Probabilistic Neural Network Implementation of the Optimum CDMA Multi-user Detector

* Ibikunle, Frank.A. and Prof. Zhong Yixin**
*Beijing University of Posts & Telecommunications, 100088 Beijing, China.
* E-mail: faibikunle2@yahoo.co.uk

Abstracts

Application of Neural Network to signal detection in CDMA multi-user communications Gaussian channel is investigated, This paper is motivated by the fact that, in a multi-user CDMA system, the conventional receiver suffers severe performance degradation as the relative powers of the interfering signals become large (i.e., "near-far problem"). Furthermore, in many cases the optimum receiver, which alleviates the near-far problem, is too complex to be of practical use. And by viewing this optimum multi-user detector problem in CDMA channel as an optimum nonlinear classification decision problem, we apply the Probabilistic Neural Network algorithm which has the abilities of arbitrary nonlinear transformations, adaptive learning and tracking to implement this classification decision optimally and adaptively. The performance of the proposed neural detector is evaluated via computer simulations in terms of probability of detection, and it is compared with those of the existing neural and conventional detector schemes in a multi-user environment.

1. Introduction

CDMA due to its high capacity potential and other desirable features such as anti-multipath, fading and anti-jamming capabilities is aim at providing a much higher capacity than the previous generations of mobile communication system. This technique permits all the users in the system to transmit simultaneously, operate at the same nominal frequency, and use the entire system bandwidth. It is achieved by spreading the spectrum of the transmitted signals using pre-assigned code waveforms. The traditional approach of demodulating the spread spectrum signals in multiuser system is that the spread spectrum signals are made to pass through a filter matched to the desired signal, and then the decision is made directly [1]. This approach ignores the multiple-access interference (MAI), or equivalently, ignoring the cross-correlation between the modulating signals of different users. For this reasons, it has stimulated the research interest in optimum receivers in multi-user communication system in [2]. This scheme has been shown to have a complexity that is exponential in the number of users. Therefore, research efforts have focused on suboptimal receivers that will; first, exhibits good near-far resistance properties, secondly, have low computational complexity to ensure their practical implementation, and thirdly, achieve good bit error rate (BER). Among the many suboptimal multi-user detectors proposed are the de-correlating detector developed by Lupas [3], that is, linear in nature and computational complexity. Another non-linear suboptimal multi-user detectors is the multistage detector developed in [4], that relies on improving each stage's estimate by subtracting the estimate of the MAI obtained by the previous stage. It is relatively complex if compared with other suboptimal detectors because it requires several decision stages to detect signal. In this paper, there are several factories that motivate us to investigate the use of neural networks as a multi-user detector. First, in addition to fast processing speed, computational efficiency, and near-optimal performance, we seek adaptivity. Secondly, the decision boundaries formed by the optimal receiver are nonlinear in nature. Thirdly, the highly structured nature of the MAI suggests that a neural network should be able to learn and remove the MAI effectively. Recently [8], [9] and [10] in their papers have proposed feed forward Back Propagation and Hopfield neural networks based multi-user detectors. While their performance is shown to be very good for a small number of synchronous and asynchronous users, their hardware complexity (i.e., number of hidden layer neurons, storage capacity and training time before convergence) appears to be exponential in the number of users. Here, we propose a new Probabilistic Neural Network (PNN) as a potential multi-user CDMA detector.

The paper is organized as follows. Section 2 describes the CDMA communications detection scheme. Section 3 describes the implementation, the architecture and a complete performance evaluation of the CDMA multiuser detector employing this neural network. In Section 4, discussion on numerical computer simulation results are presented. Conclusions and future research directions are discussed in section 5.

2. CDMA multi-user detection model

Assuming that K active transmitters shared the same Gaussian channel at a given time instance. A signature waveform, $S_k\ (t)$, (for k = 1,2,..., K) time limited in the interval [0,T], is assigned to each transmitter. Lets denote the i-th information bit of the k-th user as $b_{k,i}\ \epsilon\{+1,\ -1\}$. In general CDMA system, the signal at a receiver, r(t) is the superimposition of k transmitted signals and AWGN, given as

$$r(t) = \sum_{i=-M}^{M} \sum_{k=1}^{K} b_k^{(i)} S_k(t - iT - \tau_k) +$$

$$n(t)$$

$$t \in [iT, (iT + \quad (1)$$

In (1), r_k e [0,T] are the relative time delays associated with all users and 2M + I is the packet size, n(t) represents the AWGN with a spectral density with height σ^2. Assuming that all possible information sequences are independent and equally likely, an optimum decision on $b' = (b_1^i, b_2^i, \ldots)^T$ requires the observation of the received signal only in the i-th time interval. Without loss of generality, we therefore focus our attention on i= 0 and drop the time superscript and consider the demodulation of the vector of bits with the observation of the received signal in the time interval $\{0, T_b\}$, [l,2]. The optimum or maximum likelihood decision on **b** is chosen as $b^* = (b_1^*, b_2^*,)^T$ which maximizes the log of the likelihood function, expressed as

$$b^* =$$

$$arg\ \max_{b\in\{-1,1\}^k}\{2\sum_k \int_0^{T_b} b_k S_k(t) r(t) dt -$$

$$\int_0^{T_b} [\sum_k b_k S_k(t)]^2 dt\}$$

(2)

The optimum decision can also be written as

$$b^* = arg\ \max_{b\in\{-1,1\}^k}\{2y^T\ b - b^T \ldots\ldots(3)$$

Where $y = (y_1, y_2, \ldots, y_{k-1})^T$ is the sampled vector of output $\mathbf{y_k}$ at time $t = T_B$, which is the sampled output of the normalized correlation receiver matched to the i-th signal passed by the received signal **r (t).** The sampled vector **y** is called the observation vector and it's a sufficient statistic for demodulating **'b'**. Investigations on processing these sufficient statistics which according to equations (1), (2) and (3) depends on the transmitted bits in the following way

$$y= Hb + n= RWb+ n \qquad (4)$$

where H = *WR* is referred to as the equivalent transfer matrix in CDMA system. $H \in R^{K \times K}$ is the symmetric matrix of signal cross-correlations

$$h_{ki} = \int_0^T S_k(t) S_i(t) dt =$$

$$\int_{-\infty}^{\infty} S_k(t - \tau_k) S_i(t - jT - \tau_i) dt$$

(5)

Where **W** is called the energy matrix, and is the diagonal matrix in which the diagonal elements $W_{ki} > 0$, denoting the signal energy of the i-th user in the received signal r(t). When the transmitted power is constant, W_{ki} depends on the distance. When the k-th user changes his location, $\mathbf{W_{ki}}$ will be changed. $R = [r_{ij}]_{k \times k}$ is the correlation matrix of each user's PN code, which is a diagonal matrix, (i.e., $r_{ij} = r_{ji}$). Since each user's PN code is constant, matrix R is not changed. As there is no absolute orthogonality between PN codes, the non-diagonal dements of R is not zero, i.e., $r_{ij} \neq 0$, $(I \neq j)$. Therefore, there unavoidably exist multiple access interference in the system. And **n** is a zero-mean Gaussian K-vector with covariance matrix equal to $\sigma^2 H$.

$$n = \int_0^T n(t) S_k(t \qquad (6)$$

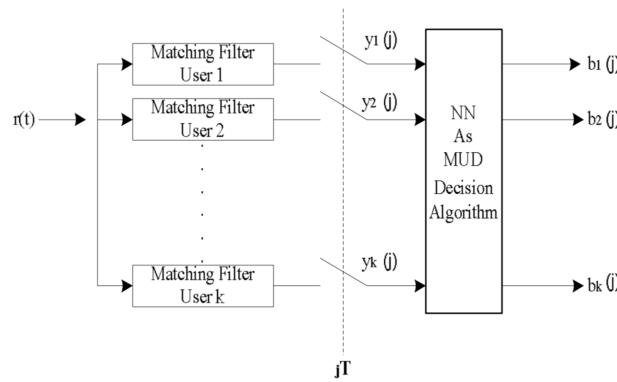

Figure 1. The general structure of a Multiuser CDMA detector

3. The PNN implementation and architecture

The CDMA multi-user detector problem is essentially to decide and estimate each user's transmitted signal from the observation vector with minimum error probability at a view time interval. So, the task at hand is of decision and classification. And, because neural networks with multi-layers perceptron (MLP) structure could implement arbitrary complex nonlinear functions by the use of proper training algorithm provided that the number of neurons in the hidden layers is adequate. So, we consider the use of Probability Neural Network algorithm based on Parzen Probability Density Function estimation theory to implement the nonlinear classification decision function problem, and to approximate a solution to the maximization problem in equation (3) that is known to be NP-hard complete [11]. **Figure 2** shows the neural network considers for the detection problem in spite of asynchronous arrivals of the active users. The training process adapts to the asynchronous nature of the system if the relative delays and phases of the active users' signals are assumed to be unknowns constant during the training and data transmission periods. And since MLPs allow only for discrete time inputs, the continuous time received CDMA signal which is used to form the hypothesis testing **problem is converted to its corresponding discrete time signal version which now form the input to the neural network. Essentially there exist two relevant possibilities: our active k-users transmit +1 or a –I.** This user's bits are embedded in noise that consists of MAI and Gaussian noise. And for the proposed network to implement a Maximum Likelihood detector, we viewed the input vector (consisting of the active user's bits embedded in MAI and Gaussian noise) to the network for detection as the following classification problem (or as a binary hypothesis-testing problem) H_1 and H_0 [8]:

$$H_1: \ y = +A_k S_k + I + \mu \qquad (7a)$$

$$H_0: \ y = -\ A_k S_k + I + \mu \qquad (7b)$$

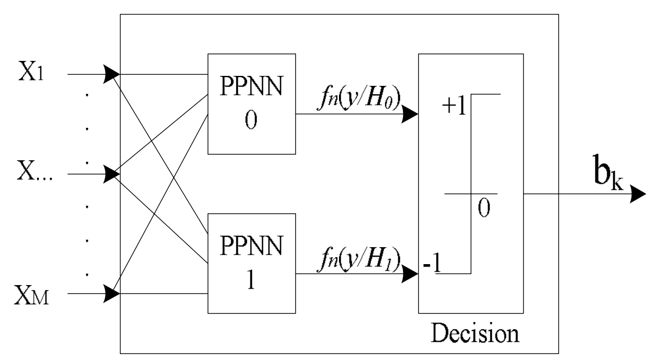

Figure 2. The Probabilistic Neural Network detector.

Where y is the observation of the active user's signal in noise, the desired signals vector is $\pm A_k S_k$. The received signal amplitude of desired users is A_k, and its spreading code denoted by sequence S_k. The sum of the interfering user's signals is 'I', and 'μ' is the length-N vector of AWGN noise samples. A Maximum Likelihood detector calculates a likelihood ratio 'λ' for a binary decision problem and compares it to a threshold 'β' [8]:

$$\lambda = \frac{f(^{y}/_{H_1})}{f(^{y}/_{H_0})} = \begin{matrix} H \\ H \end{matrix} \qquad (8)$$

Where the probability density of the observation of the desired users' signals in noise given that hypothesis (H_1) is true is $p(y/H_1)$, given as

$$p\left(\frac{y}{H_1}\right) = \frac{1}{2^{k-1}} \sum_{i=1}^{2^{k-1}} \frac{1}{(2\pi)^{\frac{N}{2}} \sigma^N} . e - \frac{1}{2\sigma^2} (\mathbf{y_m} - \mathbf{y_{+1m}})^T \mathbf{R}^{-1} (\mathbf{y_m} - \mathbf{y_{+1m}})$$

(9)

Similarly, the probability density of the observation given that hypothesis (H_0) is true is $p(y/H_0)$, given as

$$p\left(\frac{y}{H_0}\right) = \frac{1}{2^{k-1}} \sum_{i=1}^{2^{k-1}} \frac{1}{(2\pi)^{\frac{N}{2}} \sigma^N} . e - \frac{1}{2\sigma^2} (\mathbf{y_m} - \mathbf{y_{-1m}})^T \mathbf{R}^{-1} (\mathbf{y_m} - \mathbf{y_{-1m}})$$

(10)

For this work, the threshold 'β' is taken to be equal to one. Thus, the following decision rule is applied

H_1 is true: $f(y/H_1) > f(y/H_0)$ **(11a)**

H_0 is true: $f(y/H_1) < f(y/H_0)$ **(11b)**

Where y_{im} is the i-th samples (+1 or -1) of the m-th signal, and

$Y_{+1m} \in \{ y = Hb = RWb/b_k = +1\}$

$Y_{-1m} \in \{ y = Hb = RWb/b_k = -1\}$,

for $i = 1, 2, \ldots, 2^{k-1}$, which are the non noise observing vectors which has 2^{k-1} vectors when the k-th users transmits signals $b_k = +1$ and $b_k = -1$, respectively. All the vectors and the nonlinear decision function $f'_k(y)$ are only been determined by matrix H = R W, and the noise power 'σ^2. It can be seen that in other to make a decision one has to calculate $f(y/H_1)$ and $f(y/H_0)$ using equations (9) and (10). The key to using equation (8) in the network is the ability to estimate accurately the probability densities function based on the training samples. Often, the priori probabilities are known. However, if the probability densities of the samples in the hypotheses to be separated are unknown, and all that is given is a set of training samples, then it is these training samples that provide the only clue to the unknown underlying probability densities. The likelihood ratio in equation (8) is now used for the construction of the Receiver Operating Characteristics (ROC) curves for this detection and classification problem, and it is a random variable distributed according to $f(y/H_0)$ and $f(y/H_1)$ for hypotheses H_0 and H_1. An ROC curve for a binary hypotheses problem is a plot of the probability of detected signals 'P_D' versus the probability of undetected signals '$P_{No.D}$'. They are calculated from equation (12) below by simulation for different values of varying threshold, β =

{1,-1} to forms the two performance measures for the detection problem.

$$P_D = \int_\beta^\alpha f(\lambda/H_1 \text{, and } \quad P_{No.D} = \int_\beta^\alpha f(\lambda/H_0 \quad \textbf{(12)}$$

The network architecture realized for implementation of a Maximum Likelihood detector and classifier is as shown in figure 2, and described by equation (11) above. It is a multi-layer feedforward networks trained with the PPNN algorithm that uses a sum of Gaussian distributions to estimate the probability density function (pdf) for a training data samples. The trained network is then used to classify new data samples based on the learned pdfs, and further, to provide a probability factor associated with each class. The developed network though similar in structure to the Back Propagation neural network with three layers structure of: input, pattern (or hidden) and summation (or output) layers; but differs primarily by employing exponential function instead of sigmoid function as activation function in the pattern layer [7,11,12]. The input layer of the network consist of M nodes (M = n) and distributes the sampled signals to the pattern units, while the output or decision device which is a two-input neurons as shown compares the likelihood ratio functions outputted by the two networks and produces a binary signal output (-1 or +1) to declare that the input samples belongs to the hypothesis associated to the larger pdf. Hence it has one node obtaining values in the range [1,-1] for the detection problem and M nodes for the classification problem. In the pattern layer, the number of the pattern units corresponds with the training samples number. The weight coefficients of network **1** and network **0** are determined by training samples number **'n"** from both hypotheses H_0 and H_1, respectively. The output of the two networks now produces the conditional density functions $f(y/H_1)$ and $f(y/H_0)$.

4. Discussion of simulation results

Here the performance of a feedforward multilayered perceptrons referred to as Parzen Probabilistic Neural Network based Maximum Likelihood detector for detection and classification of CDMA users in an asynchronous multi-user communications Gaussian channel with near-far problem is considered. The K = 6 active users employ a spreading code of length N = 63. The signal-energy to-noise variance ratios of the users are 10 and 1, respectively. During simulations, the relative delays of the interfering users' signals at the receiver are assumed to be unknown constant during the training and data transmission periods. As depicted in **figure 2**, the neural detector input layer consists of M nodes (i.e., received discrete samples of each signal), has 2^K nodes in the hidden (or pattern units) layer, and had one node obtaining values in the range of [-1, +1] in the output layer for the detection problem. The weights in each of the pattern units are set to correspond to each of the training samples. The noise were added to the training samples during the training stage, and after the completion of the training process a large number of patterns consisting of pure Gaussian noise samples and the multi-user interference signals with added Gaussian noise was propagated through the neural detector. The values of the output nodes for each of the above samples was then compared to a threshold ' **β** ' varying in the range [-1, 1]. An output containing positive value samples which is larger than the threshold is interpreted as detection. Similarly, no-detection is indicated when the network output for a negative value samples exceeding the threshold. The number of detection and no-detection for every threshold value over the total number of propagated positive and negative value samples, respectively, yield the estimates of the pdf of detection and no-detection, which forms the PPNN receiver operating characteristic (ROC) curves.

Shown in **figure 3a and 3b** is the ROC curves obtained for the neural network multi-user CDMA detector for 6 active users with AWGN. The training samples number generated is vary from 100 to 1000 for a ratio of signal energy over noise variance of 10 and 1. The effectiveness of the proposed PNN detector in a near-far situation is fully shown in these graphs. The probability of detection versus the probability of no-detection when the network is trained with and without noise (i.e., active users and AWGN) added for different signal-energy-to-noise variance ratios are also shown. As can be seen, the distances between the curves obtained from simulations are not too far away from each other, which indicate no much performance degradation with addition of the noise interference. High detection probability were attained for even a quiet large number of active users in the system, but this tend to fall as the signal-energy-to-noise variance ratios becomes smaller, i.e., $E_s/\sigma_r^2 =$ 10. The principal conclusion from our simulation results is that, the proposed neural detector is capable of detecting users signal in a multi-user CDMA communications system in the presence of strong interfering users. It shows adaptability to some of the unknown parameter in the system, especially, the energy matrix (power) which vary with the distance between users These change of the energy matrix which is caused by the users movement are adaptively traced by the neural network As a result, the detector displays a great resistance to the near-far problem, thereby eliminating the need for strict power control.

Similarly, the neural network classifier was tested on a large number of training samples corrupted with and without noise information. The performance of the neural classifier was then evaluated on the basis of the number of misclassified patterns over the total number of transmitted sequences. There is the indication that the network classification error depends on the presence of total noise interference (i.e., MAI and Gaussian noise) in the training stage, and the error tend to be slightly change with increase in number of active users. For example, with 6 active users in the channel, the classification error when compared with that of 10 active users is much lower. Differences between classification errors of the network trained with and without noise interference tend to decrease for smaller signal energy to noise variance ratios. The network trained without noise interference for higher

E_s/σ_n^2 =1 gives lower classification errors over that of network trained with total noise interference for a lower E_s/σ_n^2 =10. Also, the optimal classifier with K = 1, which offers the lower bounds for the classification error and used in our work as a criterion for the evaluation of the Probabilistic Neural Network in the classification problem attest to the simulation results. Indication from extensive simulation results shows that for an infinitely large number of training samples, the classification error of the neural classifier is minimum compared to that of small number of training samples.

5. Conclusion

In the paper, we have proposed a Probabilistic Neural Network detector and classifier in which simulation results show an improved performance over the conventional CDMA detector. And when compared with other neural network multi-user detector schemes in a multi-user environment, our proposed P'NN detector exhibits a number of attractive properties and significantly outperforms them in terms of low computational complexity, fast processing speed, low classification error and near-far resistance. The proposed neural detector provides strong performance even in the absence of complete system parameter knowledge. Also, because memory is dense and inexpensive with the development and availability of ASICs and programmable general-purpose digital signal processing techniques, the need for acceptable buffer size for the training samples will not pose a problem to the neural detector hardware implementation. More research work into its capability to be applied in non-symmetric noise case, and in multipath fading communications environment is a further research area and it's still going on.

References

[I] S.Verdu, ''Minimum Probability of Error for Asynchronous Gaussian Multiple Access Channels''. IEEE Trans. on Information Theory, vol. IT-32, no. I, Jan. 1986.
[2] S.Verdu,"Optimum multi-user signal detection", Ph.D. dissertation, Dept of Elect & Computer Eng. Univ. of Illinois, Urban-Champaign, 1984.
[3] R.Lupas, "Near-far resistant linear multi-user detection", Ph.D dissert, Princeton Univ, Jan 1989.
[4] M,K,Varanasi and B, Aazhang, "Multistage detection in asynchronous code-division multiple-access systems", IEEE Trans. Comrn., vol.38, pp.509-519, Apr. 1990.
[5] H. L.Van Trees, Detection, Estimation, and Modulation Theory, Part I, Wiley, 1971.
[6] A.D Whalen. Detection of Signals in Noise. New York: Academic 1971.
[7] D.E.Rumelhart, and et.al,' Parallel Distributed Processing, vol. I: Foundations, The MIT Press, Cambridge, MA: 1986.
[8] James W. Watterson, "An Optimum Multilayer Perceptron Neural Receiver for Signal Detection", IEEE Trans, on NN, Vol. I, 1990.
[9] B. Aazhang, B. P. Paris, G. C. Orsak, "Neural Networks for Multiuser Detection in Code Division Multiple-Access Comms", IEEE Trans, on comm. vol. 40,, July 1992.
[10] J G. Kechriotis and E.Manolakos, "Hopfield Neural Networks Implementation of the optinral CDMA Multiuser Detector", IEEE Trans, on Neural Networks , vol.7, No. 1, Jan. 1996.
[II] D.F. Specht, Probabilistic Neural Networks. NN, vol. 3, 1990, pp. 109-118.
[12] E.Parzen, "On estimation of a probability density function and mode", Annal Math.

Table 1. Shows PNN classification errors

K=6 users	**Classification Error for Optimum Classifier**	**Classification Error for PNN Classifier trained without Noise**	**Classification Error for PNN Classifier trained with Noise**
$\frac{E_s}{\sigma_n^2} = 10$	0.0370	0.4210	0.0820
$\frac{E_s}{\sigma_n^2} - 1$	0.4190	0.4120	0.4910
K=10 users	**Classification Error for Optimum Classifier**	**Classification Error for PNN Classifier trained without Noise**	**Classification Error for PNN Classifier trained with Noise**
$\frac{E_s}{\sigma_n^2} = 10$	0.0821	0.0950	0.2150
$\frac{E_s}{\sigma_n^2} = 1$	0.4500	0.4721	0.7880

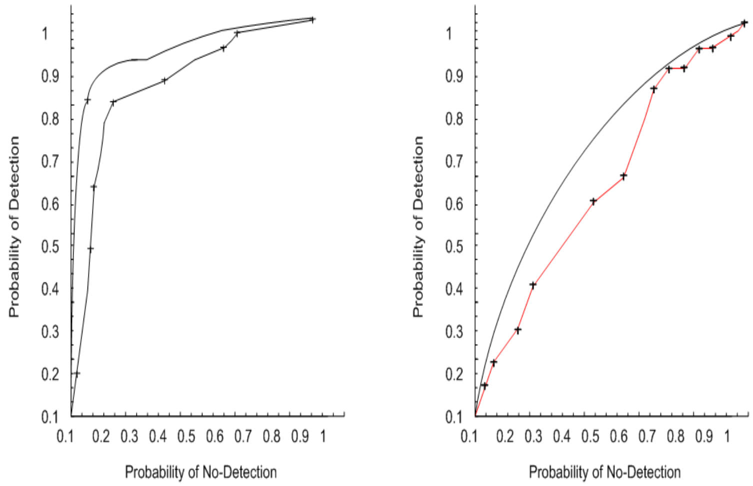

Fig 3a. The ROC curves showing the performance of PNN detector trained without noise (dark line) and with noise (dotted line) for K=6 CDMA user and for a signal-to-noise variance ratio of (a) 10 and (b) 1

Figure 3a. Performance of PNN detector trained without noise (dark line) and with noise (dotted lines) for K=6 CDMA users and for a signal-to-noise ratio of (a) 10 and (b) 1.