	
	AM	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 12.5%	0 0.0%	100% 0.0%	
	FM	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 12.5%	0 0.0%	100% 0.0%
	8PSK	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%	
		100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	0.0% 12.5%	87.5% 15.1%
	BPSK	QPSK	GMSK	16QAM	64QAM	AM	FM	8PSK		
	Actual modulation									

(a) Using MSE.

Classified modulation	BPSK	100 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 0.0%
	QPSK	0 0.0%	98 12.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 12.5%	49.5% 50.5%
	GMSK	0 0.0%	0 0.0%	100 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	
	16QAM	0 0.0%	1 0.1%	0 0.0%	93 11.6%	12 1.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	87.7% 12.3%
	64QAM	0 0.0%	1 0.1%	0 0.0%	7 0.9%	88 11.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	91.7% 8.3%
	AM	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 12.5%	0 0.0%	0 0.0%	100% 0.0%	
	FM	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100 12.5%	0 0.0%	100% 0.0%	
	8PSK	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%	
		100% 0.0%	98.0% 2.0%	100% 0.0%	93.0% 7.0%	88.0% 12.0%	100% 0.0%	100% 0.0%	0.0% 12.5%	100% 15.1%	84.9%
	BPSK	QPSK	GMSK	16QAM	64QAM	AM	FM	8PSK			
	Actual modulation										

(b) Using HOS.

Fig. 8. Validation results for the classifier and testbed implementation using over the air transmission with WARP boards.

Classified modulation	QPSK	73 18.2%	0 0.0%	16 4.0%	36 9.0%	58.4% 41.6%
	GMSK	4 1.0%	100 25.0%	0 0.0%	0 0.0%	96.2% 3.8%
	16QAM	13 3.2%	0 0.0%	59 14.8%	43 10.8%	51.3% 48.7%
	64QAM	10 2.5%	0 0.0%	25 6.2%	21 5.2%	37.5% 62.5%
		73.0% 27.0%	100% 0.0%	59.0% 41.0%	21.0% 79.0%	63.2% 36.8%
	QPSK	GMSK	16QAM	64QAM		
	Actual modulation					

(a) For MSE features.

Classified modulation	QPSK	45 11.2%	4 1.0%	27 6.8%	10 2.5%	52.3% 47.7%
	GMSK	13 3.2%	85 21.2%	3 0.8%	11 2.8%	75.9% 24.1%
	16QAM	16 4.0%	3 0.8%	49 12.2%	23 5.8%	53.8% 46.2%
	64QAM	26 6.5%	8 2.0%	21 5.2%	56 14.0%	50.5% 49.5%
		45.0% 55.0%	85.0% 15.0%	49.0% 51.0%	56.0% 44.0%	58.8% 41.2%
	QPSK	GMSK	16QAM	64QAM		
	Actual modulation					

(b) For HOS features.

Classified modulation	QPSK	64 16.0%	0 0.0%	9 2.2%	19 4.8%	69.6% 30.4%
	GMSK	3 0.8%	100 25.0%	0 0.0%	0 0.0%	97.1% 2.9%
	16QAM	10 2.5%	0 0.0%	60 15.0%	27 6.8%	61.9% 38.1%
	64QAM	23 5.8%	0 0.0%	31 7.8%	54 13.5%	50.0% 50.0%
		64.0% 36.0%	100% 0.0%	60.0% 40.0%	54.0% 46.0%	69.5% 30.5%
	QPSK	GMSK	16QAM	64QAM		
	Actual modulation					

(c) For hybrid features.

Fig. 9. Classification performance in the vehicular mobility related channel model for different feature choices.

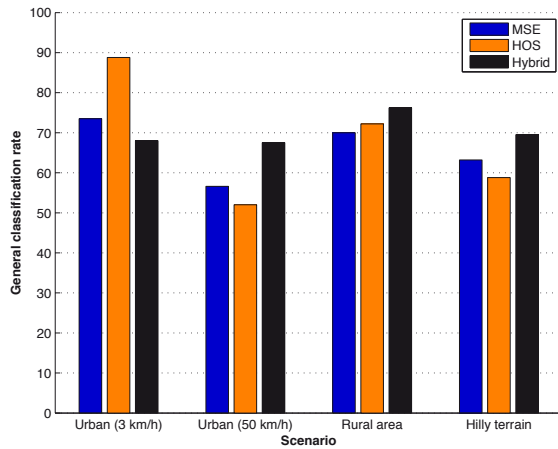


Fig. 10. Impact of channel and feature selection on classification performance in the channel emulation experiments.

- [13] J. Richman and J. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology- Heart and Circulatory Physiology*, vol. 278, no. 6, pp. 2039–2049, 2000.
- [14] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale Entropy Analysis of Complex Physiologic Time Series," *Physical Review Letters*, vol. 89, no. 6, p. 068102, Aug. 2002.
- [15] —, "Multiscale entropy analysis of biological signals," *Physical Review E*, vol. 71, no. 2, p. 021906, Feb. 2005.
- [16] A. Khattab, J. Camp, C. Hunter, P. Murphy, A. Sabharwal, and E. W. Knightly, "Warp: a flexible platform for clean-slate wireless medium access protocol design," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 12, no. 1, pp. 56–58, 2008.
- [17] Elektrobit Multi-Channel Emulator, "Propsim C8," 2008.