Click here to view PowerPoint presentation in PDF format

Classification of Communication Signals and Detection of Unknown Formats Using Artificial Neural Networks

Alexander Iversen, Nicholas K. Taylor and Keith E. Brown
Intelligent Systems Laboratory
Heriot-Watt University
Edinburgh EH14 4AS
UK

{alex, nkt}@macs.hw.ac.uk, k.e.brown@hw.ac.uk

Jørn Kårstad
Information Management Division
Norwegian Defence Research Establishment
N-2027 Kjeller
NORWAY

jorn.karstad@ffi.no

ABSTRACT

An important element in many wireless communication activities is distinguishing between different radio signals. In this paper, we address some important problems within radio communication signal classification, one of which is the detection of unknown signal formats. To tackle some of these problems, we propose a combined classifier, consisting of two different neural network types, and evaluate its performance on a variety of semi-realistic radio communication signals. Experimental results indicate that the proposed classifier can exploit the individual strengths of the neural networks and achieve both good discrimination between known, and reliable detection of unknown signal formats. We also argue that combining classifiers may be beneficial in terms of adapting to changing requirements.

1.0 INTRODUCTION

Today's military operations depend on extensive use of wireless communication technologies. Effective communication is a necessity in network-centric operations but hostile monitoring of radio transmissions may also be utilised to the benefit of any opponents. In a military setting an important objective is to establish an overview of the situation. Monitoring of radio signals may reveal vital information regarding the detection, localisation and identification of an opponent. Traditionally, operators who were trained to recognise various signal formats based on manual 'listen in' techniques performed the identification. Conventional analogue FM radios will still be around for a number of years but the increasing use of various digital radio systems calls for a different approach to classification of radio communication signals. The time aspect of the signal analysis is crucial. Also, the extensive use of digital encryption effectively hinders signal identification based on the information content. This requires automatic signal classification based on technical measurements rather than manual user intervention.

The term 'signal classification' implies that there exists some *a priori* knowledge about the various types of communication signals that are to be analysed. Usually this may be true but modern radio environments are more diverse as the radio systems are capable of adapting their characteristics to changing traffic

Iversen, A.; Kårstad, J.; Taylor, N.K.; Brown, K.E. (2006) Classification of Communication Signals and Detection of Unknown Formats Using Artificial Neural Networks. In *Military Communications* (pp. P6-1 – P6-20). Meeting Proceedings RTO-MP-IST-054, Poster 6. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

requirements or radio propagation channel effects. In a non-co-operative environment one must also take into account that unknown signal formats may turn up, perhaps introduced by a counterpart in order to confuse or deceive. Thus, in a military setting the ability to detect unknown signal formats will be highly relevant.

This paper gives a short overview of the main principles of pattern recognition and how they can be utilised for automatic classification of radio communication signal formats. The issue of detecting unknown signal formats is specifically addressed. Numerous comparative studies have indicated that artificial neural network approach is promising in terms of classification performance, speed of execution and robustness to noise and propagation channel degradation. A novel combined neural network classifier is proposed. The performance of this classifier is evaluated on sampled radio-signal sequences.

2.0 ISSUES ON SIGNAL MODULATION CLASSIFICATION

This paper is based on the issues raised by the use of communication signal classification in a military context. However, the ability to recognise various radio communication signal formats is of interest in several fields within the civilian sector as well. For instance, the introduction of Software Defined Radios may allow for radio equipment that is able to adapt the radio interface according to changing traffic requirements and radio environments. The reconfiguration of the radios may in principle be a result of manual user intervention. A more sophisticated approach would be to enable the radio terminals to be auto-adaptive, i.e., the terminals are able to autonomously reconfigure their radio interface [1]. This concept will require the introduction of some degree of 'situation awareness' in the receivers. Characterisation of the radio environment in terms of recognising available radio signal formats is one important part of such 'situation awareness'. Another example of non-military applications is governmental bodies, like the telecommunication authorities, whose task is to coordinate the use of the radio spectrum on national level. Methods for automatic classification of radio communication signals will be of great benefit for their surveillance of the radio spectrum. Depending on the application the signal classifier will have to meet different requirements with respect to performance, complexity, cost and speed of operation. Nevertheless, most of the basic principles may be identical.

In a military context the recognition of communication is of interest as input to the ongoing situation evaluation. By surveillance of the electromagnetic spectrum one may deduct vital information on the activities in the field. Timely access to such information is imperative, which implies that the recognition process must be automated yet maintaining the possibility of manual analysis of new or unknown signal formats. We can thus view a communication signal classifier system as a sensor aimed at providing information about which communication systems or units are active within a certain area. Depending on the application one may want to emphasise the issues of error control and classification accuracy differently:

- A surveillance system should ideally be more sensitive towards detecting hostile communication activity rather than own forces communication activity. Unknown communications should be detected and 'tagged' for further off-line analysis.
- Full signal identification may not necessarily be required if the information from the signal classifier is used to select appropriate EW countermeasures. A rather rough classification will probably be sufficient to find a proper jamming signal format.
- If the classifier is used for controlling a software radio the capability of detecting unknown formats is of minor interest. Such a classifier should ideally be able to distinguish among a set of known signal formats with good accuracy.

The proposed combined signal classifier scheme will allow for introduction of mechanisms for error control that can be tailored to the application.

3.0 APPROACHES TO MODULATION CLASSIFICATION

A vast range of modulation recognition methods has been proposed in the research literature. Much of these are based on standard pattern classification approaches, which we describe in Section 3.1. Section 3.2 describes issues regarding pre-processing and Section 3.3 shows examples of features that can be extracted from intercepted signal segments. Section 3.4 describes approaches to automatic classification.

3.1 Pattern Classification Approach

We find computer-based pattern classification important in many areas and applications, such as fault detection, medical diagnosis, biometric identification, speech and optical character recognition and, as we will focus on, communication signal recognition. Despite this myriad of types of data (e.g., sensor data, digital images, acoustic or radio signals) there exists a generic pattern classification approach (Figure 1). The figure shows that first some form of pre-processing is performed. This might be, for example, signal or image processing operations, such as segmentation, transformation and filtering, in order to obtain a uniform representation of the data. Next is the process of reducing the dimensionality of the data by extracting a smaller number of features that emphasise the characteristics of, and distinctions between, the classes. The features then serve as inputs to the classifier, which performs the final classification task. In order for the classifier to perform classification, it must obtain information on what data represent what class. This is often referred to as *learning*. That is, based on pre-labelled examples from each class (training examples) the classifier learns to map the data to their corresponding class. It also learns to generalise, such that new and unclassified examples are classified accordingly.

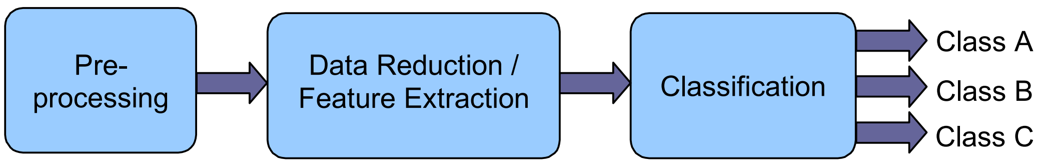

Figure 1: Generic Pattern Classification Approach

3.2 Communication Signal Pre-processing

The basic pre-processing requirements for modulation classification are to obtain a signal representation that is as consistent as possible and to reduce degradations that may confuse the classifier. Thus, the pre-processor block in Figure 1 is responsible for quality assurance and for adapting the data input into a suitable form before feeding it to the feature extractor block. A fundamental restriction in the pre-processing of unknown radio signals is that both the signal format and the transmission channel are unknown to the intercept receiver. The pre-processor may nevertheless be able to estimate some of the main signal parameters like carrier frequency and bandwidth that is required for adapting the signal processing to the intercepted signal. Typically, at the pre-processor stage of a radio signal classifier system the following issues are addressed:

- Selection of time segments that contain signals of potential interest to the classifier. This will, for instance, include removal of segments that are heavily degraded by noise.
- Fading signal channels will introduce amplitude variations that may be interpreted as amplitude modulation by the classifier. For example, attenuation from hydrometeors will cause slow amplitude variations, whereas for mobile communication channels the received signal may exhibit fast signal level variations. Such unwanted signal variations must be compensated for.
- Signal matched filtering will improve the signal/noise ratio. Estimating the optimal bandwidth of the post-detection filter will thus be a very important part of the pre-processor stage.

- Carrier and symbol synchronisation may be required for classifier algorithms that are based on coherent detection. The synchronisation should, in these cases, be included in the pre-processing stage.
- Interference from unwanted signals may disturb the reception of the target signal. Such interference should ideally be suppressed as far as possible. This may be achieved by appropriate filtering in the frequency domain by gating out strong interference events in the time domain, or, if feasible, by the use of directive antennas.
- Multi-path transmission channels will introduce propagation degradations that may confuse the classifier. To some extent, various equalisation techniques may suppress the effect of the degradation but equalisation of unknown signal formats is a non-trivial task.

3.3 Feature Extraction

A communication signal represented in the time domain contains vast redundancy and contains, without further analysis, little information that can aid the discrimination of modulation formats. This necessitates some form of feature extraction. A natural choice of dimensionality reduction is to represent a time-domain signal in the frequency domain through some form of spectral analysis. For example, periodograms, Welch periodogram and bispectrum techniques have been proposed. The spectrum samples can then be used as features either directly [2] or through the extraction of some statistical properties [3]. Statistical features - such as the mean, standard deviation and higher order statistical moments - can also be directly extracted in the time domain, i.e., from the instantaneous amplitude (envelope), phase and frequency of the signal [3]. Features extracted from the time domain are less computationally intensive and have shown to be effective for many modulation classification tasks [4]. An alternative to the above feature extraction approaches is to obtain the signal's constellation shape and then perform some form of pattern matching in order to determine the modulation format [5].

Whereas some of the feature extraction methods proposed in the literature assume full *a priori* knowledge about the communication signal, i.e., frequency, bandwidth, symbol rate and synchronisation, other features can be extracted with little or no *a priori* knowledge. Choosing an appropriate set of features thus depends on both the type of modulation formats that the classifier is being trained to classify and in which scenarios (co-operative or non-co-operative) the system is intended to operate.

3.4 Classification Methods

Having obtained a feature vector, a classifier must then use that to output a class label or, alternatively, a set of probabilities or confidence levels that indicate the prediction results. One way of categorising different types of classifiers is into statistical, decision theoretic, fuzzy, or neural network-based approaches. There is no classifier type that is superior to the others for all types of problems. Thus, in the research literature on signal classification, all approaches are represented.

The different statistical approaches have in common that they seek to create distribution models for each class based on the training data. The predicted class can then be found by selecting the highest posterior probability when evaluating the input on these distributions. The probability distributions are based on either *parametric* or *non-parametric* techniques [6]. The former assumes a known distribution model (e.g., Gaussian), and is concerned with finding suitable parameter values, based on the training data, in order to approximate the true distribution. The latter technique is concerned with creating distribution models solely based on the training data and is not constrained by the standard distributions. It can thus offer a more representative distribution model on the expense of more parameter setting. Both parametric (e.g., [7]) and non-parametric (e.g., [5]) techniques have been proposed for modulation classification.

For decision theoretic classification approaches, modulation formats are determined by traversing a decision tree where the features are tested against thresholds at the tree nodes until reaching an end-node, which indicates a format. A decision tree requires few resources and executes very quickly. It is therefore suitable for online classification and for implementation in resource-limited systems. Due to its simplicity and good classification abilities, the decision theoretic approach has been popular for modulation classification [3], [8]. Recently, fuzzy logic-based modulation classifiers have also been proposed. These can, contrary to the decision tree, make 'soft' decisions. I.e., provide varying degrees of 'certainty' to the modulation formats [9].

An approach that has proved to be very effective for modulation classification, is that based on artificial neural networks (neural networks for short). These are loosely based on the operation of the brain, and have been applied to a wide range of engineering applications throughout the last two and a half decades. It is their fast execution (once trained) and robustness to noise that have made them popular for modulation classification. Comparative studies have also shown that neural networks outperform statistical and decision tree-based modulation classifiers [3], [10].

Rather unfortunately, much of the existing work on modulation classification has been overly focused on reporting classification success rates that exceed that of others by tweaking classifier parameters. What we believe is lost in this 'race' is the realisation that many of these results are merely hypothetical as they are based on ideal computer-generated signals simulated in ideal transmission channel models. In fact, some of the problems discussed in Section 2 would in many cases swamp these insignificant differences in classification performance. In fairness, some have rightly addressed some of these challenges. For example, Liedtke [4] highlights the importance that one cannot assume full *a priori* knowledge of signals in a non-co-operative setting and that in these situations, the procedure should be able to adapt to handle new signal formats. Kim *et al.* [10] implement and evaluate classifiers on a digital signal processor and also test the classifiers on real off-air signals. Furthermore, Hatzichristos [11] and Venalainen *et al*. [12] evaluate modulation classification in a multi-path environment. In this paper, we will look at another challenge, namely that of detecting unknown formats in addition to handling the classification between known formats. Due to previously reported benefits of neural network-based signal classification we too focus on neural networks.

4.0 ARTIFICIAL NEURAL NETWORKS

The general function of a neural network is to produce an output pattern when given a particular input pattern, and is loosely related to the way the brain operates. Learning these mappings is done in conceptually the same way as the brain. That is by generalising from a number of examples. Neural networks consist of a number of fairly simple computational devices that resemble the brain's neurons, interconnected with weighted connections that resemble dendrites and axons. Several types of neural networks exist but the most common one used for modulation classification has been the Multi-Layer Perceptron.

4.1 Multi-Layer Perceptron

A Multi-Layer Perceptron (MLP) is a network of non-linear Perceptrons. A Perceptron has n inputs and n corresponding weights, and first calculates the weighted sum of inputs. This sum is then input to a non-linear activation function that produces the Perceptron's output response; typically a value between –1 or 0 and +1. In a MLP, the Perceptrons are organised in layers, such as illustrated in Figure 2.

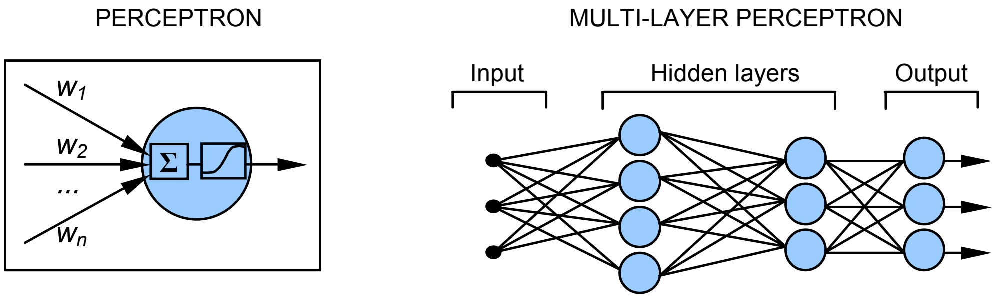

Figure 2: Example of a Perceptron and a 2-hidden-layer MLP with 3 inputs and 3 outputs

The MLP has one input layer, one or more hidden layers and one output layer. When the network is operating, the data is propagated through the network in a forward direction layer by layer. For classification tasks, the input layer usually consists of the feature vector (c.f. Section 3). The hidden and output layers map the data from feature space to an output space that represents the predicted classification. Usually in a C-class classification task there are C outputs, each of which represents a class. The network is then trained to produce a high (active) value on the output that corresponds to the class, whereas the other outputs are low (inactive). This is called 1-of-C representation.

MLPs are commonly trained using the *back-propagation learning algorithm*. This requires a set of training examples, i.e., pre-labelled examples from each of the classes that we want to classify. Essentially, the algorithm consists of two passes through the different layers of the network: forward and backward. When training data are presented to the input layer, they are propagated through the network with fixed (initially random) weights on the connections. This is the forward pass. When this is completed, an error is computed, which is effectively the difference between the actual and desired outputs. This error is then propagated back through the network from the output layer to the input layer. During this back-propagation the weights on the connections are adjusted according to the error, in order to lower the error in the next forward pass. These forward and backward passes continue until an acceptably small error is obtained.

To get an appreciation of how the MLP operates, we construct a problem classifying the five modulation formats 2- and 4-level ASK and PSK, and MSK. Assuming that signal examples can be represented in terms of two features, Figure 3 (left) graphically visualises the distribution of signal examples[1]. With these training examples, the MLP can be trained to construct decision boundaries, such that, when in operation, it can classify new examples accordingly. An essential requirement for any classifier is that, instead of correctly classifying only examples that look exactly like the training examples, it learns to generalise. Thus, the classifier should be able to correctly classify examples that are situated in principally the same area in feature space as the training examples. Figure 3 (right) visualises the decision space as created by a MLP, where each marker indicates how the MLP will classify that particular input. The white 'lines' and 'spaces' indicate that the MLP does not provide a clear classification. The lines can be interpreted as decision boundaries, at which the MLP goes from classifying one class to classifying another class. The larger white spaces can be interpreted as sections where the MLP provide no class (low values on all outputs).

[1] *Details of signal generation and feature extraction will be covered in Section 6. For now, it suffices to regard the different signal examples as points in the two-dimensional feature space.*

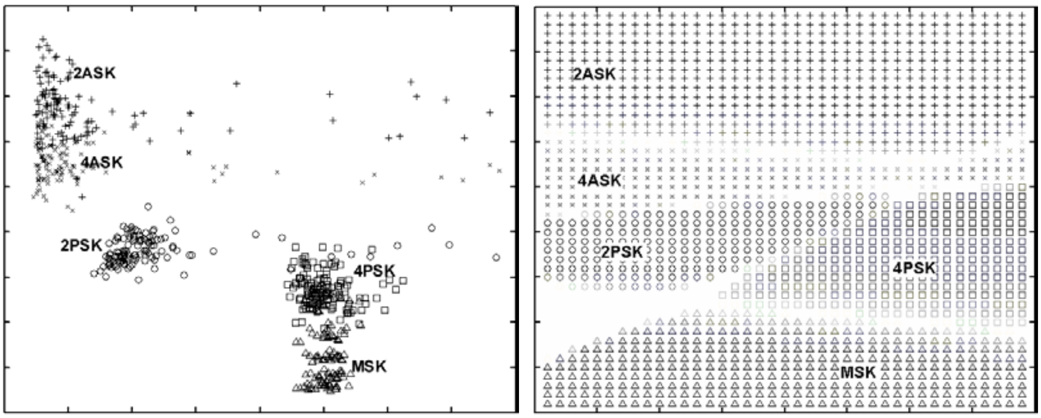

Figure 3: Left hand side: Training examples represented in 2D feature space. 2ASK signals are represented as plusses, 4ASK as crosses, 2PSK as circles, 4PSK as squares and MSK as triangles. Right hand side: Decision space as created by a MLP.

The figure illustrates a couple of important points. First, we can see that the decision boundaries are generalised well, and that new examples of the five modulation formats are likely to be correctly classified. Secondly, we note that the MLP's decision regions do not reflect the distribution of training data but rather seem to be unbounded. That is, new inputs that lie far from any training data cluster may still be classified as being one of the modulation formats. For example, an input example that lies in the lower left corner of the graph will be classified as MSK even though this example looks nothing like the MSK training examples. The significance of this weakness is highlighted in Figure 4, where we, in addition to the training data, plot examples of 2- and 4-level FSK and 16-level QAM.

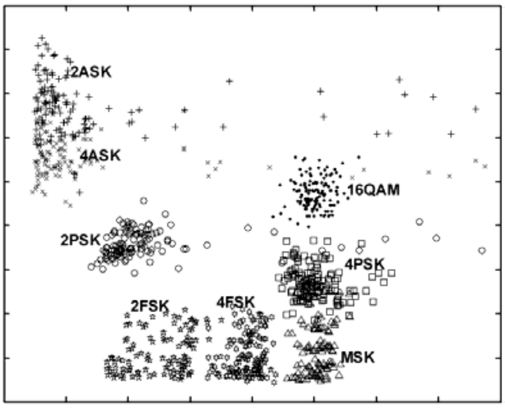

Figure 4: Training examples together with 'unknown' signal types 2FSK, 4FSK and 16QAM

Now we see that, in the event of receiving signal types that are not in the training set, the MLP is incapable of detecting the presence of these. In fact, the MLP will misclassify both 2FSK and 4FSK signals as MSK and will also misclassify 16QAM as either 2PSK or 4ASK. In case no other actions are taken *a priori* to detect the presence of unknown signal types, this can limit the usage of the MLP to co-operative modulation classification only. To address this problem we will, in the next section, look at an adaptation of the MLP.

4.2 Auto-association Neural Network

The MLP has, in addition to classification, typically been applied to problems such as function approximation and dimension reduction, e.g., non-linear principal component analysis [13]. The latter can be achieved using the MLP for auto-association. Auto-association implies that the input data are reconstructed at the output and thus that the number of output nodes must be the same as the number of inputs. Also, in order to perform dimension reduction, a hidden layer must consist of fewer nodes than the number of nodes in the input/output layers. If the network is successful in reconstructing the data, the activations at the hidden layer thus represents a compressed representation of the data. Figure 5 shows an example of what we call an auto-association neural network (AANN).

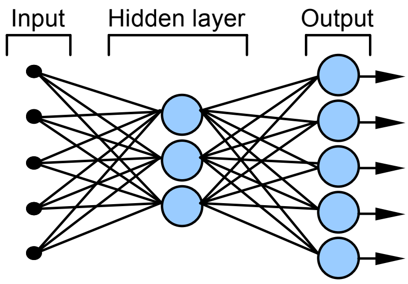

Figure 5: Example of an auto-association neural network (AANN) with 5 inputs and 3-node hidden layer.

The last decade has also seen an increased use of the AANN for detection and classification tasks. It has been shown that by training the network to recognise data of a particular type, it will reconstruct similar data with only a small error margin. If atypical data are input, however, the reconstruction error will be large. By thresholding the reconstruction error the network can be used to detect anything that deviates from the 'norm'. This is often referred to as *novelty detection* [14] and has the benefit that it only requires the 'normal' data for training in order to detect anything that deviates from this norm.

For modulation classification we see the immediate benefit of the AANN as it can detect unknown signal types of which we have no available training data. Furthermore, by creating one AANN for each of the known formats, we can not only detect unknown modulation formats but also classify the known ones. The resulting decision space obtained by using 5 AANNs for the 5-format modulation classification problem is shown in Figure 6. The decision regions for the modulation formats are now bounded around the training examples and thus anything that falls outside these sections can be detected and classified as unknown. For example, the 'unknown' 2FSK, 4FSK and 16QAM examples shown in the figure will be correctly detected. On the other hand, there are also drawbacks to using AANNs for classification. Because the AANNs are trained independently from each other, the decision regions of known formats may intersect when the classes are similar. Therefore, for inputs situated at the border between modulation formats, more than one AANN may indicate that the input belongs to their format. The AANNs may therefore be more imprecise than the MLP in discriminating between classes. For example, see how the decision boundary between 2ASK and 4ASK created by the MLP, (Figure 3) reflects the distribution of training data better than that created by the AANNs (Figure 6).

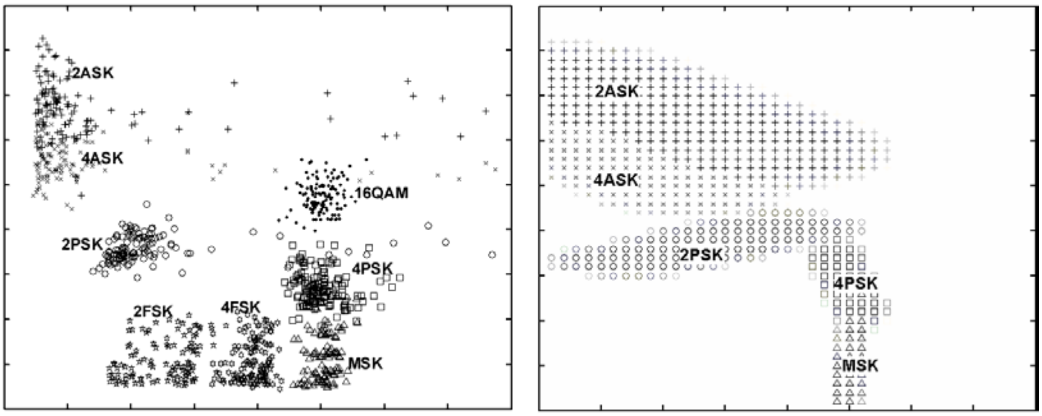

Figure 6: Left hand side: Training examples together with the 'unknown' signal types 2FSK, 4FSK and 16QAM. Right hand side: Decision space as created by 5 AANNs.

In summary, comparing the MLP results (Figure 3) with the results from the AANNs (Figure 6), we can conclude that the MLP is best suited for discriminating between modulation formats, whereas the AANNs are useful for detecting unknown formats. The next section looks at the combination of these two neural network types.

5.0 COMBINED ARTIFICIAL NEURAL NETWORK CLASSIFIER

As we have seen examples of above, stand-alone classifiers may be suitable for some types of problems but not for others. Such limitations have raised the awareness of combining classifiers for solving more complex classification problems [15]. The problem to solve may either be naturally modular, for which a combined classifier structure is obvious, or be of such size that de-composing the problem into sub-problems is required. As for military modulation recognition, there may be different requirements according to the type of scenario (e.g., surveillance, counter-measures or demodulation of friendly communication). We can break those requirements down to (a) good classification among known modulation formats and (b) reliable detection of unknown formats. To accommodate these we propose a method of combining the MLP and AANN such that we maximise the strength and minimise the weakness of each neural network type.

In order to explain the rationale behind the combination method we need to formalise the output representation of the two network types. Using the 1-of-*C* representation, we can let the *c*-th output of a MLP classifier denote a certainty factor, μ_c, where $-1 < \mu_c < 1$ and where each c represents one of C modulation formats. We let $\mu_c > 0$ indicate evidence supporting the hypothesis that the input belongs to modulation format c, and let $\mu_c < 0$ indicate evidence against the input belonging to class c. The magnitude of μ_c will indicate the degree of evidence for or against. For a C-class modulation classification problem C AANNs are required. Each AANN will, after being trained, provide a reconstruction error e_n for any given input n. By using a threshold, τ_c, we can say that $e_n < \tau_c$ indicates evidence for the hypothesis that the input, n, belongs to modulation format c. Similarly, $e_n > \tau_c$ indicates evidence against the hypothesis. To get a unified representation of the MLP and AANN outputs, we map the AANN reconstruction error, e_n, onto a certainty factor, α_c, where $-1 < \alpha_c < 1$ and where α_c can be interpreted in the same way as μ_c. This mapping function is illustrated in Figure 7 (Appendix A contains details on the extraction of the certainty factors).

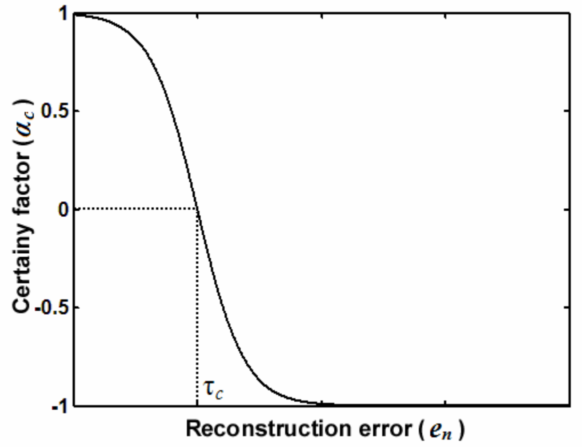

Figure 7: AANN Certainty function.

Now, for each of the C modulation formats, the combined neural network must evaluate the advice from both the MLP and AANNs, that is μ_c and α_c respectively, before making an overall decision, in the form of a *combined certainty factor,* φ_c. The decision-making is based on the following:

- A negative α_c, regardless of μ_c, indicates that the input, n, is not of modulation format c.
- A negative μ_c, regardless of α_c, indicates that the input, n, is not of modulation format c.
- A positive μ_c and a positive α_c indicate that the input, n, is of modulation format c.

Based on this reasoning we can express the combined certainty factor as $\varphi_c(\mu_c, \alpha_c) = \min(\mu_c, \alpha_c)$. The combined neural network is illustrated in Figure 8.

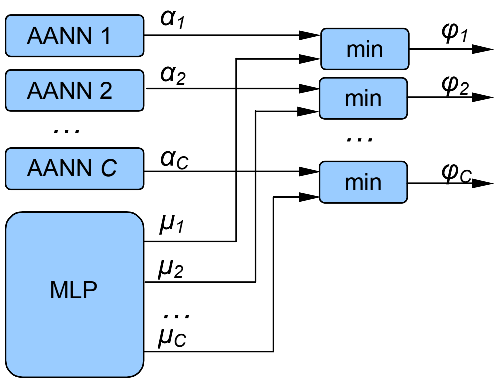

Figure 8: Combined Neural Network.

Figure 9 depicts the result from applying this combined neural network to the 5-class modulation classification problem. The figure illustrates how the combined neural network uses the MLP's decision boundaries between the classes, which ensures good discrimination of modulation formats, whereas it uses the AANNs to create bounded decision regions, which ensures detection of unknown formats.

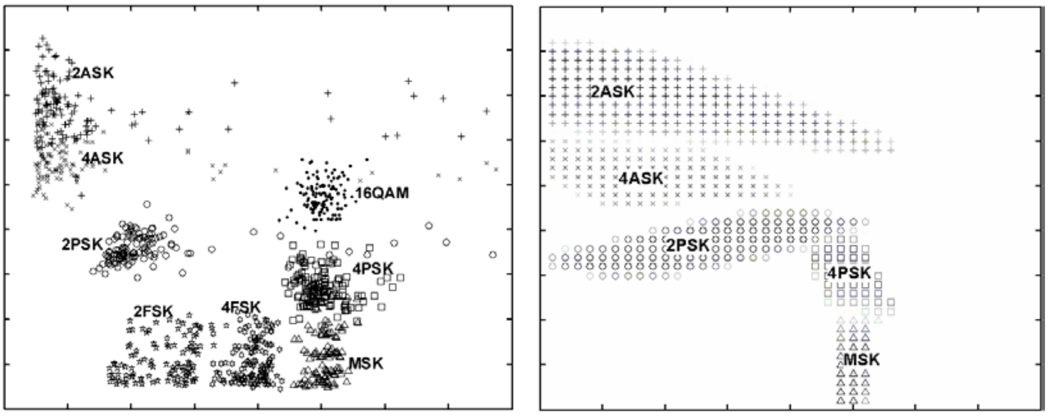

Figure 9: Left hand side: Training examples together with the 'unknown' signal types 2FSK, 4FSK and 16QAM. Right hand side: Decision space as created by the combined neural network.

The next section presents a more comprehensive experiment, which illustrates the performance of the combined neural network approach.

6.0 EXPERIMENTAL ANALYSIS

A main goal for the experiment was to test and assess the proposed classification methodology under controlled, yet realistic conditions. Section 6.1 describes the signal specifications such as generation, transmission and receiver set-up, and feature extraction. In Section 6.2, the classification results are presented and assessed.

6.1 Signal Specifications

The signals used for the training and testing of the neural network were generated directly at RF frequency using a vector signal generator. Dedicated software on an external PC provided full control over the generation of waveforms in terms of defining the information bit-streams, the modulation constellations and the signal/noise ratio. The set-up also allows for the introduction of signal degradations due to interference and multi-path propagation effects, though this was not exploited in this experiment. On the receiver side the RF signal was down-converted to a fixed IF of 21.4 MHz before digitalisation in a high-speed A/D converter. Finally the received signals were down-converted to 50 kHz and decimated digitally by a factor of 1/256 producing I- and Q-samples at a rate of ≈250 kHz. An overview of the set-up for signal generation and recording is given in Figure 10.

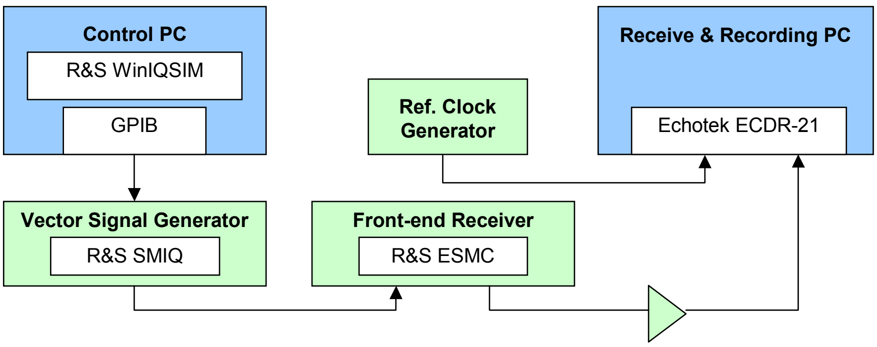

Figure 10: Signal generation and receiver set-up.

6.1.1 Data Source

In contrast to radar signals, radio communication signals normally carry some information content. The classification algorithms must not be sensitive to the content itself. The training and testing of the proposed signal classifier is therefore based on randomised signal sequences. Real, live communication signals cannot be expected to be fully randomised. Yet, as long as all symbol states are statistically equally represented within the signal sequence one may assume that the classifier should perform without additional degradation. Thus, a detached pseudo random bit sequence generator was developed which use a time-seeded random initial status. This ensured that each data sequence used for modulation of the radio signal was unique.

6.1.2 Modulation

For this experiment the classifier was intended to recognise and classify the six basic digital modulation formats 2- and 4-level ASK, PSK and FSK. These signal formats were thus used for both training and testing. In order to test the classifier's capability of detecting unknown signal formats, signal sequences of 16- and 32-level QAM, MSK and π/4DQPSK were generated. All signals had a fixed data rate of 10 kSymbols/s and were distorted with additive white Gaussian noise in the range 24-3 dB signal/noise ratio with 3 dB intervals.

6.1.3 Receiver and Data Recording

At the receiver side, care was taken to adapt the power level to utilise the full dynamic range of the A/D converter. The IF bandwidth of the receiver was set at 100 kHz, well above the signal bandwidth. The intention was to allow for optimised post-detection digital filtering of the received sequences. Each signal sequence was about 2 seconds long, corresponding to 500 kSamples at a sampling rate of ≈250 kHz. The recorded files contained a header section that described the parameters used for data creation, receiving and recording.

6.1.4 Classification Pre-processing and Feature Extraction

For the purpose of training and testing of the classifier, signals were divided into 2048-sample (8 ms) segments. With a symbol rate of 10 kHz, the segment thus consisted of approximately 80 symbols. From each individual segment the bandwidth was estimated for digital filtering. The training set consisted of 400 signal segments (50 segments per signal/noise ratio level) for each of the six modulation formats (2ASK, 4ASK, 2PSK, 4PSK, 2FSK and 4FSK). That made, in total, 2400 signal segments for training. The testing set consisted of 1144 signal segments (143 segments per signal/noise ratio level) for each of the ten modulation formats (16QAM, 32QAM, MSK and π/4DQPSK in addition to the ones above).

Eight features were extracted from each signal segment before being applied to the classifier. These are predominantly time-domain features that are shown to be effective for classifying ASK, PSK and FSK formats. The features are described in Table 1.

Table 1. Feature extraction (obtained from [3] and [16]).

Feature	Description
γ_{max}	Maximum power spectral density of normalised-centred instantaneous amplitude
σ_{ap}	Standard deviation of the absolute value of the centred instantaneous phase
σ_{dp}	Standard deviation of the centred instantaneous phase
σ_{aa}	Standard deviation of the absolute value of the normalised-centred instantaneous amplitude
σ_{af}	Standard deviation of the absolute value of the normalised-centred instantaneous frequency
σ_{da}	Standard deviation of the normalised-centred instantaneous amplitude
σ_{df}	Standard deviation of the normalised-centred instantaneous frequency
γ_{maxf}	Maximum power spectral density of normalised-centred instantaneous frequency

6.1.5 Neural Network Specifications

For this experiment we chose to use a MLP with one hidden layer consisting of 10 nodes. We can thus represent the MLP structure as 8I-10H-6O, where 'I', 'H' and 'O' represent the input, hidden and output layers respectively. We also used six AANNs with one hidden layer consisting of 5 nodes, thus 8I-5H-8O. The networks were trained using the back-propagation learning algorithm.

6.2 Results

Assessment of the combined neural network classifier is achieved by comparing it with the individual neural network types that it comprises. Useful metrics are

- *Classification rate of known formats*. This indicates how well the classifier is able to discriminate the modulation format that it has been trained to recognise.
- *Detection rate of unknown formats*. This indicates how capable the classifier is at detecting signal formats that are unknown to the classifier. (These are classified as "unknown").
- *False rejection rate of known formats*. This is the rate of known formats being falsely rejected (classified as unknown).
- *Mix-up rate of known formats*. This is the rate of known formats being misclassified as another known format.

These metrics are presented in the following sub-sections. The results are obtained from the confusion matrices contained in Appendix B.

6.2.1 MLP Results

In order to make the MLP detect unknown formats we accept its classification only if at least one $\mu_c > 0$. If all $\mu_c < 0$ we classify the input as "unknown". The results from applying only the MLP to the classification problem is presented in Table 2.

Table 2. MLP Results

Modulation formats	Correctly classified	Falsely rejected	Mixed up
Known	98.12 %	0.57 %	1.31 %
Modulation formats	**Correctly detected**	**Falsely classified**	
Unknown	38.22 %	61.78 %	

From the table we see that the MLP is very good at classifying known signal formats but that it is unreliable in detecting unknown signal formats. This is also reflected in how misclassified formats are distributed: 1.31 per cent of the known formats are mistaken for another format, whereas the remaining 0.57 per cent are incorrectly rejected.

6.2.2 AANN Results

The results from applying six AANNs to the classification problem are presented in Table 3. The results are based on a rejection threshold, τ_c, which is set to the 85th percentile of the reconstruction errors of the training data set.

Table 3. AANN Results

Modulation formats	Correctly classified	Falsely rejected	Mixed up
Known	82.20 %	15.40 %	2.40 %
Modulation formats	**Correctly detected**	**Falsely classified**	
Unknown	71.88 %	28.12 %	

For the AANN classifier, the classification rate of known formats is lower than for the MLP. This coincides with the conclusions reached by visually inspecting decision boundaries in Section 4, namely that the AANNs perform worse than the MLP in classifying (closely situated) modulation formats. However, the detection of unknown formats is considerably better. We also see that a larger proportion of misclassified known formats are rejected rather than classified as another format.

6.2.3 Combined Neural Network Results

By combining the MLP and the six AANNs above, according to Section 5, we obtain the results presented in Table 4.

Table 4. Combined Neural Network Results

Modulation formats	Correctly classified	Falsely rejected	Mixed up
Known	82.56	16.74	0.70
Modulation formats	**Correctly detected**	**Falsely classified**	
Unknown	88.18	11.82	

The table shows that we have obtained a similar classification rate as the AANN classifiers. We also see that the detection rate of unknown formats is notably better than both the AANN and the MLP. This is because the combined neural network detects both the unknown formats detected by the AANNs, but also the (fewer) formats detected by the MLP. Of misclassified known formats, the combined neural network regards 16.74 per cent as unknown whereas only 0.70 per cent are mistaken for another format.

Now, the overall performances of the classifiers depend on the weighting of the importance of classifying known formats and detecting unknown ones. In this experiment we have assumed an equal importance, thus we obtain the overall results presented in Table 5.

Table 5. Overall Classification Results

Classifier	Overall Success Rate
MLP	68.17 %
AANNs	77.04 %
Combined Neural Network	85.37 %

7.0 DISCUSSION AND CONCLUSION

Depending on the application (cf. Section 2), the importance of discriminating between known formats versus the detection of unknown formats may vary. The combined neural network can accommodate these changing requirements by adjusting the parameters in the network. In cases where the detection of unknown formats is more important than discriminating between known formats, the detection thresholds of the AANNs can be reduced. This will ensure that fewer unknown formats are misclassified as known at the expense of a reduced discrimination performance within known formats. If discrimination is more important, the AANN thresholds can be relaxed, which will increase the influence of the MLP and hence improve the discrimination performance. This, of course, will be at the expense of possibly detecting fewer unknown formats. Furthermore, in situations where we want to find appropriate jamming techniques, we would rather that the classifier, in case of misclassification, provided a 'second guess' rather than refusing to classify altogether. For example, the combined classifier could be adjusted to output possible solutions in cases where individual classifiers disagreed rather than labelling them as 'unknown'.

These issues illustrate the need for addressing problems beyond simply classifying a number of modulation formats. Even though we do not imply that, at this stage, the proposed combined neural network can readily handle all these problems, our results highlight some important points:

- Neural networks display good performance with a variety of additive Gaussian noise levels. This confirms previous results (cf. Section 3.4).
- Combining learning systems with different capabilities has been shown to be useful for handling more complex problems.
- The variety of requirements within modulation recognition necessitates adaptable classifiers. In this sense, combined classifier systems may be more versatile than stand-alone systems.

This work has addressed only some of the challenges involved in automatic modulation classification, and focused mainly on the last stage of the pattern classification approach (see Figure 1). It is important to note that any classification performance is also highly dependent on the effectiveness and accuracy at the receiver, pre-processing and feature extraction stages. To address this, comprehensive testing on more realistic signals will be necessary in future research.

8.0 REFERENCES

[1] J. Palicot, C. Roland, "A new concept for wireless reconfigurable receivers", IEEE Communications Magazine, July 2003, pp. 124-132.

[2] N. Ghani, R. Lamontagne, "Neural networks applied to the classification of spectral features for automatic modulation recognition", in Proceedings of the Military Communication Conference - MILCOM 1993, pp. 111-115.

[3] E. E. Azzouz, A. K. Nandi, Automatic Modulation Recognition of Communication Signals, Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996.

[4] F. Liedtke, "Adaptive procedure for automatic modulation recognition", Journal of Telecommunications and Information Technology, vol. 4, 2004, pp. 91-97.

[5] B. G. Mobasseri, "Digital modulation classification using constellation shape", Signal Processing, vol. 80, 2000, pp. 251-277.

[6] R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification – 2nd Edition, New York: John Wiley and Sons Inc., 2001.

[7] C.- Y. Huang, A. Polydoros, "Likelihood methods for MPSK modulation classification", IEEE Transactions on Communications, vol. 43, 1995, pp. 1493-1504.

[8] D. Boudreau, C. Dubuc, F. Patenaude, M. Dufour, J. Lodge, R. Inkol, "A fast automatic modulation recognition algorithm and its implementation in a spectrum monitoring application, in Proceedings of the Military Communication Conference - MILCOM 2000, pp. 732-736.

[9] J. Lopatka, M. Pedzisz, "Automatic modulation classification using statistical moments and a fuzzy classifier, in Proceedings of the International Conference on Signal Processing – ICSP 2000, pp. 1500-1506.

[10] N. Kim, N. Kehtarnavaz, M. B. Yeary, S. Thornton, "DSP-based hierarchical neural network modulation signal classification", IEEE Transactions on Neural Networks, vol. 14, 2003, pp. 1065-1071.

[11] G. Hatzichristos, "Classification of digital modulation types in multipath environments", Electrical Engineer Thesis, Naval Postgraduate School, Monterey, California, 2001.

[12] J. Venäläinen, L. Terho, V. Koivunen, "Modulation classification in fading multipath channel", in Proceedings of the Conference on Signal, Systems and Computers, 2002, pp. 1890-1894.

[13] C. M. Bishop, Neural Networks for Pattern Recognition, Oxford: Oxford University Press, 1995.

[14] M. Markou, S. Singh, "Novelty detection: A review – part 2: Neural network approaches", Signal Processing, vol. 83, 2003, pp.2499-2521.

[15] A. J. C. Sharkey, "Multi-Net Systems", In A. J. C. Sharkey (editor), Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems, Springer-Verlag, 1999.

[16] G. Arulampalam, V. Ramakonar, A. Bouzerdoum, D. Habibi, "Classification of digital modulation schemes using neural networks", in Proceedings of the International Symposium on Signal Processing and its Applications – ISSPA 99, pp. 649-652.

[17] G. L. Luger, Artificial Intelligence: Structures and Strategies for Complex Problem Solving, Harlow, UK: Addison Wesley, 2002.

A EXTRACTION OF CERTAINTY FACTORS

The interpretation of the certainty factors, μ_c and α_c, extracted from the MLP and AANNs, respectively, is based on the *Stanford Certainty Factor Algebra* [17], which is concerned with the combination of certainties from different 'experts'. Due to the nature of MLP outputs, y_c, when using the 1-of-C representation, we find it adequate to transform the outputs to certainty factors directly as

$$\mu_c(y_c) = y_c \tag{1}$$

if $-1 < y_c < 1$, or as

$$\mu_c(y_c) = 2y_c - 1 \tag{2}$$

if $0 < y_c < 1$. The extraction of certainty factors from an AANN, on the other hand, requires more calculation, as the AANN does nothing more than trying to reconstruct the input at the output. The first step to extract the certainty factor, α_c, from the AANN is to obtain a reconstruction error, e_n, from an example, n. If z_k is the kth output of the AANN, the reconstruction error can be expressed as

$$e_n = \left(\sum_{k=1}^{N_I} (x_k - z_k)^2 \right) / N_I \tag{3}$$

where x_k is an input and N_I is the number of inputs/outputs. If the AANN has been trained using a training set T_c, we can obtain a set of training reconstruction errors, S_c, where

$$S_c = \{e_n, n \in T_c\} \tag{4}$$

S_c can now be used to determine the error threshold τ_c, below which we can assume that n belongs to the modulation format c. For example, we can set τ_c to be the maximum error obtained from the training data set, i.e., $\tau_c = \max(S_c)$, or, if we want the system to be more sensitive to unknown modulation formats, the threshold can be set to the p-th percentile of S_c, $p < 100$.

To obtain a certainty factor in the range $-1 < \alpha_c < 1$, we choose to transform the e_n according to a hyperbolic tangent function (cf. Figure 7), which is similar to the activation function of the output nodes of an MLP. Thus

$$\alpha_c(e_n) = \frac{2}{1 + \exp(K(e_n - \tau_c))} - 1 \tag{5}$$

where K is a constant. As this function never reaches the extremes of -1 or 1, we can define an R, where $\alpha_c(0) = R$. Instead of experimenting with different values of K, we can represent K in terms of R

$$\alpha_c(0) = \frac{2}{1 + \exp(-K\tau_c)} - 1 = R \tag{6}$$

$$K = -\frac{\ln(\frac{2}{R+1} - 1)}{\tau_c} \tag{7}$$

and, finally, replace K in (5) such that

$$\alpha_c(e_n) = \frac{2(R-1)}{\left(\frac{-R+1}{R+1} \right)^{\frac{e_n}{\tau_c}} (R+1) - R + 1} + 1 \tag{8}$$

Now, we can control the shape of α_c by selecting an appropriate R. If R approaches 1, α_c will approach a step function that changes between 1 to –1 at τ_c. By reducing R, α_c will approach linearity in the range $[0, 2\tau_c]$. Typically, we want R to be close to 1 as it represents the certainty value when the reconstruction error is 0. In Figure 7, $R = 0.99$.

B CONFUSION MATRICES

Table 6. Confusion matrix for the MLP classifier

Actual Format	Predicted Format						
	2ASK	**4ASK**	**2PSK**	**4PSK**	**2FSK**	**4FSK**	**Unknown**
2ASK	95.98	3.58	0.00	0.00	0.00	0.00	0.44
4ASK	2.10	95.80	0.44	0.00	0.00	0.00	1.66
2PSK	0.00	0.00	97.64	1.31	0.00	0.00	1.05
4PSK	0.00	0.00	0.26	99.56	0.00	0.00	0.17
2FSK	0.00	0.00	0.00	0.00	99.74	0.17	0.09
4FSK	0.00	0.00	0.00	0.00	0.00	100.00	0.00
MSK	0.00	0.00	0.00	100.00	0.00	0.00	0.00
π/4DQPSK	0.00	0.00	2.10	97.47	0.00	0.00	0.44
16QAM	0.44	0.35	13.81	7.95	0.00	0.00	77.45
32QAM	0.61	0.70	14.95	8.74	0.00	0.00	75.00

Table 7. Confusion matrix for the AANN classifier

Actual Format	Predicted Format						
	2ASK	**4ASK**	**2PSK**	**4PSK**	**2FSK**	**4FSK**	**Unknown**
2ASK	83.13	1.92	0.00	0.00	0.00	0.00	14.95
4ASK	8.22	77.36	0.00	0.00	0.00	0.00	14.42
2PSK	0.00	0.44	83.57	0.26	0.00	0.00	15.73
4PSK	0.00	0.09	3.50	81.03	0.00	0.00	15.39
2FSK	0.00	0.00	0.00	0.00	85.58	0.00	14.42
4FSK	0.00	0.00	0.00	0.00	0.00	82.52	17.48
MSK	0.00	0.00	0.52	14.86	0.00	0.00	84.62
π/4DQPSK	0.00	0.09	5.42	22.90	0.00	0.00	71.59
16QAM	0.17	18.71	9.09	4.20	0.00	0.00	67.83
32QAM	0.00	19.76	11.45	5.33	0.00	0.00	63.46

Table 8. Confusion matrix for the Combined Neural Network Classifier

Actual Format	Predicted Format						
	2ASK	4ASK	2PSK	4PSK	2FSK	4FSK	Unknown
2ASK	81.73	2.10	0.00	0.00	0.00	0.00	16.17
4ASK	1.75	80.68	0.00	0.00	0.00	0.00	17.57
2PSK	0.00	0.00	82.43	0.26	0.00	0.00	17.31
4PSK	0.00	0.00	0.09	82.43	0.00	0.00	17.48
2FSK	0.00	0.00	0.00	0.00	85.58	0.00	14.42
4FSK	0.00	0.00	0.00	0.00	0.00	82.52	17.48
MSK	0.00	0.00	0.00	15.21	0.00	0.00	84.79
π/4DQPSK	0.00	0.00	0.09	25.09	0.00	0.00	74.83
16QAM	0.00	0.00	1.75	1.22	0.00	0.00	97.03
32QAM	0.00	0.09	1.40	2.45	0.00	0.00	96.07

C ABBREVIATIONS

AANN	Auto-Association Neural Network
A/D	Analogue to Digital
ASK	Amplitude Shift Keying
FM	Frequency Modulation
FSK	Frequency Shift Keying
IF	Intermediate Frequency
MLP	Multi-Layer Perceptron
MSK	Minimum Shift Keying
π/4DQPSK	π/4 Differential Quadrature Phase Shift Keying
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
RF	Radio Frequency