

UNIVERSIDADE FEDERAL DA BAHIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MECATRÔNICA

TIAGO ANDRADE MOTA

**SELEÇÃO DE ESTRUTURA NEURAL PARA O PROBLEMA
DE EQUALIZAÇÃO UTILIZANDO ALGORITMO GENÉTICO**

Salvador

2014

TIAGO ANDRADE MOTA

**SELEÇÃO DE ESTRUTURA NEURAL PARA O PROBLEMA
DE EQUALIZAÇÃO UTILIZANDO ALGORITMO GENÉTICO**

Dissertação apresentada ao Programa de Pós-
Graduação em Mecatrônica da Universidade
Federal da Bahia para obtenção do título de
Mestre em Mecatrônica.

Orientador: Prof. Dr. Antônio Cezar de Castro
Lima

Salvador

2014

TERMO DE APROVAÇÃO

TIAGO ANDRADE MOTA

SELEÇÃO DE ESTRUTURA NEURAL PARA O PROBLEMA DE EQUALIZAÇÃO
UTILIZANDO ALGORITMO GENÉTICO

Dissertação aprovada como requisito parcial para obtenção do título de Mestre em
Mecatrônica, Universidade Federal da Bahia, pela seguinte banca examinadora:

Salvador, 21 de outubro de 2014.

Antonio Cezar de Castro Lima



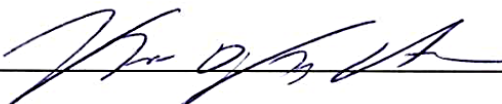
Doutor em Engenharia Eletrônica, University of Kent

Angelo Amâncio Duarte



Doutor em Ciência da Computação, Universitat Autònoma de Barcelona

Karcius Day Rosário Assis



Doutor em Engenharia Elétrica, Universidade de Campinas

Leizer Schnitman



Doutor em Engenharia Eletrônica e Computação, Instituto Tecnológico de Aeronáutica

A Clara, minha filha, que nasceu durante execução da pesquisa e a
Luciana, minha esposa, que me apoiou incondicionalmente em todos
os momentos.

AGRADECIMENTOS

A Antonio Cezar de Castro Lima, orientador, pela atenção em todos os momentos.

A Jorgean Ferreira Leal, amigo e compadre, por toda a ajuda no desenvolvimento da pesquisa: simulação do canal de comunicações, sugestões de soluções, discussão e interpretação dos resultados, revisão da dissertação e de artigos ...

Aos professores Ângelo, Leizer e Karcus por terem participado da banca e dado sugestões que melhoraram a qualidade da dissertação.

A Luciana, esposa, por ter me apoiado nos momentos mais difíceis e pela compreensão nas ausências.

Aos meus pais, Maria Auxiliadora e Samuel, por terem me presenteado com o dom da vida e por terem me transmitido o maior dos ensinamentos: o valor da educação.

Aos meus irmãos, Samuel Jr. e Daniel, por terem sempre me acolhido.

RESUMO

Os sistemas de telecomunicações sem fio vêm evoluindo rapidamente desde o início do século XXI, período no qual os dispositivos de comunicação móveis se tornaram muito populares. Hoje em dia as pessoas querem estar conectadas todo tempo, em qualquer lugar. Essa necessidade tornou o telefone celular um dos aparelhos eletrônicos mais populares do mundo. A cada nova geração tecnológica, uma série de inovações é incorporada, proporcionando taxas de transmissão de dados cada vez maiores, de forma a atender à demanda crescente. Alguns fatores conhecidos degradam a qualidade da comunicação móvel, dentre eles se destacando o ruído térmico, interferência intersimbólica e o desvanecimento. Uma das formas de mitigar esses problemas é o uso de equalizadores que, idealmente, implementam a função inversa do canal de transmissão, permitindo que o sinal recebido apresente mínimo erro em relação ao transmitido.

Muitos artigos da literatura científica tratam canais invariantes no tempo, contudo, atualmente, a maioria dos canais é variante, o que prejudica o desempenho dos sistemas e torna seus efeitos difíceis de mitigar. Assumindo que o canal de comunicações é variante no tempo, algoritmos de otimização são usualmente utilizados para atualização dos pesos neurais de forma a adaptar o equalizador às condições do canal, levando em conta que as estruturas utilizadas nesse trabalho são inspiradas em redes neurais.

Os trabalhos científicos que utilizam redes neurais geralmente empregam estruturas escolhidas sem uma investigação sistemática que justifique sua aplicabilidade. Não foi encontrada na literatura uma abordagem que comparasse algumas estruturas e indicasse a melhor para ser aplicada a um problema específico. É proposta nesse trabalho uma metodologia para comparar tipos diferentes de estruturas utilizando algoritmo genético.

Foi desenvolvido também um equalizador neural capaz de lidar com o sinal proveniente de canal sujeito a desvanecimento severo, além de interferência intersimbólica, causada por multipercursos, e ruído térmico. Após a escolha da estrutura, o próximo passo foi a seleção de um algoritmo de otimização para atualização dos pesos neurais.

Palavras-chave: Equalizadores – Desvanecimento – Interferência Intersimbólica – Comunicações – Redes neurais.

ABSTRACT

The wireless telecommunication systems have evolved quickly since the beginning of the current century, when the mobile communication devices have become very popular. Nowadays, people want to be connected all the time in everywhere, which makes the cell phone one of the most popular electronic devices around the world. Thermal noise, Intersymbolic Interference (ISI) and fading degrade the performance of mobile communication systems; therefore, equalizers are usually employed at the receiver to combat those distortions.

Most works in the scientific literature deal with time-invariant channels, but the majority of channels are time-variant, which diminish system performance and their effects are difficult to mitigate. Assuming the communication channel is time-variant, optimization algorithms usually update the neural network weights to adapt the equalizer to the channel conditions, considering that neural structures have been used in this research.

The scientific works that apply neural networks commonly employ an *ad hoc* structure without a systematic investigation that explains their results. Approaches that measure how much a structure fit to a specific problem are rare in scientific literature. In this work we propose a methodology to compares different structures, using genetic algorithm.

Develop a neural equalizer to deal with signals from fast fading channels, which also present thermal noise and intersymbolic interference (ISI), is another aim of this work. The selection of the optimization algorithm, which updates the synaptics weights of the chosen structure, is last step.

Key-words: Equalizer – Fading – Intersymbolic interference – Communication – Neural networks.

Lista de Ilustrações

Figura 1- (a) Modelo do sistema de telecomunicações; (b) Sinal gerado no transmissor; (c) Sinal degradado pelo canal de transmissão.	14
Figura 2- Modelo Cenário de multipercuro.	16
Figura 3- Comportamento do desvanecimento lento (<i>slow fading</i>) e do desvanecimento severo (<i>fast fading</i>).	17
Figura 4- Filtro <i>tappet delay line</i>	22
Figura 5 - Filtro <i>lattice</i>	22
Figura 6- Exemplo de <i>DFE</i> baseado em <i>TDL</i>	23
Figura 7- Ilustração de canais sujeitos a nulos espectrais na frequência de 1kHz	26
Figura 8- Estrutura MLP.	27
Figura 9- Estrutura PPN.	27
Figura 10- Estrutura FLANN.	28
Figura 11- Rede Radial Basis Function (<i>RBFN</i>).	29
Figura 12- RNN modelo espaço-estado.	30
Figura 13- Comportamento dos coeficientes do canal de transmissão para desvio padrão de: a) 0; b) 0,1; c) 0,3; d) 0,5.	36
Figura 14- Exemplo de comportamento da variação do erro numa estrutura neural em função da variação dos pesos sinápticos.	39
Figura 15- Curva do BER vesus taxa de cruzamento para SNR fixo em 10dB.	43
Figura 16- Esquema que ilustra a utilização dos indivíduos na função de aptidão.	43
Figura 17- Sistema de transmissão digital.	45
Figura 18- Comparação de desempenho entre estruturas neurais diferentes com NL=0.	47
Figura 19- Comparação de desempenho entre estruturas neurais diferentes com NL=1.	47
Figura 20- Comparação de desempenho entre estruturas neurais diferentes com NL=2.	48
Figura 21- Comparação de desempenho entre estruturas neurais diferentes com NL=3.	48
Figura 22- Representação scatter plot: a) NL=0; b) NL=1; c) NL=2; and d) NL=3.	49
Figura 23- Comparação de desempenho entre estruturas neurais diferentes.	51
Figura 24- Comportamento dos coeficientes do canal de transmissão para desvio padrão de: a) 0; b) 0,1; c) 0,3; d) 0,5.	54
Figura 25- Modelo de neurônio utilizado para equalização.	55
Figura 26- Modelo do sistema utilizado.	56
Figura 27- (a) resposta no tempo e (b) pulso modelado pelo filtro do cosseno elevado.	57
Figura 28- Comparação de desempenho de quatro equalizadores com decisão realimentada (<i>DFE</i>)... ..	58
Figura 29- Desempenho do equalizador proposto sob diferentes condições do canal.	59

Sumário

1 INTRODUÇÃO	10
2 PROBLEMAS DE EQUALIZAÇÃO.....	14
2.1 PROBLEMAS DE EQUALIZAÇÃO	15
2.2 MODELOS DE CANAIS.....	18
2.2.1 Canais com Multipercurso Discreto.....	18
2.2.2 Canais com Multipercurso Contínuo.....	20
2.3 TIPOS DE EQUALIZADORES.....	21
3 EQUALIZADORES NEURAIIS	25
3.1 TIPOS DE EQUALIZADORES NEURAIIS.....	25
3.1.1 MLP	26
3.1.2 PPN.....	27
3.1.3 FLANN.....	28
3.1.4 RBF.....	29
3.1.5 RNN	30
3.2 ALGORITMOS DE OTIMIZAÇÃO.....	31
3.2.1 Algoritmos determinísticos	31
3.2.2 Algoritmos meta-heurísticos.....	33
3.3 MOTIVAÇÃO PARA O EMPREGO DE EQUALIZADORES NEURAIIS.....	34
4 ESCOLHA DA ESTRUTURA DO EQUALIZADOR NEURAL	38
4.1 ESCOLHA DO EQUALIZADOR NEURAL.....	40
4.2 MÉTODO PROPOSTO.....	41
4.2.1 Avaliação de desempenho.....	46
4.3 ESCOLHA DA ESTRUTURA NEURAL.....	50
5 EQUALIZAÇÃO DE CANAL VARIANTE UTILIZANDO ESTRUTURA NEURAL RECURSIVA	53
5.1 AVALIAÇÃO DE DESEMPENHO	57
6 CONCLUSÃO / TRABALHOS FUTUROS	60
REFERÊNCIAS BIBLIOGRÁFICAS.....	63

1 INTRODUÇÃO

Os sistemas de telecomunicações sem fio vêm evoluindo rapidamente desde o início do século. Quatro gerações móveis foram implantadas no Brasil em uma década e meia. Cada nova geração incorpora uma série de inovações tecnológicas, proporcionando taxas de transmissão de dados cada vez maiores, de forma a atender à demanda do mercado.

Alguns fatores conhecidos degradam a qualidade da comunicação, impondo sérios desafios à busca da maior eficiência espectral, dentre eles se destacando o ruído térmico, interferência intersimbólica e o desvanecimento. Uma das formas de mitigar esses problemas é o uso de equalizadores que, idealmente, implementam a função inversa do canal de transmissão, permitindo que o sinal recebido apresente mínimo erro em relação ao transmitido.

Equalizadores baseados em filtros adaptativos, cujos pesos são alterados em função da variação da função de erro, apesar de serem amplamente utilizados, não são capazes de lidar com não linearidades satisfatoriamente. A utilização de funções de ativação não lineares em equalizadores neurais resultou em ganhos de desempenho principalmente ao tratar canais não lineares (HAYKIN, 1994).

Estruturas neurais continuam sendo utilizadas como ferramentas para solução de problemas em diversas áreas, inclusive como estruturas de equalizadores. No entanto, o auge do estudo específico do tema ocorreu há algumas décadas, o que justifica a utilização de referências do século passado, no que concerne ao estudo da teoria de redes neurais. Portanto, são citadas, ao longo da dissertação, além de referências atuais, nas quais estruturas neurais são utilizadas para solução de problemas, também referências antigas que tratam da teoria de redes neurais.

Dois pontos determinantes no desempenho dos equalizadores neurais são o algoritmo de otimização, responsável pela atualização dos pesos neurais, e a estrutura neural. A escolha adequada desses elementos é uma tarefa complexa, que exige conhecimento prévio do problema e das estruturas neurais existentes.

Para dar prosseguimento ao projeto, foi necessária a escolha da estrutura neural adequada para resolver o problema de equalização proposto, no entanto, geralmente os trabalhos partem de uma estrutura predefinida, sem uma justificativa prévia para sua escolha. Nesse ponto foi notada a existência de uma lacuna na literatura científica: não foi identificado nenhum trabalho que aborde uma análise comparativa entre estruturas neurais aplicadas à

resolução de problemas específicos. A adaptação dinâmica da estrutura neural é realizada pelos algoritmos de *pruning* (XINGCHENG; YANQIN; PENGFEI, 2014) e pelo *Self Adaptive Growing Neural Network* (SAGNN) (BARAKAT et al, 2012). Esses métodos trabalham com a redução ou crescimento de estruturas neurais, baseados na relevância dos seus elementos (ligações neurais, neurônios e camadas) dentro da estrutura, com o objetivo de encontrar estrutura ideal para solução do problema. Contudo, essas abordagens normalmente partem de um determinado tipo de estrutura, geralmente MLP (*Multilayer Perceptrons*) ou RBF (*Radial Basis Function*), não considerando a possibilidade de utilização de outros tipos. O desenvolvimento de metodologia para comparar o potencial de diferentes tipos de estruturas, quando aplicadas ao problema de equalização, é um dos objetivos desse trabalho.

Burse, Yadav e Shrivastava (2010) descreveram quatro tipos de redes neurais usualmente aplicadas a problemas de equalização: *Multilayer Perceptrons* (MLP), *Functional Link Artificial Neural Network* (FLANN), *Radial Basis Function* (RBF) e *Recursive Neural Networks* (RNN). As três primeiras são classificadas como *Feedforward Neural Networks* (FNN) por possuírem apenas conexões diretas em suas estruturas. Por outro lado, estruturas RNN apresentam pelo menos uma conexão recursiva, o que pode gerar instabilidade. No entanto, conexões recursivas provêm capacidade para responder rapidamente a erros na saída, permitindo a esse tipo de estrutura lidar melhor com canais sujeitos a desvanecimento severo e aqueles que possuem nulos espectrais profundos¹. Além disso, RNNs são conhecidas por superarem o desempenho de FNNs tais como redes MLP e RBF (BACK; TSOI, 1991; KECHRIOTIS; ZERVAS; MANOLAKOS, 1994).

Atualmente, muitos pesquisadores têm aplicado, com sucesso, equalizadores baseados em redes neurais para tratar canais de comunicação sujeitos a desvanecimento (SIVAKUMAR, 2011; BARUAH; SARMA, 2011; DING; HIROSE, 2013, 2014; BHUYAN; SARMA, 2013). Esses trabalhos exemplificam como as redes neurais podem ser aproveitadas para equalizar tais tipos de canais, os quais representam o comportamento da maioria dos sistemas de comunicações móveis atuais.

Choi, Lima e Haykin (2005) e Corral, Ludwig e Lima (2010) analisaram sistemas envolvendo canais sujeitos a desvanecimento lento, propondo RNN para tratar esse tipo de distorção. O primeiro utiliza uma estrutura composta por três camadas, com três neurônios na camada escondida e 31 pesos sinápticos atualizados por algoritmo baseado em filtro de kalman, fatores responsáveis por elevar o custo computacional. O segundo utiliza uma

¹ Nulo espectral profundo é a atenuação severa imposta pelo canal de transmissão ao sinal em determinada faixa de frequência. A definição detalhada será vista no capítulo 3.

estrutura composta por apenas um neurônio, do qual os pesos são atualizados pelo algoritmo de Gauss-Newton (GN), capaz de produzir instabilidade de convergência. Essa instabilidade é preocupante quando se trata de canais com variações rápidas.

A maioria dos trabalhos científicos considera canais invariantes no tempo, contudo, os canais de comunicação móveis são variantes. Dentre os trabalhos que tratam de canais variantes, não foi encontrado nenhum que tratasse o problema do desvanecimento severo. Esse tipo de desvanecimento é comum, principalmente quando os dispositivos móveis estão em deslocamento.

Outro objetivo desse trabalho é indicar uma estrutura neural que, utilizando algoritmo de otimização confiável, seja capaz de tratar com menor erro possível, além das distorções em um canal sujeito a desvanecimento lento, sinais provenientes de canais com desvanecimento severo. Após a seleção de algumas estruturas mais indicadas na literatura científica para tratar problemas de classificação em tempo real, foi desenvolvida metodologia para comparar o potencial delas, utilizando algoritmo genético. Após a seleção da estrutura, o próximo passo foi escolha de um algoritmo de otimização que produzisse os melhores resultados.

A utilização conjunta de redes neurais e algoritmos genéticos data das últimas décadas do século passado, geralmente com o segundo suprindo alguma deficiência do primeiro. Dar suporte à escolha de parâmetros e da topologia, funcionar como algoritmo de treinamento e selecionar os dados para treinamento da rede neural são algumas das aplicações do algoritmo genético apontadas por Schaffer, Whitley e Eshelman (1992). Recentemente, Ning e Cai-yun (2010) combinaram o algoritmo genético ao *backpropagation* com a finalidade de atualizar os pesos neurais em uma rede utilizada para conversão de coordenadas geográficas. Peixian et al. (2011) aplicaram algoritmo genético para indicar os pesos iniciais em uma rede MLP (*Multilayer Perceptron*) utilizada na previsão de séries temporais. Apesar de existirem trabalhos combinando as duas teorias (redes neurais e algoritmos genéticos), geralmente eles são aplicados a redes MLP, sem considerar a existência de outros tipos de arquitetura neurais. No presente trabalho o algoritmo genético é utilizado para comparar o desempenho de diferentes tipos de estruturas neurais, além da MLP.

Visando alcançar as metas propostas, os seguintes objetivos específicos foram trilhados: compreensão os principais fatores que prejudicam o desempenho de canais de comunicação móveis; pesquisa dos modelos de canais de comunicações mais indicados para simulação computacional; seleção de estrutura, dentre as utilizadas na literatura para equalização de canais móveis, as mais eficientes com baixo custo computacional; desenvolvimento metodologia para comparar estruturas selecionadas para equalização de

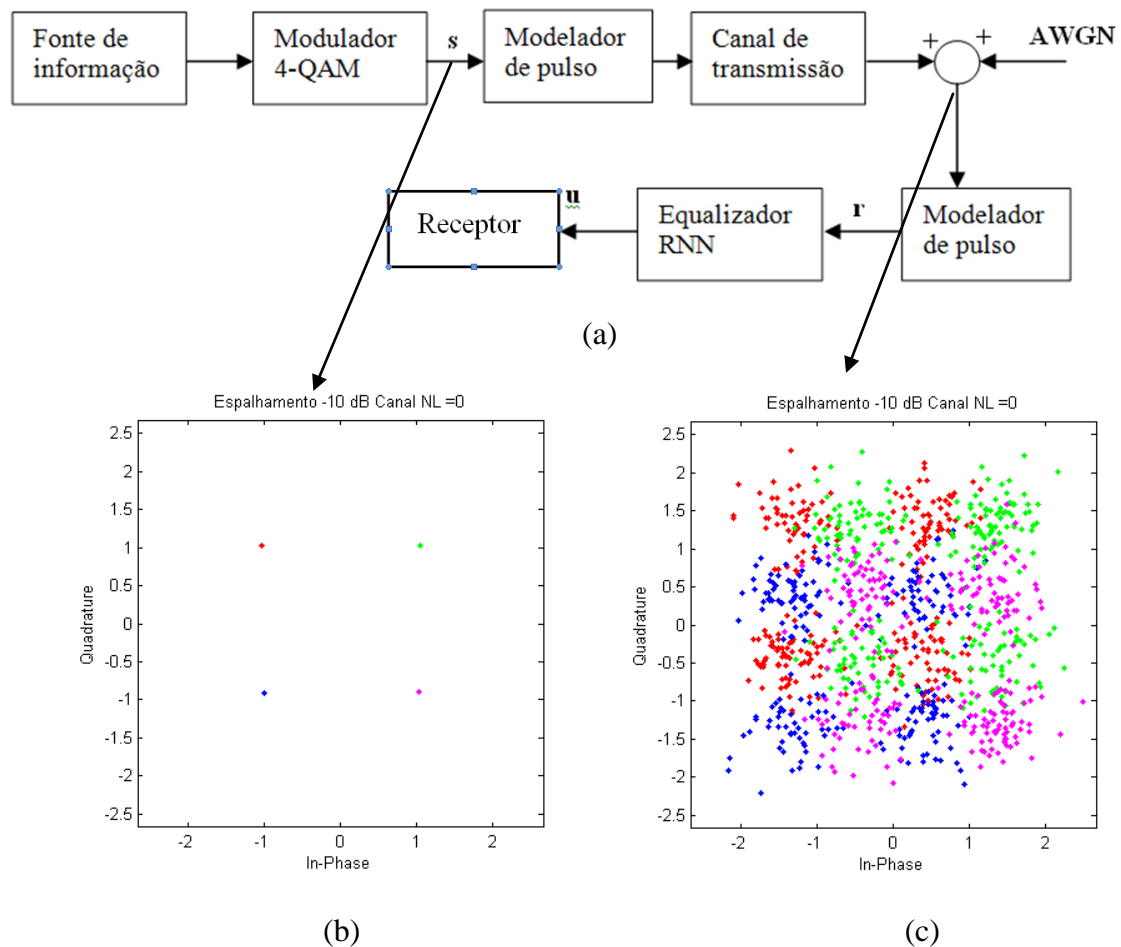
canais de comunicação; utilização de metodologia desenvolvida para encontrar a estrutura mais indicada para lidar com as distorções causadas por canal de simulação proposto; pesquisa de algoritmo de otimização adequado para atualização dos pesos da estrutura neural escolhida.

Essa dissertação está organizada da seguinte forma: no capítulo 2 são apresentados alguns conceitos necessários para o perfeito entendimento do trabalho, tais como, os principais elementos de um sistema de telecomunicações, a função básica e os mais comuns tipos de canais, as distorções as quais são acometidos esses tipos de canais e os modelos de canais utilizados em simulação; no capítulo 3 são descritos os principais tipos de equalizadores neurais, os algoritmos mais utilizados para atualização dos pesos de uma rede neural e a motivação para o uso de equalizadores neurais; no capítulo 4 é descrita a metodologia desenvolvida para a comparação de estruturas neurais. Um artigo contemplando a pesquisa desenvolvida nesse capítulo foi submetido à Revista IEEE Latino Americana; no capítulo 5 é mostrada a solução para o problema proposto nessa dissertação, além de comparar o resultado com outros trabalhos da literatura científica. Foi publicado no *IJCNS: International Journal of Computer Science and Network* (MOTA; LEAL; LIMA, 2014), artigo apresentando os resultados do trabalho realizado neste capítulo; e no capítulo 6 são descritas algumas conclusões e sugestões para futuros trabalhos na área.

2 PROBLEMAS DE EQUALIZAÇÃO

Um sistema de telecomunicações típico é composto, no mínimo, por transmissor ou fonte de informação, canal de comunicação e receptor, conforme pode ser observado na figura 1(a). O papel desses elementos é, respectivamente, gerar a informação a ser transmitida, servir de meio para o fluxo da informação e captar a informação recebida. Nos sistemas práticos, além dos três elementos já citados, alguns outros são inseridos de forma a aproveitar da melhor forma as características do canal de transmissão, fazendo com que a informação atinja o receptor com o mínimo erro. Dos elementos ilustrados na figura 1(a), três deles, modulador 4-QAM, modelador de pulso e equalizador, serão descritos com mais detalhes durante o desenvolvimento da dissertação.

Figura 1- (a) Modelo do sistema de telecomunicações; (b) Sinal gerado no transmissor; (c) Sinal degradado pelo canal de transmissão.



Fonte: elaborado pelo autor

O desempenho dos sistemas de comunicações móveis é degradado por ruído térmico, interferência intersimbólica (ISI) e desvanecimento. Um exemplo de sinal sujeito a esses tipos de degradação pode ser visualizado na figura 1. A figura 1(b) mostra os símbolos gerados antes de atravessarem o canal de transmissão e a figura 1(c) mostra os símbolos distorcidos no momento em que atingem o receptor, após atravessarem o canal de transmissão.

Os problemas causados pelos canais aos sinais de comunicações são mitigados com o uso de técnicas de equalização, ou seja, a informação degradada que atinge o receptor, tal qual visto na figura 1(c), deve ser corretamente classificada dentre uma das quatro opções vistas na figura 1(b). O mapeamento da informação recebida no receptor para um dos quatro símbolos gerados pelo transmissor é um trabalho realizado majoritariamente pelo equalizador.

Neste capítulo serão tratadas, na seção 2.1, as principais distorções a que estão sujeitos os canais de comunicação sem fio e a função do processo de equalização, na seção 2.2, são descritos os modelos de canais mais utilizados na literatura científica para simulação e, na seção 2.3, são discutidos os principais tipos de equalizadores.

2.1 PROBLEMAS DE EQUALIZAÇÃO

Equalização é um processo crítico quando se trata de canais de telecomunicações. Os dados que trafegam em redes sem fio estão sujeitos a distorções que podem inviabilizar o processo de comunicação. Os principais tipos de distorção são ruído térmico, interferência intersimbólica e desvanecimento.

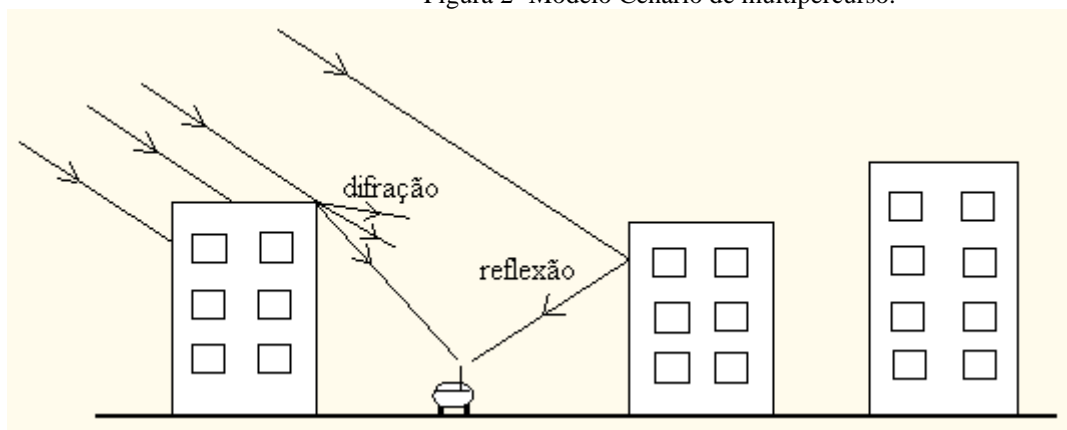
O ruído térmico é o nome dado ao fenômeno eletromagnético que surge devido ao movimento aleatório dos elétrons em um condutor, interno ou externo ao sistema de telecomunicações, e é proporcional à temperatura a qual o sistema está submetido (HAYKIN, 2004). Para efeito de simulação computacional o ruído térmico é geralmente representado pelo Ruído Branco com Distribuição Gaussiana (AWGN – *Additive white Gaussian noise*) e é inserido no receptor, onde provoca maiores danos, apesar de se manifestar também no transmissor.

A interferência intersimbólica (ISI – *Intersymbolic Interference*) consiste no dano causado pelo símbolo atual em símbolos subsequentes e é causado principalmente por multipercurso (HAYKIN, 1988). O sinal transmitido em determinado momento percorre, devido à refração, à difração e principalmente à reflexão, um percurso maior e atinge o receptor no mesmo momento do sinal transmitido em momento posterior, que percorre um

percurso direto ou com menos reflexões que o símbolo interferente (JERUCHIM; BALABAN; SHANMUGAN, 2000), conforme ilustrado na figura 2. Esse comportamento do canal pode ser visualizado por meio da função de transferência do modelo do canal, representada por meio da sua *Transformada z*, como exemplificado na equação 2.1, na qual o número de coeficientes representa a quantidade de percursos relevantes e os seus valores determinam a atenuação sofrida pelo sinal ao atravessar cada percurso. Outro fator que pode provocar ISI é a limitação do sinal em determinada banda de frequência. A limitação em frequência provoca espalhamento no domínio do tempo, podendo fazer com que o símbolo atual interfira no subsequente.

$$H = 0,3482 + 0,8704z^{-1} + 0,3482z^{-2} \quad (2.1)$$

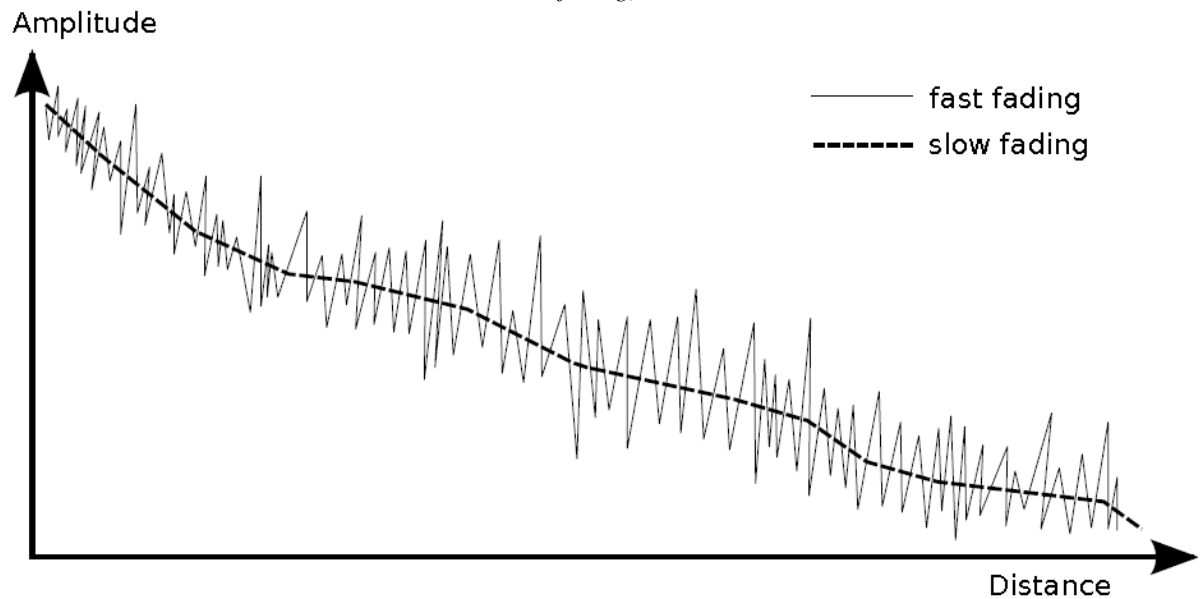
Figura 2- Modelo Cenário de multipercurso.



Fonte: <www.teleco.com.br>.

O desvanecimento é a variação da atenuação sofrida pelo sinal ao percorrer o caminho entre o transmissor e o receptor. Tal variação pode ser atribuída às mudanças sofridas pelo canal de transmissão (PROAKIS, 1995) devido, principalmente, à movimentação da estação móvel. O desvanecimento lento consiste numa variação gradual da amplitude do sinal que atinge o receptor, causada pelo aparecimento de obstáculos que bloqueiam a linha de visada, o que geralmente pode ser mitigado pelo controle de potência do sistema de comunicações. O desvanecimento rápido usualmente é causado pela interferência destrutiva em canais sujeitos à multipercurso, podendo resultar também de variações bruscas do sinal devido ao surgimento de obstruções que bloqueiam abruptamente a linha de visada (JERUCHIM; BALABAN; SHANMUGAN, 2000). A comparação entre desvanecimentos lento (*slow fading*) e severo (*fast fading*) pode ser vista na figura 3.

Figura 3- Comportamento do desvanecimento lento (*slow fading*) e do desvanecimento severo (*fast fading*).



Fonte: < <http://www.intechopen.com/books/matlab-a-fundamental-tool-for-scientific-computing-and-engineering-applications-volume-2/mobile-radio-propagation-prediction-for-two-different-districts-in-mosul-city>>.

Para lidar com as distorções descritas é comum o uso de equalizadores, cuja função básica é mitigar as distorções causadas pelo canal de comunicações, permitindo que o sinal transmitido, após atravessar o conjunto canal de transmissão e equalizador, seja decodificado no receptor com o mínimo de erro.

O problema de equalização consiste, principalmente, na classificação dos símbolos recebidos após atravessarem o canal de transmissão, tais quais mostrados na figura 1(c), numa das quatro opções geradas pelo transmissor ($1+j$, $1-j$, $-1+j$ ou $-1-j$), conforme figura 1(b). Para tanto, o equalizador é posicionado no receptor, onde o sinal já se encontra distorcido pelos efeitos destrutivos do canal de transmissão.

Alguns modelos de canais de comunicação são utilizados para simular o desempenho de equalizadores. Os canais de comunicações móveis sem fio geralmente são muito complexos, dificultando o desenvolvimento de uma análise matemática adequada (GUIMARÃES; PINTO, 2001). No entanto, existem alguns modelos que serão explorados na seção 2.2 e que conseguem simular as distorções mais relevantes. Os coeficientes da função de transferência desses canais, representada por sua *Transformada z*, indicam o multipercorso discreto que, se somado a uma parcela variável, pode simular também o desvanecimento (LING; PROAKIS, 1985).

2.2 MODELOS DE CANAIS

Os canais sujeitos a multipercurso podem ser classificados como: canais com multipercurso discreto, nos quais uma quantidade finita de caminhos é percorrida pelo sinal até atingir o receptor; e canais com multipercurso contínuo, no qual existe uma quantidade não mensurável de caminhos percorridos pelo sinal até atingir o receptor. Canais com multipercurso discreto são comuns em áreas predominantemente rurais, onde a urbanização é distribuída, resultando numa quantidade limitada de percursos. Canais com multipercurso contínuo são aplicáveis a regiões altamente urbanizadas, nas quais uma grande quantidade de obstruções gera inúmeras possibilidades para reflexão, refração e difração do sinal transmitido (JERUCHIM; BALABAN; SHANMUGAN, 2000).

Devido à sua relativa simplicidade, o modelo de canal com multipercurso discreto tem sido amplamente utilizado para simulação de desempenho de sistemas de transmissão digital (GUIMARÃES; PINTO, 2001). Um exemplo é o modelo estabelecido pela TIA (*Telecommunications Industry Association*) nas especificações dos padrões americanos IS-136 (TIA/EIA Standard 136-270-B, 2000) para telefonia celular de segunda geração. Este modelo também foi proposto para avaliação de alguns serviços previstos no padrão UWC-136, de terceira geração. Outro exemplo é a recomendação ITU-R M.1225 (ITU-R M.1225 Recommendation, 1997) para a terceira geração de sistemas móveis (IMT-2000) que determina que a resposta do canal móvel deva ser simulada de acordo com o modelo de multipercurso discreto, com parâmetros definidos em função do ambiente de propagação considerado.

2.2.1 Canais com Multipercurso Discreto

Uma forma simples de se representar canais sujeitos à multipercurso discreto é por meio da equação 2.2.

$$y(t) = \sum_n a_n(t)s(t - \tau_n(t)) \quad (2.2)$$

Onde $s(t)$ é o sinal transmitido ao atingir o receptor, $a_n(t)$ é a atenuação sofrida pelo sinal em determinado percurso, n é o número de caminhos percorridos pelo sinal e $\tau_n(t)$ é o atraso sofrido pelo sinal em determinado percurso.

O canal mostrado na equação 2.2 geralmente é representado por meio da sua transformada z , conforme equação 2.3.

$$Y(z) = \sum_n a_n z^{-\tau_n} \quad (2.3)$$

Muitos autores utilizam canais desse tipo com o objetivo de testar e comparar o desempenho de equalizadores baseados em redes neurais. O trabalho desenvolvido em Patra (1999) é importante no que concerne à metodologia desenvolvida nessa dissertação devido a três motivos principais: ele também trata sinal gerado por modulador 4-QAM; três tipos equalizadores neurais são utilizados para tratar as distorções causadas pelo canal de transmissão; e utiliza canais sujeitos à multipercurso discreto para simulação, conforme mostrado nas equações 2.4 a 2.9.

$$CH = 1: 1 \quad (2.4)$$

$$CH = 2: 0,447 + 0,8894z^{-1} \quad (2.5)$$

$$CH = 3: 0,209 + 0,995z^{-1} + 0,209z^{-2} \quad (2.6)$$

$$CH = 4: 0,260 + 0,930z^{-1} + 0,260z^{-2} \quad (2.7)$$

$$CH = 5: 0,304 + 0,903z^{-1} + 0,304z^{-2} \quad (2.8)$$

$$CH = 6: 0,341 + 0,876z^{-1} + 0,341z^{-2} \quad (2.9)$$

CH=1 representa um canal que não está sujeito a multipercurso, sendo o ruído térmico a única fonte de distorção. CH=2 corresponde a um canal de fase não mínima (CHEN; GIBSON; COWAN, 1990), ou seja, um canal que, se comparado a um canal de fase mínima com a mesma contribuição em amplitude, proporciona maior influência na variação da fase, com dois percursos, sendo um refletido e um direto. Os CH=3 a 6 representam canais sujeitos a um percurso direto e dois refletidos.

No presente trabalho, foram adicionados coeficientes variáveis, conforme proposto por Ling e Proakis (1985) para simular o desvanecimento, a um canal sujeito a multipercurso discreto com três coeficientes, sendo um percurso direto e dois refletidos. O canal representado genericamente pela função de transferência H , apresentado na equação 2.10, com $a_0(t) = a_1(t) = a_2(t) = 0$ e $c_0 = c_2 = 0,3482$, $c_1 = 0,8704$, corresponde a um canal invariante recomendado pelo UIT (União Internacional de Telecomunicações) para testar o

desempenho de equalizadores, que é usualmente utilizado na literatura científica (CHOI; LIMA; HAYKIN, 2005) e (CORRAL; LUDWIG; LIMA, 2010).

$$H = [c_0 + a_0(t)] + [c_1 + a_1(t)]z^{-1} + [c_2 + a_2(t)]z^{-2} \quad (2.10)$$

Trata-se de um canal linear de fase não mínima que causa severas distorções no sinal transmitido. Os coeficientes variáveis são gerados de forma independente, conforme proposto por Ling e Proakis (1985), passando um sinal de ruído branco através do filtro passa baixa de Butterworth de segunda ordem com frequência de corte normalizada de 0,1. Os coeficientes da função de transferência têm distribuição Gaussiana com média c_i e variância proporcional ao desvio padrão (σ) do ruído branco. Os valores dos coeficientes são proporcionais a σ , i.e., um valor alto de σ representa um canal com desvanecimento severo.

2.2.2 Canais com Multipercurso Contínuo

As distorções causadas por multipercurso contínuo seguem a distribuição de Rayleigh ou de Ricean. Nos casos em que não existe uma portadora dominante atingindo o receptor, o efeito do multipercurso pode ser modelado pela distribuição de Rayleigh. Nos casos em que uma portadora dominante existe, tal qual em linha de visada, a distribuição de Ricean representa melhor tal efeito (JERUCHIM; BALABAN; SHANMUGAN, 2000).

2.2.2.1 Distribuição de Rayleigh

Em sistemas de comunicações móveis, é comum a situação em que não existe linha de visada entre o transmissor e o receptor, ou seja, qualquer sinal que atinja o receptor proveniente do transmissor percorre caminhos indiretos, por meio de reflexão, refração ou difração em obstáculos. Nesses casos, o sinal recebido é descrito estatisticamente pela função densidade de probabilidade (pdf – *probability density function*) de Rayleigh, apresentada na equação 2.11, conforme descrito por Papoulis e Pillai (2001).

$$f(x; \sigma) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} \quad (2.11)$$

Onde σ^2 corresponde à variância do sinal.

2.2.2.2 Distribuição de Ricean

Nos casos em que, além das componentes do sinal refletidos, refratados e difratados que atingem o receptor, existe também uma parcela que corresponde ao sinal direto, ou seja, aquele que percorre o caminho direto que liga transmissor ao receptor, a composição que atinge o receptor é descrita estatisticamente pela função densidade de probabilidade (pdf – *probability density function*) de Ricean, apresentada na equação 2.12, conforme descrito por Proakis (1995).

$$f(x|v; \sigma) = \frac{x}{\sigma^2} e^{-(x^2+v^2)/2\sigma^2} I_0\left(\frac{xv}{\sigma^2}\right) \quad (2.12)$$

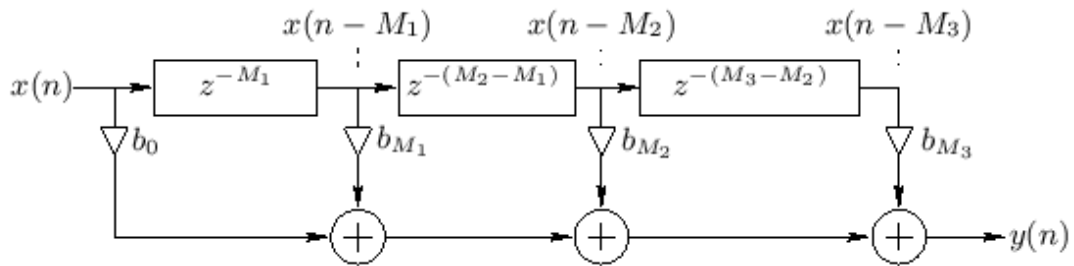
Onde σ^2 corresponde à variância do sinal, $v \geq 0$ é o parâmetro de não centralidade e I_0 é a função modificada de Bessel de ordem zero.

2.3 TIPOS DE EQUALIZADORES

A tarefa primordial de um equalizador é representar a função inversa do canal de transmissão, idealmente anulando as distorções causadas pelo mesmo. Para tal fim, os equalizadores são instalados no receptor do sistema de comunicações, conforme mostrado na figura 1.

Equalizadores baseadas em filtros adaptativos são comumente utilizados com o objetivo de tratar as distorções causadas pelos canais de comunicação. Esses filtros são compostos por pesos variáveis, que são ajustados para minimizar o erro entre o sinal transmitido e o recebido (HAYKIN, 1996). Quando tratamos de aprendizado supervisionado, uma sequência de treinamento pré-definida é enviada pelo transmissor, em intervalos de tempo regulares, para que os pesos sejam recalculados por algum algoritmo de otimização. Quando o aprendizado não é supervisionado, o equalizador tenta extrair a função de transferência do canal de transmissão baseado em características do sinal recebido.

O equalizador baseado em filtro adaptativo mais utilizado é o *tapped delay line (TDL)* ou filtro transversal, mostrado na figura 4. Esse tipo de equalizador gera uma série de versões atrasadas do sinal de entrada, as quais são multiplicadas por pesos e somadas para gerarem o sinal de saída (HAYKIN, 1996). Os pesos variam de acordo com o treinamento realizado.

Figura 4- Filtro *tapped delay line*.

Fonte: < https://crma.stanford.edu/~jos/lumped/Example_Tapped_Delay_Line.html>.

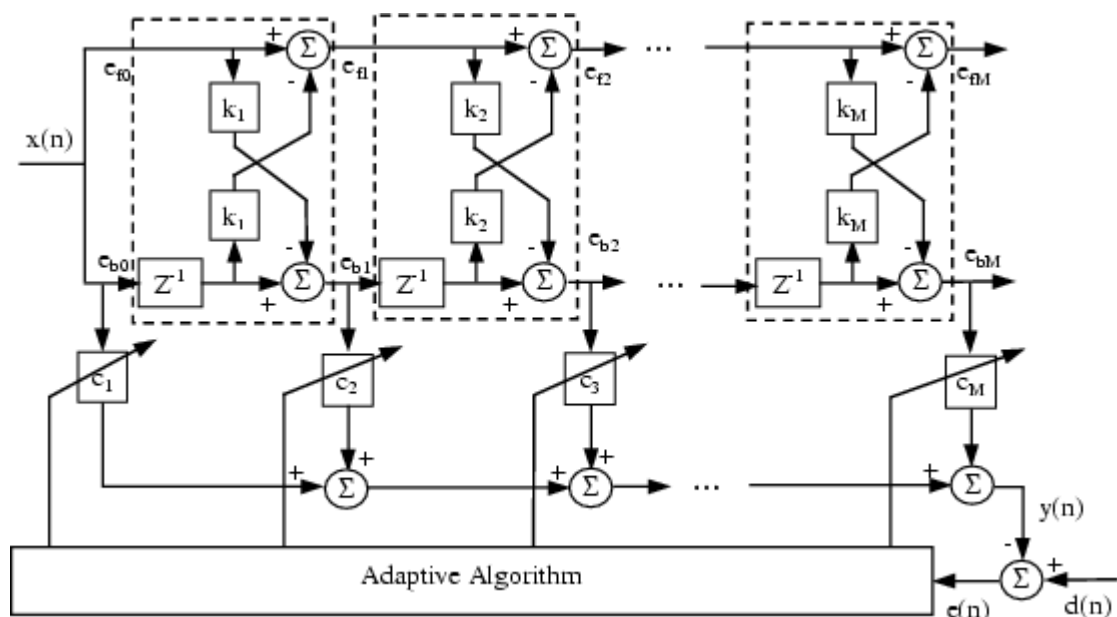
A função de transferência do equalizador é representada por:

$$H(z) = b_0 + b_{M_1}z^{-M_1} + b_{M_2}z^{-M_2} + b_{M_3}z^{-M_3} \quad (2.13)$$

Onde b_0 , b_{M_1} , b_{M_2} e b_{M_3} são os pesos variáveis, M_1 , M_2 e M_3 são os atrasos aos quais o sinal de entrada é submetido.

Além do *TDL*, a estrutura conhecida como *Lattice Predictor* também é abordada em Haykin (1996) como um tipo de filtro adaptativo.

A estrutura de um filtro do tipo *lattice* é composta por módulos ou estágios, os quais possuem uma entrada direta e_{fm} e uma realimentada e_{bm} . A entrada de cada estágio é igual a saída do anterior e as entradas do primeiro são o sinal de entrada do filtro (HAYKIN, 1996), conforme mostrado na figura 5.

Figura 5 - Filtro *lattice*

Fonte: < <http://www.dspalgorithms.com/aspt/asptnode23.html>>.

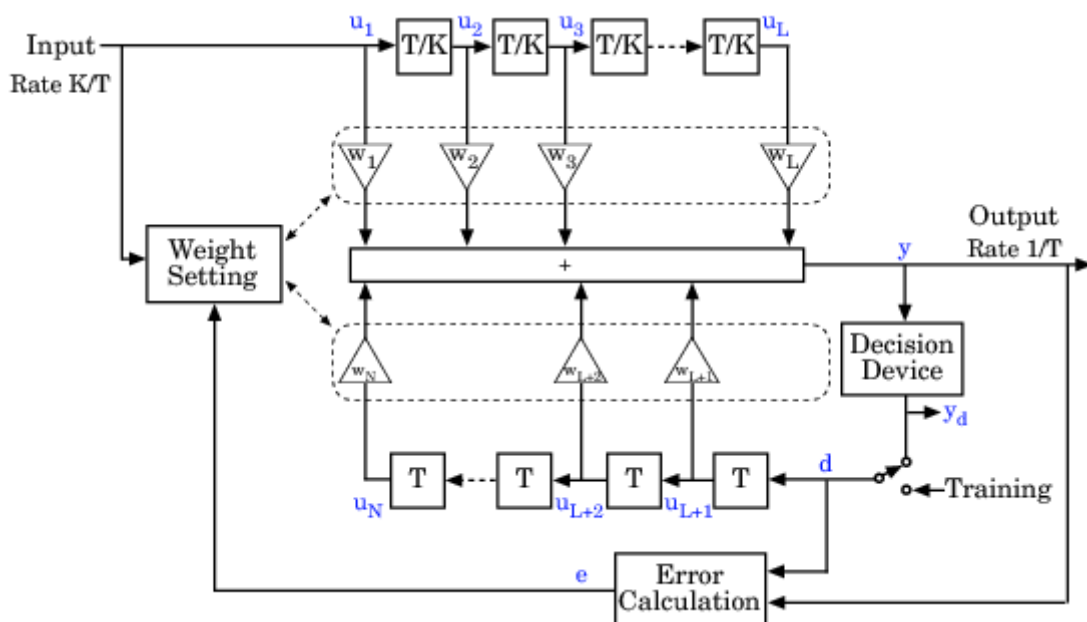
O m -ésimo estágio do *lattice predictor* é descrito por um par de relações entrada-saída, conforme equações 2.14 e 2.15.

$$e_{fm+1}(n) = e_{fm}(n) - k_{m+1}e_{bm}(n - 1) \quad (2.14)$$

$$e_{bm+1}(n) = e_{bm}(n - 1) - k_{m+1}e_{fm}(n) \quad (2.15)$$

As duas estruturas apresentadas são exemplos de filtros *FIR* (*Finite Impulse Response*), cuja estrutura contém somente percursos diretos. Qualquer uma das duas pode servir como base para uma estrutura conhecida como *DFE* (*Decision Feedback Equalizer*). Um *DFE* é um filtro do tipo *IIR* (*Infinite Impulse Response*), no qual a saída proveniente do filtro direto, após passar pelo processo decisório, é realimentado na entrada ou em algum outro ponto intermediário do mesmo. Um *DFE* baseado em *TDL* é mostrado na figura 6.

Figura 6- Exemplo de *DFE* baseado em *TDL*.



Fonte :<<http://www.mathworks.com/help/comm/ug/equalization.html>>.

Um filtro *IIR* se diferencia de um filtro *FIR* pelo fato de apresentar pelo menos um caminho realimentado. A realimentação é capaz de tornar o filtro instável, por isso, os coeficientes dos ramos realimentados devem ser cuidadosamente escolhidos. Por outro lado, o fato dos filtros *FIR* não apresentarem realimentação, os torna inerentemente estáveis. Por esse motivo, esse tipo de filtro é bastante utilizado (HAYKIN, 1996). Em problemas de

equalização, estruturas com realimentação apresentam melhor desempenho, tornando justificável o risco da instabilidade.

A utilização de funções de ativação não lineares em equalizadores neurais resultou em ganhos de desempenho principalmente ao tratar canais não lineares (HAYKIN, 1994). Os tipos de equalizadores neurais mais utilizados na literatura serão discutidos no próximo capítulo.

3 EQUALIZADORES NEURAIIS

Equalizadores associados a estruturas neurais continuam sendo objetos de trabalhos na literatura para tratar canais de comunicação sujeitos a desvanecimento (SIVAKUMAR, 2011; BARUAH; SARMA, 2011; DING; HIROSE, 2013, 2014; BHUYAN; SARMA, 2013). A maioria deles emprega alguma arquitetura neural, escolhida de forma *ad hoc*, combinada com um algoritmo de otimização determinístico, com a finalidade de atualizar os pesos sinápticos. Neste capítulo serão abordados, na seção 3.1, os principais tipos de equalizadores neurais e, na seção 3.2, os algoritmos de otimização mais utilizados na literatura científica para atualização dos pesos neurais. Na seção 3.3 serão expostas as motivações para utilização de equalizadores neurais.

3.1 TIPOS DE EQUALIZADORES NEURAIIS

As primeiras arquiteturas neurais utilizadas consistiam de arranjos contendo diversos perceptrons dispostos em camadas que deram origem as chamadas redes MLP (*Multilayer Perceptron*). Esse tipo de rede apresenta bom desempenho em tarefas de classificação, são simples de implementar, no entanto, por necessitarem de um longo tempo de treinamento (BURSE; YADAV; SHRIVASTAVA, 2010), o seu aproveitamento em aplicações em tempo real, tal qual equalização, é limitado.

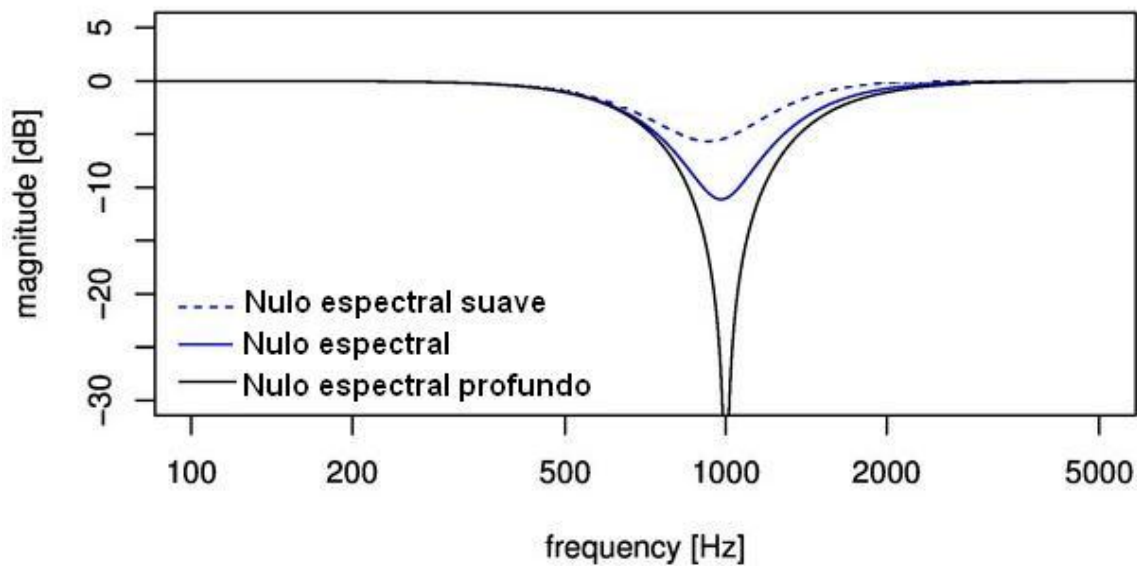
Com o objetivo de minimizar a complexidade das redes MLP foram concebidas estruturas nas quais as camadas escondidas são substituídas por mapeamento não linear, responsável por expandir as entradas, utilizando funções para combiná-las. Dentre elas estão as estruturas PPN (*Polynomial Perceptron Network*) e as FLANN (*Functional Link Artificial Neural Network*).

Redes RBF (*Radial Basis Function*) utilizam funções de ativação do tipo radial, como a função Gaussiana, que lhes confere um grande potencial em problemas de classificação. Apesar disso, as escolhas do centro e do raio da função de ativação tendem a elevar o seu custo computacional.

As RNN (*Recursive Neural Network*) implementam ramos de realimentação em sua estrutura, o que as habilita a utilizar a saída atual para influenciar nas decisões futuras. Essa característica permite a esse tipo de rede lidar bem com canais com nulos espectrais profundos (KECHRIOTIS; ZERVAS; MANOLAKOS, 1994).

A existência de nulos espectrais é comum nos canais de transmissão dos sistemas de telecomunicações. Esse tipo de problema consiste na atenuação imposta em determinadas faixas de frequência, pelo canal ao sinal. Na figura 7 podem ser observados três canais com nulo espectral na frequência de 1kHz sendo, o que apresenta maior atenuação, representado em linha contínua preta na figura, dito conter nulo espectral profundo. Esse tipo de canal gera distorções severas aos símbolos transmitidos.

Figura 7- Ilustração de canais sujeitos a nulos espectrais na frequência de 1kHz



Fonte :< <http://rtaylor.sites.tru.ca/category/modeling/>>, editada pelo autor.

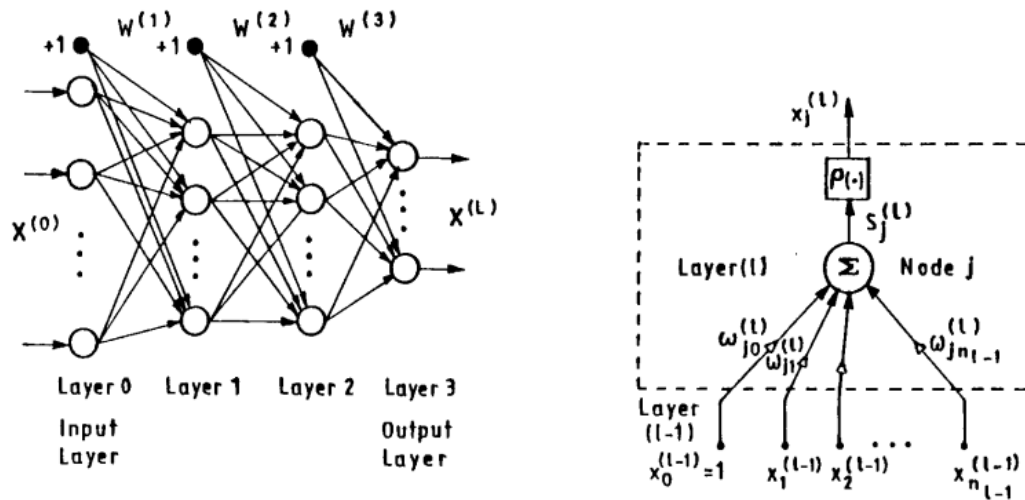
Conforme sugerido por Corral, Ludwig e Lima (2010) e pelo resultado desse trabalho, a rede neural ideal para o problema de equalização deve utilizar poucos neurônios, o que deve ser interpretado como uma quantidade mínima de pesos neurais para serem atualizados, e ainda assim ser capaz de classificar com menor erro possível.

A seguir serão descritas as arquiteturas mais utilizadas em equalizadores inspirados em redes neurais e que serão utilizadas ao longo da dissertação.

3.1.1 MLP

Uma rede MLP (*Multilayer Perceptron*) é um tipo de *feedforward neural network* (FNN) na qual neurônios são dispostos em mais de uma camada, conforme figura 8, cujo algoritmo de treinamento mais utilizado é o *Back-Propagation* (BP).

Figura 8- Estrutura MLP.



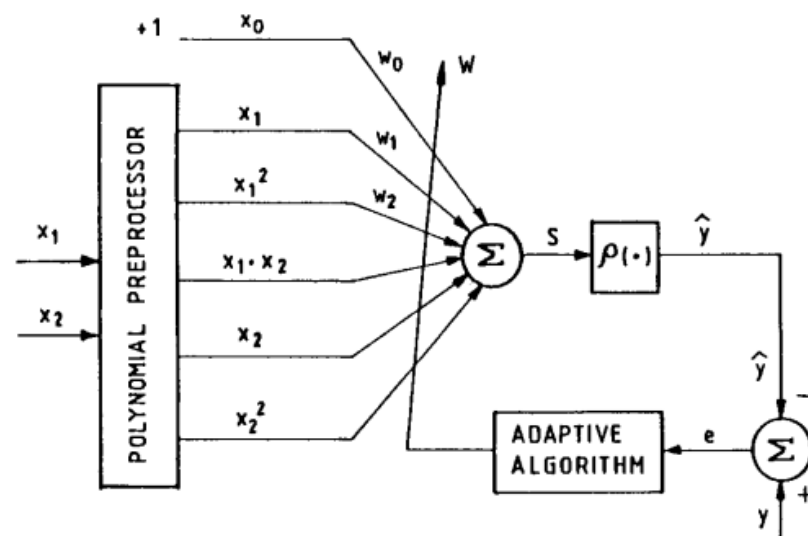
Fonte : (PATRA, 1999).

Essa estrutura é muito utilizada em equalizadores adaptativos por ser simples de implementar (BURSE; YADAV; SHRIVASTAVA, 2010).

3.1.2 PPN

As redes polinomiais de perceptrons (PPN) têm como base o teorema da aproximação de Weierstrass, que declara que qualquer função contínua dentro de um intervalo fechado pode ser aproximada, dentro desse intervalo, por polinômios (PATRA, 1999) e é estruturada conforme mostrado na figura 9.

Figura 9- Estrutura PPN.



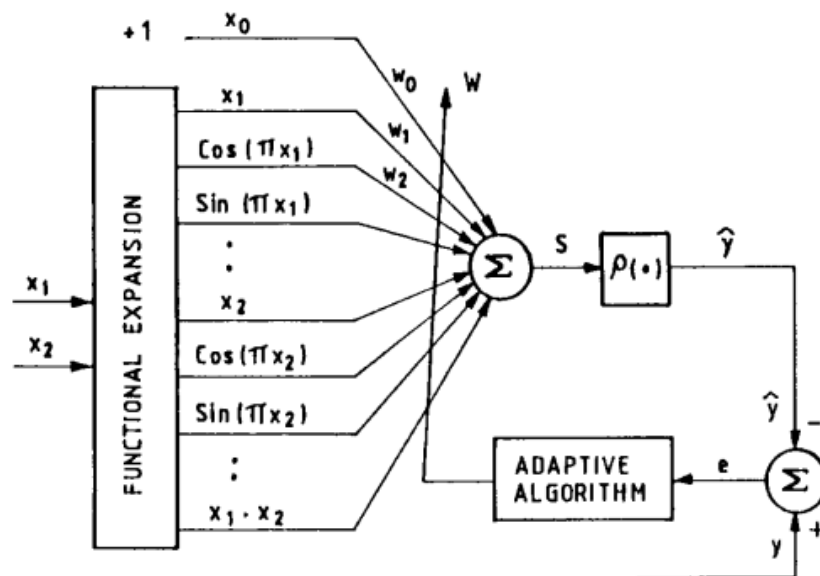
Fonte : (PATRA, 1999).

Para esse tipo de estrutura a dimensão do padrão de entrada é expandida por meio de expansão polinomial, tarefa realizada pelo bloco *POLYNOMIAL PREPROCESSOR* da figura 9, na qual ordens mais altas e produtos cruzados dos termos da entrada original são utilizados. Além do bloco de expansão polinomial, um PPN apresenta uma estrutura bem simples, possuindo apenas de 1 neurônio, com alta taxa de convergência e que utiliza menor carga computacional se comparada à estrutura MLP (PATRA, 1999).

3.1.3 FLANN

Uma Rede Neural Artificial com Ligação Funcional (FLANN) é uma estrutura que também contém apenas 1 neurônio. O padrão inicial de entrada é expandido em funções linearmente independentes, tarefa realizada pelo bloco *FUNCTIONAL EXPANSION* da figura 10, que usam como argumento os termos de entrada e então essas funções são avaliadas como argumento do neurônio que gera a saída.

Figura 10- Estrutura FLANN.



Fonte : (PATRA, 1999).

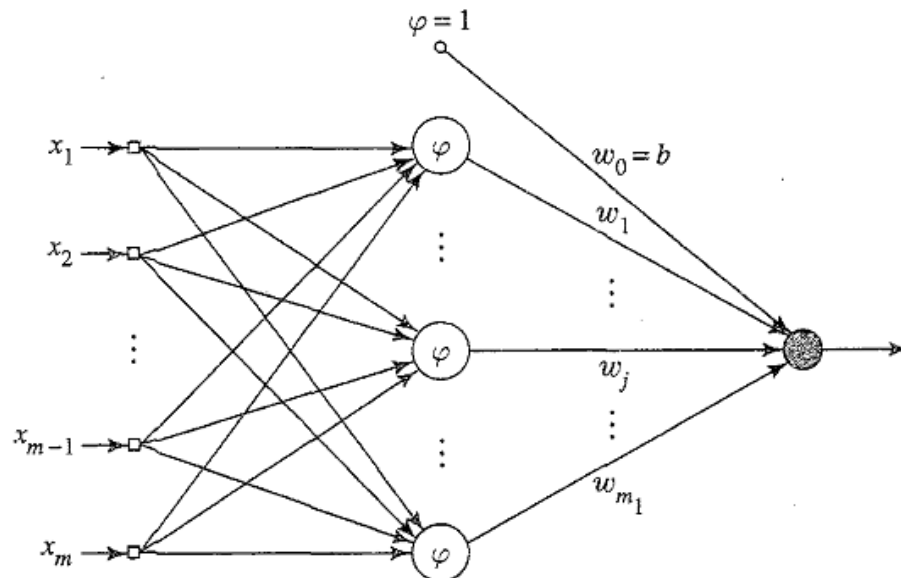
De acordo com Patra (1999), uma estrutura FLANN apresenta menor custo computacional e taxa de convergência mais rápida que uma rede MLP, devido à utilização de somente um neurônio.

3.1.4 RBF

Outra estrutura do tipo de *feedforward neural network (FNN)* é a *RBF (Radial basis function)* que trata o problema de aproximação em espaço de alta dimensão. Nesse tipo de rede, aprender equivale a encontrar uma função multidimensional que provê o melhor ajuste para os dados do treinamento (ALBU et al., 1998).

As redes do tipo RBF possuem duas camadas, conforme mostrado na figura 11, na qual a camada escondida contém n neurônios que calculam a distância Euclidiana entre um vetor centro c_i e o vetor de entrada y , cujo resultado é inserido em uma função não-linear. Normalmente a função Gaussiana $\varphi_i = \exp\left(-\frac{\|y-c_i\|^2}{r_i^2}\right)$ é escolhida, onde r_i é o desvio padrão.

Figura 11- Rede Radial Basis Function (RBFN).



Fonte : (HAYKIN, 1994) .

A saída é obtida através do mapeamento:

$$f(y) = \sum_{i=1}^n w_i \varphi_i \quad (3.1)$$

Onde w_i são os pesos neurais.

A desvantagem desse tipo de rede é a necessidade da definição dos vetores centro c_i e desvio padrão r_i , sendo i o número de neurônios da camada escondida. A definição desses parâmetros eleva o custo computacional desse tipo de estrutura, podendo torná-la inviável

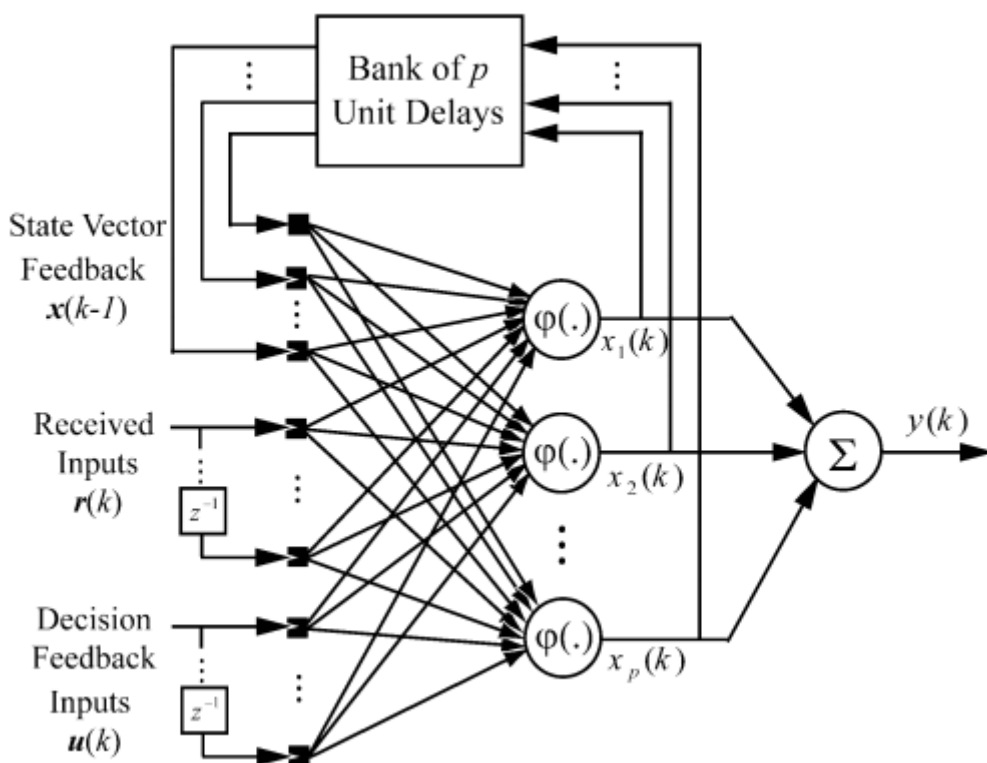
para aplicações em tempo real. Além disso, uma escolha equivocada desses parâmetros pode tornar a rede ineficiente para a tarefa de classificação.

3.1.5 RNN

Equalizadores que utilizam redes neurais recursivas (*RNN*) têm sido propostos para equalização adaptativa de canais lineares e não lineares, invariantes (KECHRIOTIS; ZERVAS; MANOLAKOS, 1994) e variantes (CHOI; LIMA; HAYKIN, 2005) no tempo. Essas redes possuem realimentação, propriedade que as torna atrativas para equalização de canais não lineares com nulos espectrais profundos (KECHRIOTIS; ZERVAS; MANOLAKOS, 1994).

As RNN podem ser configuradas em um grande número de arquiteturas, que podem variar devido ao número de camadas, número de perceptrons por camada, operações entre sinais de entrada e realimentados, característica da realimentação, entre outros. Na figura 12 é apresentada a arquitetura utilizada por Choi, Lima e Haykin (2005), de forma a ilustrar esse tipo de equalizador.

Figura 12- RNN modelo espaço-estado.



Fonte : (CHOI; LIMA; HAYKIN, 2005).

O comportamento desse modelo é descrito por

$$\mathbf{x}(k) = \boldsymbol{\varphi}(\mathbf{x}(k-1), \mathbf{u}(k), r(k)) \quad (3.2)$$

$$\mathbf{y}(k) = \mathbf{W}\mathbf{x}(k) \quad (3.3)$$

Onde $\boldsymbol{\varphi}(\cdot)$ é a função de ativação não linear da camada escondida e \mathbf{W} é o vetor de pesos sinápticos.

Uma RNN com um número reduzido de perceptrons (2 ou 3) é capaz de modelar a função inversa do canal de comunicações com desempenho superior às estruturas de equalização tradicionais (KECHRIOTIS; ZERVAS; MANOLAKOS, 1994).

3.2 ALGORITMOS DE OTIMIZAÇÃO

Definida a estrutura de um equalizador neural, a determinação dos pesos da rede é um problema de otimização. Os algoritmos de otimização são classificados como determinísticos e estatísticos (HOLTZ, 2005). Os determinísticos geralmente empregam busca por meio de derivada, o que, na maioria dos casos, conduz a solução a um mínimo local. Entre os métodos estatísticos, que geram e utilizam variáveis aleatórias, estão os algoritmos meta-heurísticos. Eles empregam métodos nos quais a busca é extrapolada para além da vizinhança do ponto, o que aumenta a probabilidade de se encontrar o mínimo global. Tendo em vista que, em geral, os algoritmos determinísticos são mais rápidos e os meta-heurísticos conseguem melhores resultados, existe um compromisso entre desempenho e custo computacional.

3.2.1 Algoritmos determinísticos

A utilização de algoritmos de otimização determinísticos muitas vezes não resulta na melhor solução para problemas multimodais, como os de equalização, por dependerem do ponto de partida (HOLTZ, 2005). No entanto, das simulações realizadas no decorrer do trabalho foi possível verificar que o mínimo global muitas vezes é desnecessário, tendo em vista que uma solução mínima local pode ser suficiente. Além disso, eles são largamente utilizados por inserirem menor custo computacional ao sistema, preenchendo um requisito básico para esses tipos de problema. A seguir são descritos alguns dos algoritmos

determinísticos mais utilizados na literatura científica em equalizadores baseados em redes neurais.

3.2.1.1 Método do Gradiente

O algoritmo é formalmente descrito por

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \eta \mathbf{g}(n) \quad (3.4)$$

Onde η é uma constante positiva chamada *taxa de aprendizado*, e $\mathbf{g}(n) = \nabla \xi(\mathbf{w})$ é o vetor gradiente da função erro avaliada no ponto $\mathbf{w}(n)$, cujo valor mínimo se deseja obter. Entre as iterações n e $n+1$ o algoritmo aplica a correção.

$$\Delta \mathbf{w}(n) = \mathbf{w}(n+1) - \mathbf{w}(n) = -\eta \mathbf{g}(n) \quad (3.5)$$

Esse algoritmo converge para a solução ótima \mathbf{w}^* lentamente e a taxa de aprendizagem tem influência significativa no comportamento da convergência.

3.2.1.2 Método de Gauss-Newton

É aplicável para função de custo que expressa o somatório dos quadrados dos erros:

$$\xi(\mathbf{w}) = \frac{1}{2} \sum_{i=1}^n e^2(i) \quad (3.6)$$

Dado o ponto de operação $\mathbf{w}(n)$, linearizada a dependência de $e(i)$ em \mathbf{w} , escrevendo:

$$e'(i, \mathbf{w}) = e(i) + \left[\frac{\partial e(i)}{\partial \mathbf{w}} \right]_{\mathbf{w}=\mathbf{w}(n)}^T (\mathbf{w} - \mathbf{w}(n)), \quad i = 1, 2, \dots, n \quad (3.7)$$

Ou, usando a notação matricial:

$$\mathbf{e}'(n, \mathbf{w}) = \mathbf{e}(n) + \mathbf{J}(n)(\mathbf{w} - \mathbf{w}(n)) \quad (3.8)$$

Onde $\mathbf{J}(n)$ é a matriz Jacobiana m -por- n de $\mathbf{e}(n)$:

$$\mathbf{J}(n) = \begin{bmatrix} \frac{\partial e(1)}{\partial w_1} & \frac{\partial e(1)}{\partial w_2} & \dots & \frac{\partial e(1)}{\partial w_m} \\ \frac{\partial e(2)}{\partial w_1} & \frac{\partial e(2)}{\partial w_2} & \dots & \frac{\partial e(2)}{\partial w_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e(n)}{\partial w_1} & \frac{\partial e(n)}{\partial w_2} & \dots & \frac{\partial e(n)}{\partial w_m} \end{bmatrix} \quad (3.9)$$

O jacobiano é a transposta da matriz gradiente $\nabla \mathbf{e}(n)$, onde:

$$\nabla \mathbf{e}(n) = [\nabla e(1), \nabla e(2), \dots, \nabla e(n)] \quad (3.10)$$

A atualização do vetor de pesos $\mathbf{w}(n+1)$ será dada pela equação 3.11.

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \left(\mathbf{J}^T(n)\mathbf{J}(n) \right)^{-1} \mathbf{J}^T(n)\mathbf{e}(n) \quad (3.11)$$

3.2.1.3 Método de Levenberg-Marquardt

O algoritmo proposto por Levenberg-Marquardt é uma modificação do Gauss-Newton que insere o termo $\lambda \mathbf{I}$, conforme descrito na equação 3.12.

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \left(\mathbf{J}^T(n)\mathbf{J}(n) + \lambda \mathbf{I} \right)^{-1} \mathbf{J}^T(n)\mathbf{e}(n) \quad (3.12)$$

A função do fator λ é rastrear a convergência da função erro e , de forma que um aumento do erro implique num aumento do fator e a redução do erro gere redução do mesmo. Um valor muito alto de λ conduz função descrita pela equação 3.12 a se comportar como o algoritmo do Gradiente que possui garantia de convergência. Uma redução acentuada do fator λ leva a equação 3.12 a se aproximar do método de Gauss-Newton, que tem como característica convergência mais rápida. Dessa forma obtém-se um método intermediário entre os algoritmos de Gauss-Newton e do Gradiente.

3.2.2 Algoritmos meta-heurísticos

Uma meta-heurística utiliza a combinação entre escolhas aleatórias e conhecimentos históricos dos resultados para guiarem suas buscas.

3.2.2.1 Algoritmos Genéticos

Algoritmos genéticos são algoritmos meta-heurísticos de otimização que utilizam estratégias inspiradas nos processos da evolução natural e da genética para encontrar o ponto de mínimo ou máximo de uma função (SOARES, 1997).

A aleatoriedade é inserida por meio de processos como a mutação, que introduz genes estranhos em alguns indivíduos da população a cada geração. O direcionamento da evolução é

conduzido pela função de aptidão, responsável por indicar os indivíduos mais aptos, e pelo processo de cruzamento que promove a mistura de genes de indivíduos sobreviventes, com o objetivo de gerar indivíduos mais aptos. O conceito de elite, no qual os melhores indivíduos são preservados de uma geração para outra, garante que, na pior das hipóteses, o melhor indivíduo da geração atual seja igual ao melhor da geração anterior.

Os maiores desafios para utilizar algoritmos genéticos são as escolhas adequadas da função de aptidão e dos valores dos parâmetros. A escolha adequada desses parâmetros vai evitar que o algoritmo convirja prematuramente, ficando preso em um mínimo local.

Do que já foi exposto até esse ponto sobre métodos de otimização, é possível compreender que o algoritmo genético não pode ser utilizado em problemas em tempo real, como os de equalização, por convergirem lentamente e inserirem um elevado custo computacional (SOARES, 1997). No entanto, o objetivo de utilizá-lo nesse trabalho é fazer com que indique a melhor estrutura para tratar um problema específico. Para isso, o algoritmo genético é executado, em modo *offline*, para que indique o valor mais próximo do mínimo global da função de aptidão de algumas estruturas aplicadas a um canal de comunicações. Após a seleção da estrutura que apresentar os melhores resultados, o próximo passo é encontrar um algoritmo determinístico, cujos resultados se aproximem dos encontrados pelo algoritmo genético, para ser executado em tempo real.

3.3 MOTIVAÇÃO PARA O EMPREGO DE EQUALIZADORES NEURAIIS

Antes de dar prosseguimento ao trabalho, nesse ponto, um questionamento se faz necessário: qual é a vantagem do uso de equalizadores neurais? Não seria um equalizador linear simples capaz de tratar o problema satisfatoriamente?

Como já foi mencionada, a utilização de funções de ativação não lineares torna os equalizadores neurais mais eficientes ao tratar as distorções inseridas por canais de comunicações, principalmente quando se trata de canais não lineares.

Atualmente, as estruturas neurais continuam sendo empregadas para a solução de problemas em áreas diversas: acústica (SIGTIA; DIXON, 2014), agronomia (XIAOHUI HUANG et al., 2014), engenharia mecânica (HAO LI et al., 2014), controle robótico (YUNDUAN CUI ; TAKAHASHI; HASHIMOTO, 2014), reconhecimento de linguagem escrita (CHATURVEDI; TITRE; SONDHIIYA, 2014).

Além disso, são inúmeros os trabalhos científicos que delineiam uma comparação entre equalizadores neurais e outros tipos de equalizadores: os resultados obtidos por Siu, Gibson e Cowan (1990) indicam que uma estrutura *DFE* (*Decision Feedback Equalizer*) baseada em perceptrons provê melhor desempenho com relação ao *BER* (*Bit Error Rate*) se comparado ao *LMS* (*Least Mean Square*) *DFE*, principalmente em condições de ruído mais elevado; Peng, Nikias e Proakis (1992) concluem que equalizadores baseados em *MLP* apresentam melhor desempenho que um equalizador linear baseado em *LMS* ao tratar distorções causadas por canal não linear; Kechriotis, Zervas e Manolakos (1994) descrevem a superioridade de *RNN* pequenas com relação a equalizadores baseados em filtros transversais lineares, baseada em resultados de simulação; Albu et al. (1998) confirma os resultados obtidos por outros autores, mostrando que redes *MLP* e *RBF* podem oferecer vantagens sobre estruturas lineares no projeto de equalizadores adaptativos. Esses são apenas alguns exemplos de trabalhos que indicam melhor desempenho na utilização de redes neurais no projeto de equalizadores.

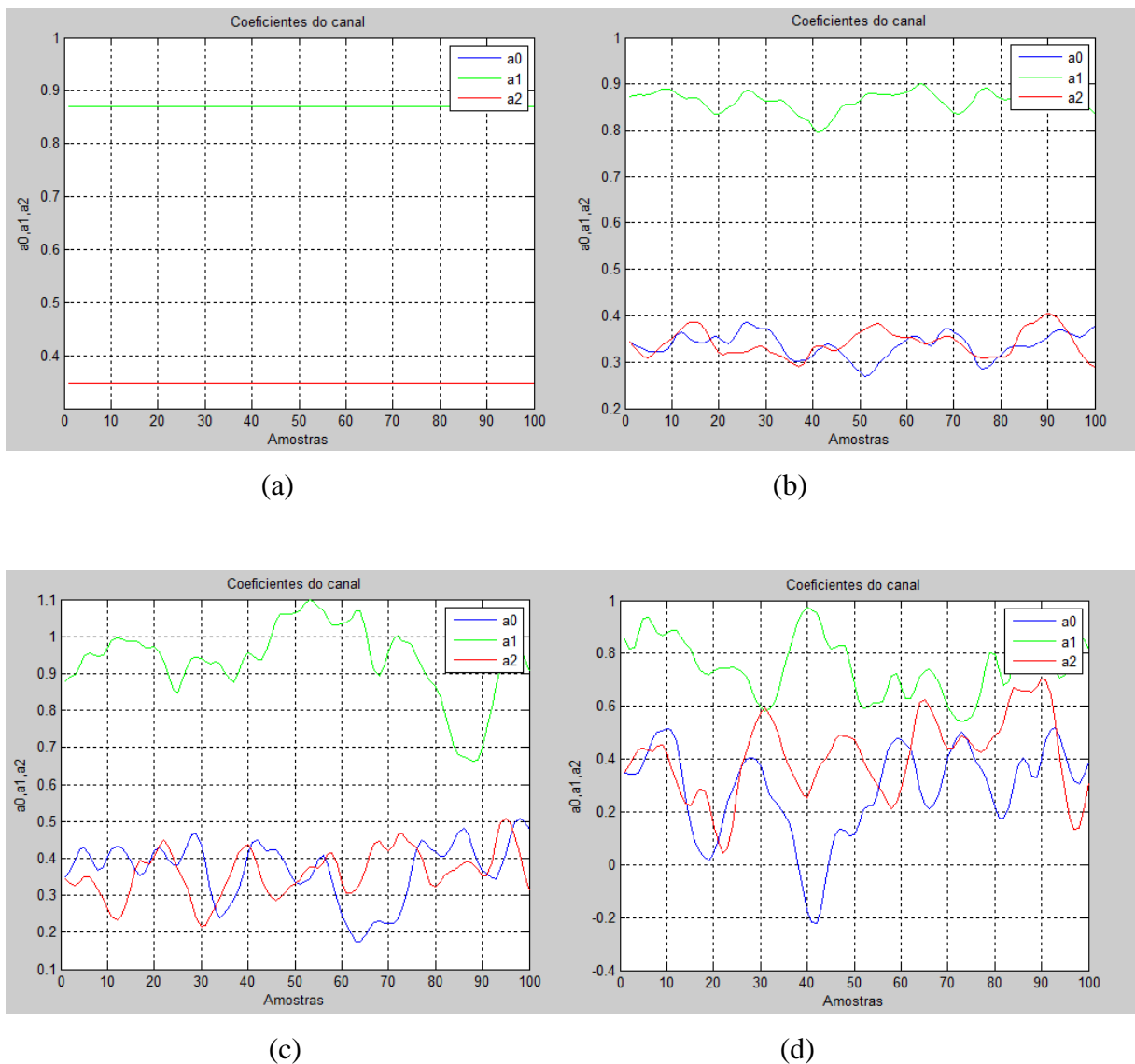
Somado ao melhor desempenho, estruturas que possuem algum tipo de realimentação, como *RNN* ou *DFE* baseados em neurônios, apresentam maior aptidão para tratar sinais provenientes de canais com nulos espectrais profundos (KECHRIOTIS; ZERVAS; MANOLAKOS, 1994). Os sistemas de comunicações móveis apresentam grandes variações, podendo vir a apresentar nulos espectrais profundos e não linearidades. Portanto, baseado nos resultados dos trabalhos realizados nessa área, parece coerente partir do pressuposto que equalizadores neurais que possuam algum tipo de realimentação estejam mais aptos a solucionar o problema proposto nessa dissertação.

São comuns na literatura trabalhos que utilizam equalizadores baseados em redes neurais realimentadas com o objetivo de tratar sinais provenientes de diversos tipos de canais, sujeitos a interferência intersimbólica, com nulos espectrais profundos ou não, lineares ou que apresentam algum tipo de não linearidade e variantes ou invariantes no tempo: Siu, Gibson e Cowan (1990) utilizaram uma estrutura *DFE* baseada em perceptrons para tratar o canal linear e invariante no tempo, recomendado pela *ITU* descrito pela equação 2.10; No trabalho de Sheeja e Sahu (2010), dois canais lineares e invariantes, com um deles apresentando nulos espectrais profundos, são equalizados por duas estruturas *DFE*: uma baseada em *RBF* e outra em *MLP*; Meyer e Pfeiffer (1993) tratam as distorções causadas por um canal variante no tempo utilizando estrutura *DFE* baseada em *MLP*.

Foi proposto por Ling e Proakis (1985) um canal para simulação, cuja função de transferência possui coeficientes variáveis para simular o desvanecimento, com valores

gerados a partir de distribuições Gaussianas, nas quais o desvio padrão é diretamente proporcional à intensidade do desvanecimento. Esses coeficientes foram somados a parcelas fixas em Choi, Lima e Haykin (2005) e Corral, Ludwig e Lima (2010), conforme canal variável representado pela equação 2.10. Os comportamentos dos coeficientes da função de transferência para desvios padrão iguais a 0, 0,1, 0,3 e 0,5 são mostrados na figura 13.

Figura 13- Comportamento dos coeficientes do canal de transmissão para desvio padrão de: a) 0; b) 0,1; c) 0,3; d) 0,5.



Fonte : Elaborada pelo autor.

A figura 13 (a) representa um canal invariante no tempo. Aumentando-se o valor do desvio padrão, figuras 13 (b), (c) e (d), é observado um desvanecimento cada vez mais agressivo. O crescimento da intensidade do desvanecimento pode ser observado se notarmos

que na figura 13 (b), com desvio padrão igual a 0,1, por exemplo, os coeficientes do canal variam 0,1 em torno do seu valor médio. Quando esse desvio padrão aumenta, para 0,5, por exemplo, figura 13 (d), a variação dos coeficientes do canal de transmissão fica em torno de 0,6 do seu valor médio. A variação mais brusca dos coeficientes do canal aumenta a probabilidade de gerar maior variação na intensidade do sinal que atinge o receptor, a depender da composição dos sinais que atravessam cada percurso.

Choi, Lima e Haykin (2005) propuseram um equalizador baseado em RNN com 4 neurônios (31 pesos) e um algoritmo baseado em filtro de Kalman, que insere alto custo computacional, para tratar o canal com desvio padrão de 0,1. Esse arranjo não parece ser o ideal para problemas em tempo real devido à sua complexidade e à consequente dificuldade em rastrear sinais que variam rapidamente. Corral, Ludwig e Lima (2010) propuseram uma estrutura com apenas 1 neurônio e algoritmo de otimização de Gauss-Newton, para tratar o sinal com o mesmo desvio padrão. Apesar dessa proposta parecer promissora, o algoritmo de Gauss-Newton não possui garantia de convergência, podendo comprometer a estabilidade do sistema.

É um objetivo desse trabalho, encontrar uma estrutura neural simples que, utilizando algoritmo de otimização confiável, seja capaz de equalizar com menor erro possível, o canal descrito acima, com desvios padrão 0,1, 0,3 e 0,5. Ou seja, além de tratar distorções em um canal sujeito a desvanecimento lento, busca-se a solução, não encontrada na literatura, para canais com desvanecimento severo.

4 ESCOLHA DA ESTRUTURA DO EQUALIZADOR NEURAL

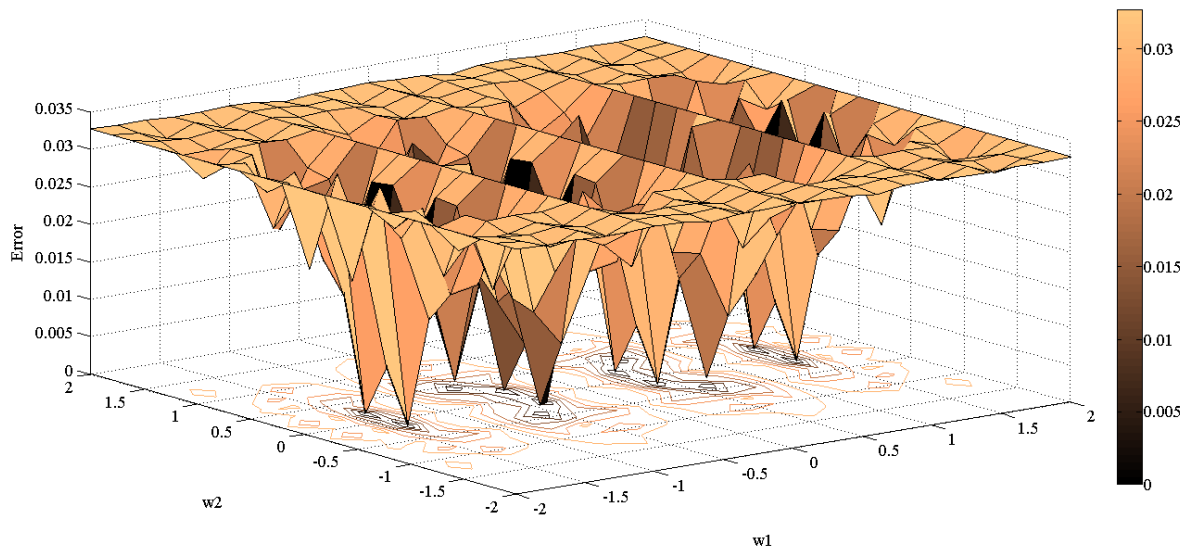
Diferentes tipos de redes neurais têm sido empregados para fins de equalização, contudo, não foi encontrado na literatura científica nenhum método que comparasse essas estruturas neurais e indicasse a que se ajusta melhor a determinado tipo de problema. Frequentemente esses trabalhos, por meio da proposta de mudanças na estrutura neural, no algoritmo de otimização ou em ambos, conseguem melhoras incrementais na curva da BER (*Bit Error Rate*) versus SNR (*Signal to Noise Ratio*), porém sem enfatizar seus critérios de aplicabilidade. Tal abordagem não contribui para obtenção da melhor estrutura neural para resolver um problema equalização específico. Existem muitos algoritmos e estruturas diferentes, porém, a busca de um algoritmo que melhor se ajuste a determinada estrutura ou uma comparação entre estruturas são assuntos ainda pouco abordados.

A necessidade de se encontrar os melhores pesos para determinada estrutura neural, conduz a problemas multimodais, tal qual o ilustrado na figura 14. Na figura é possível observar a variação da função erro em relação à variação de dois pesos sinápticos w_1 e w_2 . Ainda que os valores dos pesos tenham sido limitados dentro de uma faixa estreita de valores, entre -2 e 2, é difícil indicar visualmente o valor mínimo global da função erro. Considerando então que uma estrutura neural muito pequena possui pelo menos 5 pesos e variando os valores dos pesos dentro de uma faixa maior, entre -1000 e 1000, por exemplo, é possível imaginar que existiriam centenas de combinações de valores de pesos que levariam a valores de erro muito próximos. Esse tipo de problema é complexo e a probabilidade de um algoritmo de otimização determinístico ficar preso em um mínimo local é muito alta. Para solução de problemas dessa complexidade, um algoritmo de otimização mais robusto, tal qual o algoritmo genético, seria mais indicado.

Nesse capítulo é proposto um método para avaliar o potencial de determinados equalizadores neurais aplicados a canais de comunicação específicos. Na presente abordagem, o algoritmo genético é utilizado para tentar determinar os melhores pesos sinápticos para cada estrutura neural, o que permite avaliar os resultados obtidos pelo algoritmo determinístico originalmente proposto para o equalizador. Para tal fim, são formadas populações nas quais cada indivíduo corresponde a um conjunto de pesos sinápticos e a função de aptidão corresponde ao erro obtido ao se comparar a saída calculada pela estrutura neural, utilizando os indivíduos, e o valor esperado, conforme conjunto de treinamento. Assim, infere-se sobre

possíveis limitações nos resultados das curvas BER x SNR deste equalizador, que podem ser decorrentes da estrutura neural ou de um algoritmo de otimização inadequado.

Figura 14- Exemplo de comportamento da variação do erro numa estrutura neural em função da variação dos pesos sinápticos.



Fonte : Elaborada pelo autor.

Para testar o método proposto, foram utilizados três equalizadores neurais, idênticos àqueles empregados em Patra (1999), para equalização de sinais 4-QAM (*Quadrature Amplitude Modulation*). A escolha de Patra (1999) para teste da metodologia, apesar de se tratar de uma referência relativamente antiga, se deveu ao fato de ser a única a tratar sinal 4-QAM utilizando mais de um tipo de estrutura neural, nesse caso três. Dessa forma os resultados obtidos com o teste puderam demonstrar o potencial da metodologia desenvolvida. Na seção 4.2 é apresentada uma comparação entre essas estruturas e uma avaliação da eficácia dos algoritmos de otimização aplicados a cada uma delas. Em 4.1 são descritas as principais dificuldades para definição de uma estrutura neural adequada a um problema específico. Na seção 4.3, com o objetivo de testar o método, avaliam-se algumas estruturas já utilizadas na literatura para resolver o problema de equalização em um canal com severa ISI. O canal utilizado é o recomendado pela UIT (União Internacional de Telecomunicações) para avaliação de equalizadores, modificado na literatura para contemplar o efeito do desvanecimento.

4.1 ESCOLHA DO EQUALIZADOR NEURAL

Na maioria dos trabalhos analisados, a topologia para a rede neural é escolhida de forma *ad hoc*. Os trabalhos encontrados na literatura levam em consideração características gerais das topologias para justificar sua aplicabilidade, porém sem uma avaliação sistemática e comparação entre elas.

O número possível de estruturas neurais para solução de um problema de equalização não é mensurável. Muitas dessas topologias já foram empregadas na literatura científica, tais como: MLP, PPN, FLANN, RBF, RNN. Cada uma delas pode variar dependendo do número de perceptrons empregados, quantidade de camadas, quantidade e tipos de ligações entre perceptrons, utilização ou não de *bias*², tipos de entradas, funções de ativação.

Para atualização dos pesos das redes neurais, geralmente são utilizados algoritmos de otimização determinísticos, os quais preferencialmente apresentam reduzido custo computacional, que os permite serem utilizados para aplicações em tempo real. Foram descritos na subseção 3.1.1 alguns dos algoritmos determinísticos mais utilizados na literatura científica em equalizadores baseados em redes neurais.

A combinação entre uma arquitetura neural e um algoritmo de otimização possibilita uma grande quantidade de soluções neurais para o problema de equalização, resultando em desempenhos de BER x SNR diferentes. Nesse contexto, a escolha do melhor algoritmo de otimização está estritamente vinculada à topologia escolhida, ou seja, a melhor estrutura indicada pelo método proposto pode apresentar pior desempenho que outra, a depender do algoritmo aplicado.

Não foi encontrada na literatura uma sistematização que permita avaliar a melhor combinação entre estrutura e algoritmo aplicável a um problema de equalização específico. Neste trabalho é proposto um método que seja capaz de comparar o potencial de estruturas neurais ou, no caso da estrutura neural já definida, indicar o algoritmo de otimização que mais se aproxime do mínimo global.

² Entrada extra à camada de entrada da rede neural, sempre em valor unitário, sofrendo ponderações do peso correspondente.

4.2 MÉTODO PROPOSTO

Definida a estrutura de um equalizador neural, a determinação dos pesos da rede usualmente é resolvida utilizando algoritmos de otimização, o qual busca minimizar a BER. Portanto, para um problema específico de equalização, avaliando-se a mínima BER obtida por estruturas distintas, mantidas as condições de relação sinal ruído no canal, é possível indicar a arquitetura que possui maior potencial de mitigar os problemas associados ao canal de comunicação.

Nesse trabalho, propõe-se como metodologia de avaliação dos equalizadores, a aplicação de Algoritmo Genético (GA), em modo *offline*, para buscar o mínimo global da curva BER. A utilização conjunta de Algoritmo Genético e Redes Neurais é antiga. Segundo Schaffer, Whitley e Eshelman (1992) existiam três abordagens principais nas quais tal combinação era utilizada: o GA dando suporte à rede neural, filtrando os dados de treinamento, selecionando os parâmetros e regras de aprendizagem, avaliando o desempenho; o GA atualizando os pesos neurais; e o GA especificando a topologia da rede neural. Ning e Cai-yun (2010) combinaram o algoritmo genético ao *backpropagation* com a finalidade de atualizar os pesos neurais em uma rede utilizada para conversão de coordenadas geográficas. Peixian et al. (2011) aplicaram algoritmo genético para indicar os pesos iniciais em uma rede MLP utilizada na previsão de séries temporais.

Tal metodologia indica, dentre algumas estruturas, a que potencialmente geraria o melhor desempenho para o problema proposto. Porém, o uso de algoritmo genético em aplicações em tempo real é inviável, devido ao seu alto custo computacional, sendo necessária a escolha de um algoritmo determinístico apropriado para a estrutura escolhida.

Definida a estrutura neural de maior potencial, uma forma de avaliar a adequação de um algoritmo determinístico a esta estrutura é comparando a curva BER, alcançada por este algoritmo, àquela conseguida empregando-se o GA. Quanto mais próximas forem essas curvas, mais eficaz será a aplicação em tempo real, optando-se, portanto, pelo algoritmo determinístico que mais se aproximar da curva BER x SNR do algoritmo genético.

Estruturas muito pequenas podem gerar erros elevados de ajuste ao sistema real, provocando um desempenho ruim. Por outro lado, estruturas maiores, das quais se poderiam esperar taxas de erro menores, podem sofrer com *overfitting* (LAWRENCE; GILES; TSOI, 1997), no qual o tamanho da rede neural excede o mínimo necessário para resolução

satisfatória do problema, o que pode prejudicar o seu poder de generalização. É esperado que o método proposto indique possíveis estruturas em *overfitting*.

Estruturas grandes impõem maior dificuldade ao algoritmo de otimização na sua função de encontrar os melhores pesos sinápticos levando, muitas vezes, a resultados piores que o de estruturas neurais menores. Foi concluído por Lawrence, Giles e Tsoi (1997) que, mesmo para uma estrutura de tamanho excessivo, um tempo infinito e um algoritmo apropriado para treinamento resultariam na melhor solução possível. No entanto, para aplicações em tempo real, como as avaliadas nesse trabalho, a utilização da rede neural mínima é essencial para que o sistema de comunicação não gere atrasos que venham a inviabilizar seu funcionamento.

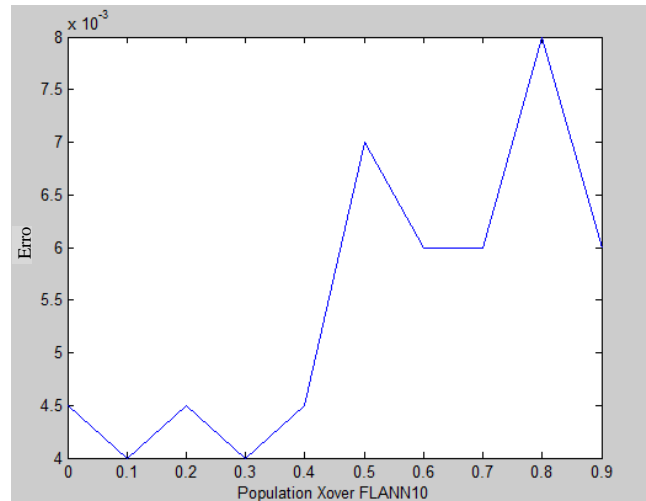
Para testar o método proposto, foram avaliadas três estruturas neurais empregadas em Patra (1999), cujo trabalho apresenta uma análise da aplicação de diversas topologias a alguns problemas específicos e tece conclusões que serão discutidas.

A abordagem desenvolvida em Patra (1999) é uma das poucas na literatura que se propõe a discutir e determinar uma estrutura, dentre algumas candidatas, que mais se adeque a um problema de equalização específico, utilizando como parâmetro as curvas BER x SNR obtidas por cada estrutura. Portanto, optou-se por utilizar os mesmos parâmetros de Patra (1999), para efeito de comparação entre os resultados obtidos por seus autores com o método aqui proposto. Além disso, Patra (1999) emprega um número considerável de estruturas neurais com o objetivo de equalizar um sinal 4-QAM, que sofre distorções causadas por multipercurso.

Seguindo o método proposto, empregou-se o GA na busca das melhores curvas BER x SNR. A escolha dos parâmetros do GA é uma questão crucial para o sucesso da metodologia, buscando-se reduzir a probabilidade do algoritmo parar num mínimo local, evitando-se a redução da diversidade genética. Desta forma, os parâmetros do GA foram escolhidos buscando-se reduzir a probabilidade de convergência prematura. O tamanho da população e a taxa de cruzamento foram escolhidos dentro dos intervalos [50;120], [0;0,9], respectivamente, e o intervalo inicial da população variou entre [-1,1] e [-100;100], sendo os intervalos definidos baseados em testes de simulação. Portanto, para os três parâmetros foram utilizados os valores que resultaram na menor BER. A figura 15 ilustra um exemplo de curva para ajuste de parâmetros, na qual os valores de BER são definidos em função da variação da taxa de cruzamento, mantendo o SNR fixo em 10dB. Nesse exemplo, o menor valor do BER foi obtido para taxas de cruzamento de 10% e 30%, indicando um valor ideal abaixo de 40% para

esse parâmetro. Os outros parâmetros, escolhidos por meio de teste de simulação, são apresentados na Tabela 1.

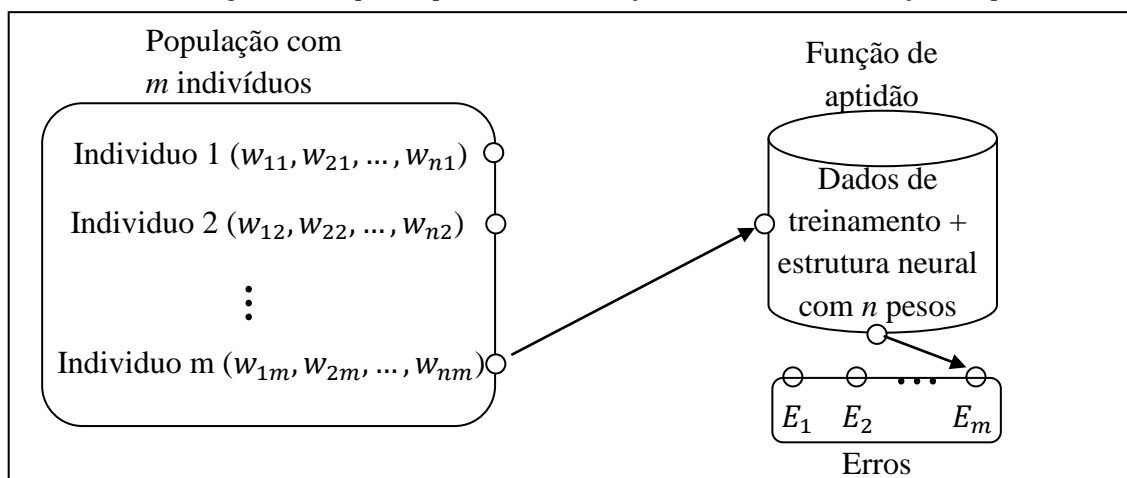
Figura 15- Curva do BER vesus taxa de cruzamento para SNR fixo em 10dB.



Fonte : Elaborada pelo autor.

Além disso, a função de aptidão consiste na comparação entre o resultado esperado e o obtido pela estrutura neural, utilizando o conjunto de símbolos de teste. Apesar da determinação da função de aptidão ser simples para o problema de equalização, ela pode ser de grande complexidade, conforme apresentado por Juan, Carlos e Jorge (2006), sendo crucial para a solução de qualquer problema que utilize GA.

Figura 16- Esquema que ilustra a utilização dos indivíduos na função de aptidão.



Fonte: Elaborada pelo autor.

Apesar de simples, a definição da população e da função de aptidão é de grande importância para o entendimento da metodologia proposta:

- Cada indivíduo que forma a população corresponde a um conjunto de pesos sinápticos que será testado na função de aptidão;
- A função de aptidão corresponde ao conjunto que contém dados de treinamento (entradas e saída esperada) e função correspondente à estrutura neural a ser testada. Cada indivíduo é aplicado à estrutura que, quando submetida aos dados de treinamento, geram um valor de erro E , que indicará o potencial do indivíduo. Ou seja, quanto maior o erro gerado, menor será a aptidão. O esquema descrito pode ser visualizado na figura 16.

TABELA 1
PARÂMETROS UTILIZADOS PELO ALGORITMO GENÉTICO.

Parâmetro	Valor
Tipo de cruzamento	<i>scattered</i>
Tamanho da Elite	2
Escala de desempenho	ranque
Número máximo de gerações	100
Algoritmo determinístico híbrido	baseado na Hessiana
Tipo de mutação	Gaussiana
Função de seleção	Roleta

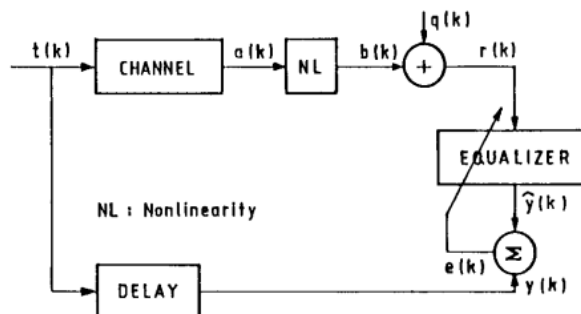
Os valores apresentados na Tabela 1 representam:

- O tipo de cruzamento é a regra segundo a qual os genes dos progenitores são combinados. O valor *scattered* implica na geração aleatória de uma sequência binária, do mesmo tamanho do cromossomo do descendente, segundo a qual as posições ocupadas por bits “1” são preenchidas com os genes de um progenitor e as ocupadas com bit “0” com a do outro;
- A elite, da qual os indivíduos são preservados para a próxima geração, contém dois indivíduos;
- A escala de desempenho consiste na regra estabelecida para atribuição de nota baseada na aptidão de cada indivíduo. O valor ranque atribui uma numeração sequencial, sendo atribuído 1 para o mais apto, 2 para o segundo melhor e assim por diante;
- O número máximo de 100 gerações foi definido como limite para a convergência do GA;

- Um algoritmo determinístico, baseado no cálculo da Hessiana, foi utilizado, paralelamente, para facilitar a convergência local, o que torna o algoritmo híbrido utilizado mais rápido;
- A mutação gaussiana consiste na geração de uma sequência numérica aleatória do tamanho de um cromossomo, baseada em distribuição gaussiana, que determina os genes que sofrem mutação;
- A função de seleção indica a regra segundo a qual os pares de progenitores serão formados para gerar a próxima geração. O valor roleta corresponde à formação de um gráfico no formato de pizza, desenhada sobre uma roleta, na qual cada fatia é atribuída a um indivíduo e o tamanho da fatia é proporcional à aptidão do mesmo. A seleção dos progenitores acontece como em um jogo de roleta e, portanto, os indivíduos que possuírem as maiores fatias terão maior probabilidade de propagar seus genes.

As estruturas foram avaliadas utilizando-se o canal linear apresentado na equação 4.1, submetido a 4 tipos de não linearidades, mostradas nas equações (4.2) a (4.5), conforme proposto em Patra (1999) e mostrado na figura 17, usando 10.000 símbolos para teste em cada relação Sinal Ruído (SNR). A comparação entre os resultados obtidos nas simulações e aqueles obtidos em Patra (1999) propiciaram algumas conclusões importantes sobre as estruturas neurais e o algoritmo de otimização utilizados. Foi possível mensurar o nível de adequação entre a estrutura e o algoritmo empregado.

Figura 17- Sistema de transmissão digital



Fonte : (PATRA, 1999).

$$CH = 2: 0,447 + 0,8894z^{-1}$$

(4.1)

O canal utilizado corresponde a um canal de fase não mínima (CHEN; GIBSON; COWAN, 1990) e foi escolhido para testar o método.

$$NL = 0: b(k) = a(k) \quad (4.2)$$

$$NL = 1: b(k) = \tanh(a(k)) \quad (4.3)$$

$$NL = 2: b(k) = a(k) + 0,2a^2(k) - 0,1a^3(k) \quad (4.4)$$

$$NL = 3: b(k) = a(k) + 0,2a^2(k) - 0,1a^3(k) + 0,5\cos(\pi a(k)) \quad (4.5)$$

Na abordagem utilizada em Patra (1999), um canal linear é aplicado em série com a não linearidade, de forma a representar o comportamento não linear, conforme proposto no modelo de Hammerstein (DING; CHEN, 2005).

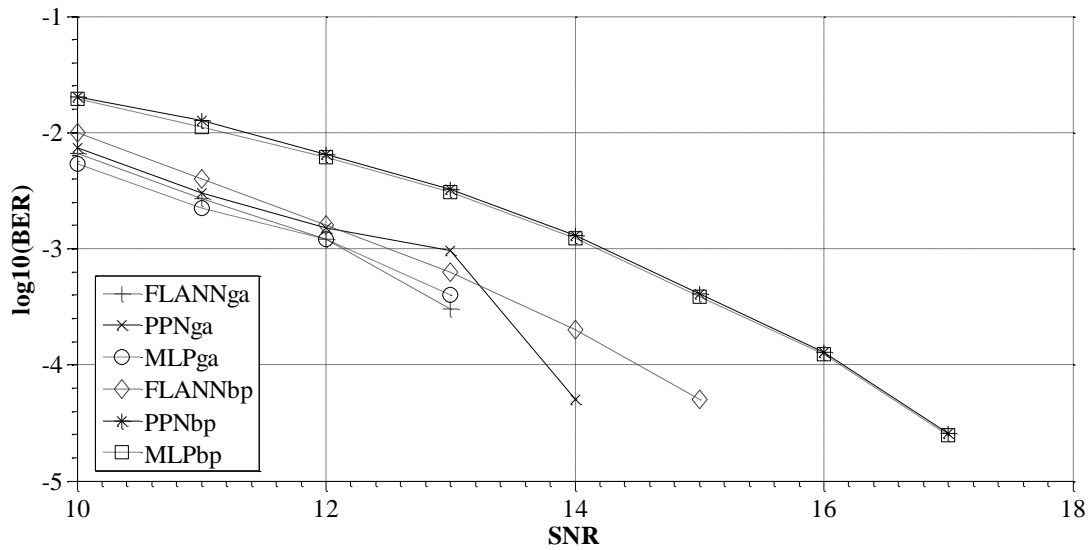
Em Patra (1999) são abordados: um equalizador Multicamada MLP (*Multilayer Perceptron*), composto por 6 entradas, 8 neurônios na camada escondida e 2 na camada de saída; um equalizador baseado em Rede Polinomial de Perceptrons PPN (*Polynomial Perceptron Network*), com 4 entradas que se transformam em 19 após a fase de processamento polinomial e 2 neurônios na saída; e uma Rede Neural Artificial com Link Funcional FLANN (*Functional Link Artificial Neural Network*) composto por 4 entradas que se transformam em 19 após a fase de expansão funcional e 2 neurônios na saída. Os pesos sinápticos são atualizados pelo algoritmo *backpropagation*.

4.2.1 Avaliação de desempenho

Os resultados das simulações geraram curvas da BER versus SNR para as três estruturas propostas em Patra (1999), FLANN, MLP e PPN, usando os algoritmos BP e GA para atualização dos pesos sinápticos. Todos eles foram avaliados utilizando-se o modelo de canal apresentado na equação 4.1, aplicando-se as quatro não-linearidades descritas anteriormente, usando 10.000 símbolos para teste.

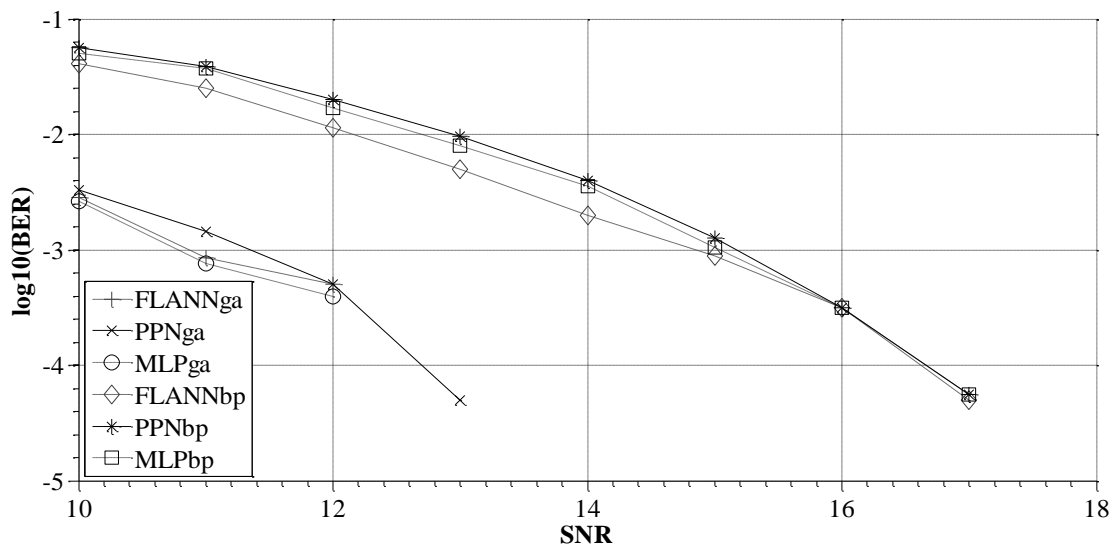
Na figura 18, aplicou-se NL=0 (canal linear). O algoritmo BP mostrou estar bem ajustado para a estrutura FLANN, já que a curva ficou próxima à obtida quando o algoritmo genético foi aplicado. Por outro lado, o algoritmo BP utilizado em Patra (1999) não está bem ajustado às estruturas MLP e PPN, já que existe um deslocamento considerável entre as curvas quando utilizam o algoritmo genético e o *backpropagation*.

Figura 18- Comparação de desempenho entre estruturas neurais diferentes com NL=0.



Fonte : Elaborado pelo autor.

Figura 19- Comparação de desempenho entre estruturas neurais diferentes com NL=1.



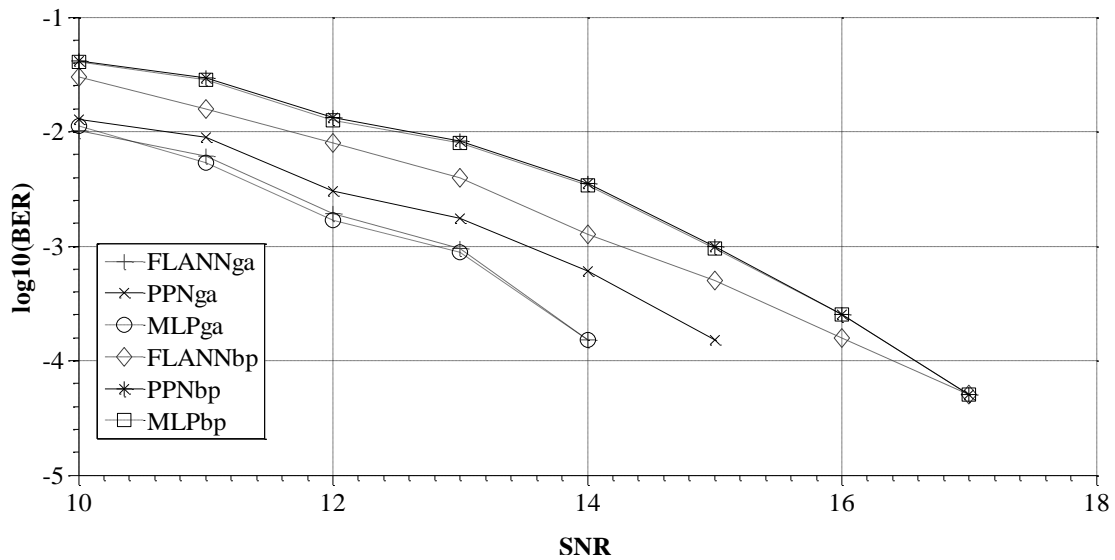
Fonte : Elaborado pelo autor.

Na figura 19, aplicou-se NL=1. O algoritmo BP aplicado não está bem ajustado a nenhuma das três estruturas, pois existe um grande deslocamento entre as curvas das estruturas utilizando o algoritmo genético e o BP. A comparação entre a figura 18 com a figura 19 mostra resultados melhores para o algoritmo genético no canal com NL=1, se comparado ao canal linear (NL=0), o que pode parecer equivocado. A análise detalhada mostrou que, como pode ser visto nos gráficos de espalhamento das figuras 22 (a) e (b), a não-linearidade NL=1 facilita o processo de equalização, pois separa áreas de indecisão. O mesmo comportamento

ocorre com a não-linearidade $NL=3$, como pode ser observado na figura 22(d). A única não-linearidade que realmente dificulta o trabalho do equalizador é a $NL=2$, conforme figura 21(c).

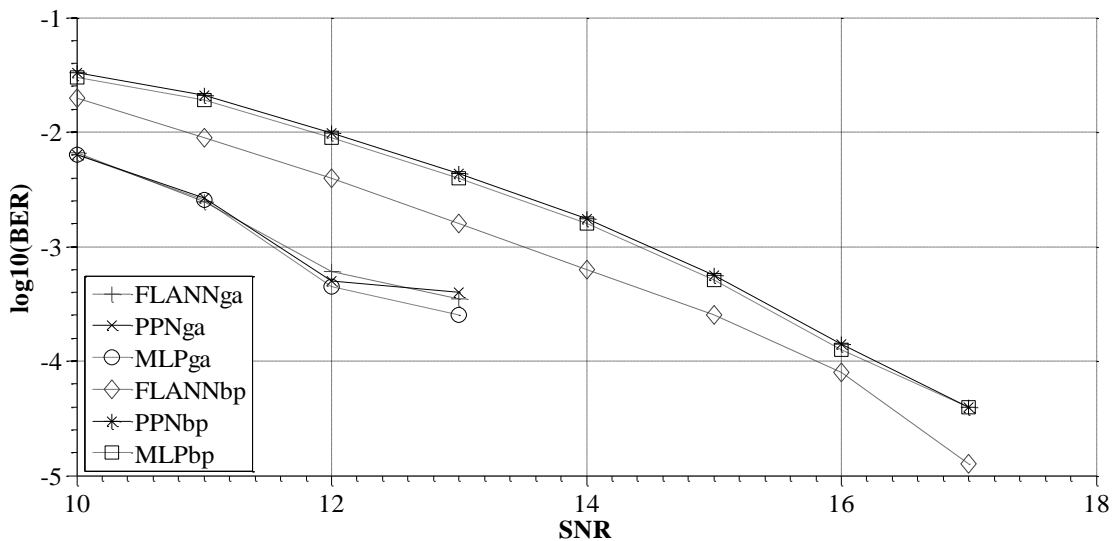
Na figura 20, aplicou-se $NL=2$. As três estruturas apresentam deslocamento quando utilizado os algoritmos BP e o GA para atualização dos pesos sinápticos. Entre as estruturas analisadas o maior deslocamento ocorreu para o MLP.

Figura 20- Comparação de desempenho entre estruturas neurais diferentes com $NL=2$.



Fonte : Elaborado pelo autor.

Figura 21- Comparação de desempenho entre estruturas neurais diferentes com $NL=3$.

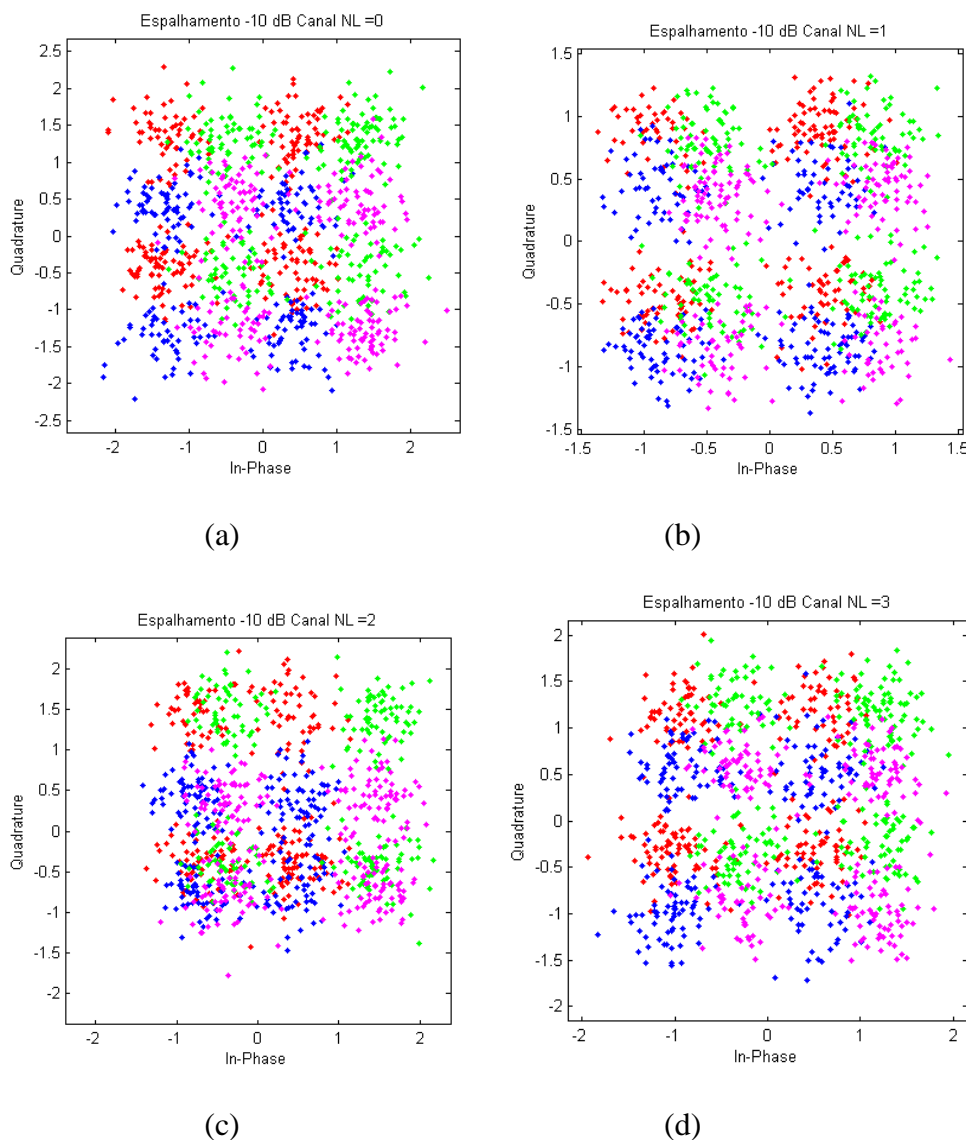


Fonte : Elaborado pelo autor.

Por fim, são mostrados na figura 21 os resultados quando a $NL=3$. Novamente é possível concluir que o algoritmo BP está mais bem ajustado à estrutura FLANN. As três estruturas apresentam deslocamento para ambos os algoritmos, porém com menor deslocamento para a estrutura FLANN.

Na figura 22 são mostradas as representações em formato *scatter plot* do sinal gerado, conforme representação mostrada na figura 1 (b), após atravessar o canal de transmissão representado pela equação 4.1, sujeitos às não linearidades $NL=0$, $NL=1$, $NL=2$ ou $NL=3$.

Figura 22- Representação scatter plot: a) $NL=0$; b) $NL=1$; c) $NL=2$; and d) $NL=3$.



Fonte : Elaborado pelo autor.

Os resultados obtidos em Patra (1999) levaram seus autores à conclusão de que o desempenho da estrutura FLANN proposta seria melhor, em termos da BER, que o das

estruturas PPN e MLP e que essas duas teriam um desempenho próximo entre si. No entanto, avaliando as três estruturas por meio da metodologia aqui proposta, foi possível identificar uma proximidade entre os desempenhos das redes FLANN e MLP, por sua vez um pouco superiores à PPN. A conclusão diferente pode ter resultado do melhor ajuste do algoritmo *backpropagation*, utilizado em Patra (1999), à rede FLANN, ou seja, seria necessário encontrar um algoritmo de otimização que se ajustasse melhor às estruturas MLP e PPN, explorando melhor suas potencialidades.

4.3 ESCOLHA DA ESTRUTURA NEURAL

Nesta seção será utilizada a metodologia apresentada na seção anterior para comparar o desempenho de seis estruturas neurais quando aplicados ao canal representado na equação 4.6, descrita na seção 2.2, problema objetivo dessa dissertação, que corresponde a um canal recomendado pela UIT para estudos de equalização, com $c_0 = c_2 = 0,3482$, $c_1 = 0,8704$, acrescido de termos variáveis a_i , com $i=0, 1$ e 2 , que inserem o efeito do desvanecimento.

$$H = [c_0 + a_0(t)] + [c_1 + a_1(t)]z^{-1} + [c_2 + a_2(t)]z^{-2} \quad (4.6)$$

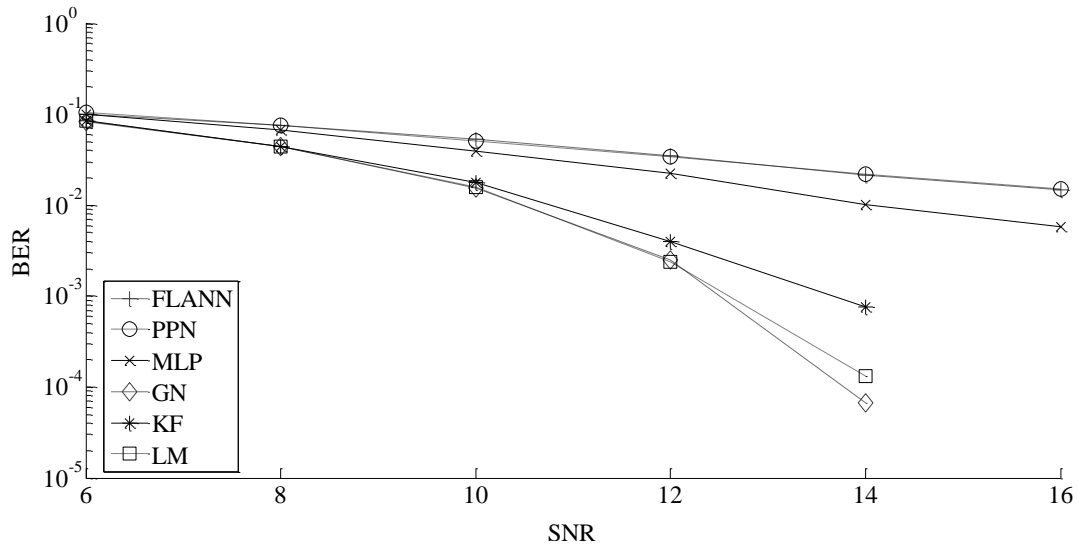
Das seis estruturas analisadas, três foram propostas em Patra (1999), conforme descrito na seção anterior, sendo uma rede MLP, uma PPN e uma FLANN. Outra estrutura consiste em uma rede neural recursiva (RNN) com três camadas, contendo uma camada escondida com três neurônios e o total de 31 pesos sinápticos, conforme proposto por Choi, Lima e Haykin (2005). A quinta estrutura consiste em apenas um neurônio recursivo, tal qual proposto por Corral, Ludwig e Lima (2010), com apenas cinco pesos sinápticos complexos. A última estrutura consiste de dois neurônios recursivos independentes, cada um tratando a parte real ou imaginária do sinal recebido, com dez pesos neurais reais.

A cada uma das estruturas foi aplicado o algoritmo genético para encontrar os melhores pesos sinápticos, conforme metodologia explicada na seção anterior. Os resultados obtidos indicarão a estrutura a ser utilizada na solução do problema proposto nessa dissertação.

Na figura 23 é exibida a comparação do desempenho das seis estruturas descritas acima. Todas foram avaliadas utilizando-se o modelo de canal (4.6), com desvio padrão dos coeficientes variáveis (σ) iguais 0,1. Foram usados 3.000 símbolos para testar as estruturas avaliadas. Cada curva é resultado da média de 10 simulações estatisticamente independentes.

Os parâmetros escolhidos para simulação são semelhantes aos utilizados por Choi, Lima e Haykin (2005) e Corral, Ludwig e Lima (2010), de forma que fosse possível obter resultados comparáveis aos da literatura correlata.

Figura 23- Comparação de desempenho entre estruturas neurais diferentes.



Fonte : Elaborado pelo autor.

As estruturas que utilizam redes recursivas (RNN), sendo elas a GN (Gauss-Newton), a KF (Kalman *filter*) e a LM (Levenberg-Marquardt), apresentam nítida vantagem sobre as que não utilizam nenhum tipo de realimentação, a FLANN (*Functional Link Artificial Neural network*), a PPN (*Polynomial Perceptron Network*) e a MLP (*Multilayer Perceptrons*). Essa vantagem pode ser atribuída à facilidade que redes com realimentação têm em tratar canais com nulos espectrais profundos (KECHRITIS; ZERVAS; MANOLAKOS, 1994), como o representado pela equação 4.6, conforme descrito no capítulo 2.

Dentre as três estruturas classificadas como RNN é possível observar o melhor desempenho das GN e LM com relação à estrutura KF, notadamente nos pontos com maior SNR. Essa diferença pode ser atribuída à maior dificuldade que o algoritmo genético tem em encontrar o mínimo para uma estrutura mais complexa, como a KF, que possui 31 pesos sinápticos. O excesso de graus de liberdade que uma estrutura neural maior apresenta, combinado à representatividade limitada dos símbolos de treinamento, pode reduzir o seu poder de generalização. Uma estrutura nessas condições é dita estar em *overfitting*, conforme descrito por Lawrence, Giles e Tsoi (1997) e Mahdavian et al. (2008). Uma estrutura em *overfitting*, como a KF, necessitaria de um tempo de treinamento muito elevado, o que seria

incompatível com aplicações em tempo real, como é o problema da equalização de canais de comunicação.

O mesmo raciocínio pode ser empregado para justificar a pequena diferença entre a GN e a LM, em favor da primeira, já que ela possui apenas 5 pesos enquanto a LM possui 10. Essa dificuldade também prevaleceria caso se usasse um algoritmo determinístico para encontrar os pesos sinápticos.

Apesar da pequena vantagem da estrutura GN, a estrutura empregada neste trabalho será a LM, pois confere maior estabilidade e menor custo computacional ao equalizador. Maior estabilidade porque ela evita a aproximação por definição $\tanh(x + j \cdot y) \triangleq \tanh(x) + j \cdot \tanh(y)$ (LEUNG; HAYKIN, 1991; BENVENUTO; PIAZZA, 1992) que pode gerar erro na atualização dos pesos sinápticos e menor custo computacional porque evita operações entre números complexos, já que todos os seus pesos possuem valores reais.

5 EQUALIZAÇÃO DE CANAL VARIANTE UTILIZANDO ESTRUTURA NEURAL RECURSIVA

Neste capítulo será apresentada a solução de um dos problemas objeto dessa dissertação: a equalização do sinal 4-QAM proveniente do canal de comunicação sujeito a desvanecimento severo e lento, interferência intersimbólica (devido à existência de multipercurso) e ruído térmico, utilizando uma estrutura neural simples e um algoritmo de otimização estável, que possibilite a convergência da solução para um mínimo global ou para um mínimo local que atenda alguns requisitos e apresente baixo custo computacional.

Corral, Ludwig e Lima (2010) utilizaram um neurônio recorrente para equalizar o sinal proveniente de um canal sujeito a multipercurso, desvanecimento e ruído térmico. O algoritmo de otimização de Gauss-Newton (GN) foi empregado para atualização dos pesos neurais, dando origem a uma estrutura chamada de DFE-GN (Decision Feedback Equalizer – Gauss-Newton). A instabilidade desse algoritmo, que não possui em suas características a garantia de convergência, inviabilizou a sua utilização na solução do problema de equalização de canal sujeito a desvanecimento severo.

Para efeito de comparação, os resultados obtidos nesse trabalho foram avaliados com base nos resultados obtidos por Choi, Lima e Haykin (2005), que implementa uma rede neural recursiva de três camadas, contendo uma camada escondida com três neurônios e o total de 31 pesos sinápticos que são atualizados por algoritmos baseado em filtro de kalman. Desse trabalho, duas estruturas conhecidas como DFE-EKF e DFE-UKF são utilizadas. Além delas, a estrutura aplicada por Corral, Ludwig e Lima (2010), descrita anteriormente, também foi comparada.

Dentre os tipos de canais de comunicações apresentados no capítulo 2, foi utilizado o canal variante no tempo, modelado pela função de transferência apresentada na equação 5.1, para avaliar o desempenho da estrutura escolhida.

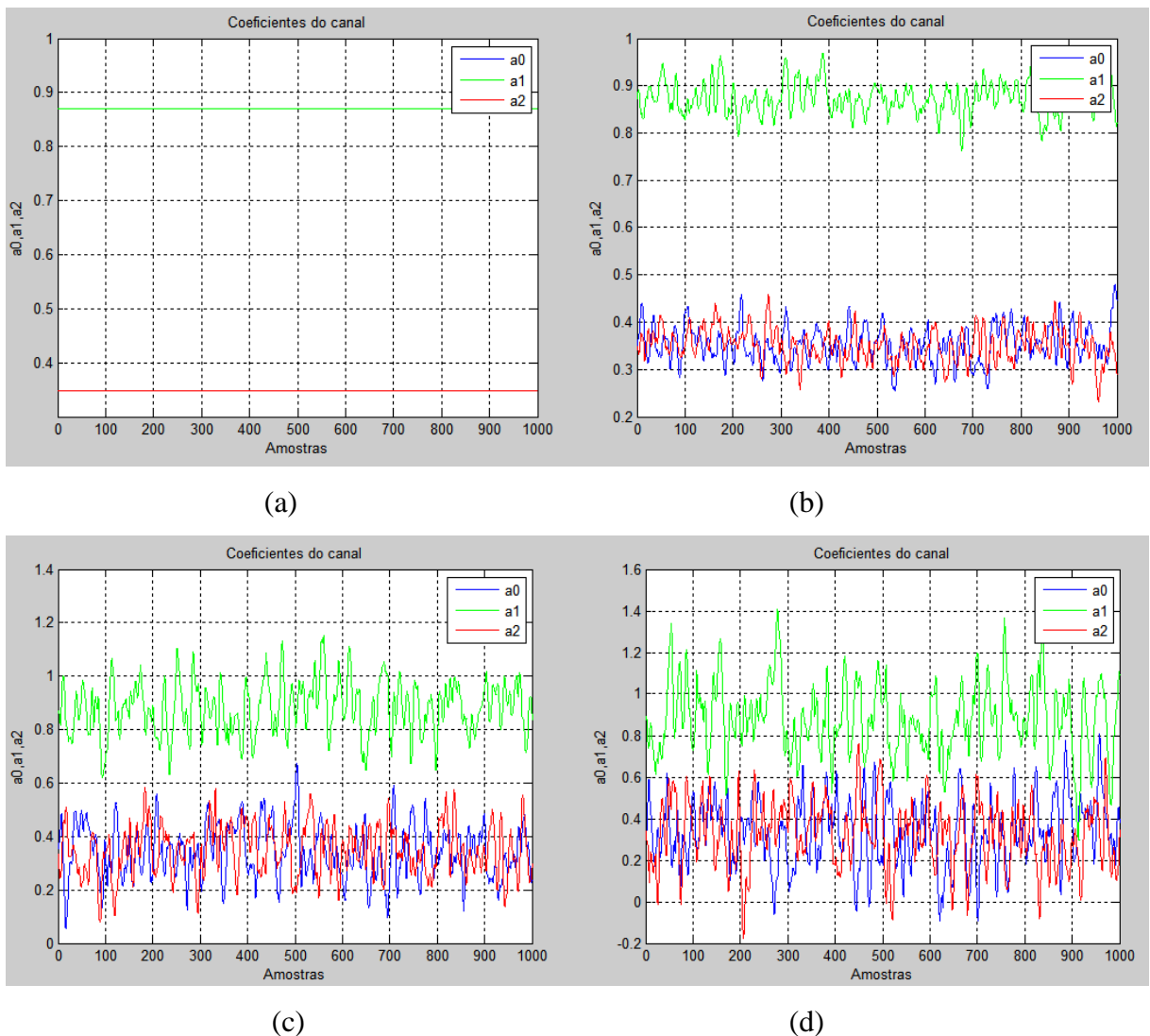
$$H = [c_0 + a_0(t)] + [c_1 + a_1(t)]z^{-1} + [c_2 + a_2(t)]z^{-2} \quad (5.1)$$

A função de transferência H , com $a_0(t) = a_1(t) = a_2(t) = 0$ e $c_0 = c_2 = 0,3482$, $c_1 = 0,8704$, é um canal invariante recomendado pelo ITU para testar o desempenho de equalizadores, que é usualmente utilizado na literatura científica (PENG; NIKIAS; PROAKIS, 1992) e (SHAFI; ZERGUINE; BETTAYEB, 1999). Trata-se de um canal linear de fase não mínima que causa severas distorções no sinal transmitido. Os coeficientes

variantes no tempo foram propostos por Ling e Proakis (1985) para transformar o canal invariante, recomendado pelo UIT, num canal variante no tempo.

Os coeficientes variáveis são gerados de forma independente, conforme proposto por Ling e Proakis (1985), passando um sinal de ruído branco através do filtro passa baixa de Butterworth de segunda ordem com frequência de corte normalizada de 0,1. Os coeficientes da função de transferência têm distribuição Gaussiana com média c_i e variância proporcional ao desvio padrão (σ) do ruído branco. Os valores dos coeficientes são proporcionais a σ , i.e., um valor alto de σ representa um canal com desvanecimento severo. A figura 24 mostra o comportamento dos coeficientes da função H para quatro valores de σ , escolhidos de forma a possibilitar a comparação com os resultados obtidos na literatura.

Figura 24- Comportamento dos coeficientes do canal de transmissão para desvio padrão de: a) 0; b) 0,1; c) 0,3; d) 0,5.

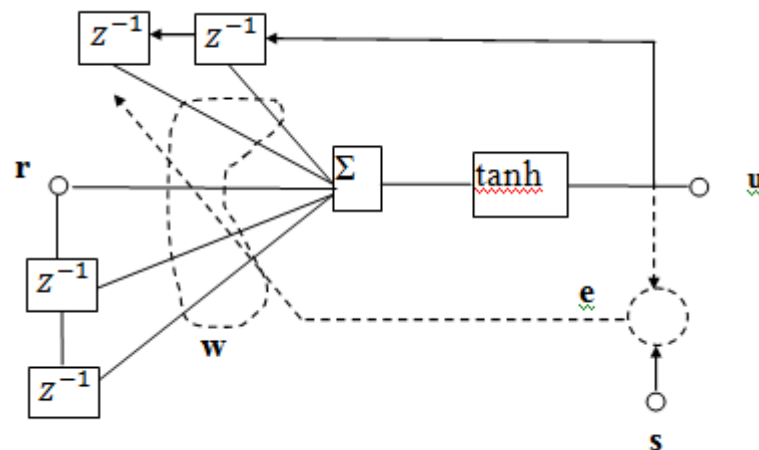


Fonte : Elaborado pelo autor.

Com o objetivo de lidar com as distorções geradas pelo canal proposto, foi empregado um equalizador neural recursivo composto por apenas dois neurônios independentes, tendo a tangente hiperbólica como função de ativação e sem o uso de *bias*. Um par de neurônios foi utilizado com a finalidade evitar a aproximação por definição $\tanh(x + j \cdot y) \triangleq \tanh(x) + j \cdot \tanh(y)$, utilizada por Leung e Haykin (1991) e Benvenuto e Piazza (1992) e analisada em Kim e Adali (2002), que pode inserir erro durante o processo de atualização dos pesos sinápticos. Dentro da estrutura proposta o sinal proveniente do canal de comunicações é dividido em partes real e imaginária e cada uma dessas partes é processada por um dos neurônios, os quais possuem a estrutura mostrada na figura 25.

Além de evitar possíveis erros gerados pela aproximação, essa abordagem reduz o custo computacional, pois evita operações entre números complexos. Parece contraditório que a utilização de estrutura que possua o dobro de pesos sinápticos possa reduzir o custo computacional, contudo essa aparente contradição pode ser esclarecida lembrando que a operação de multiplicação entre dois números complexos equivale a quatro multiplicações e duas adições entre números reais.

Figura 25- Modelo de neurônio utilizado para equalização.



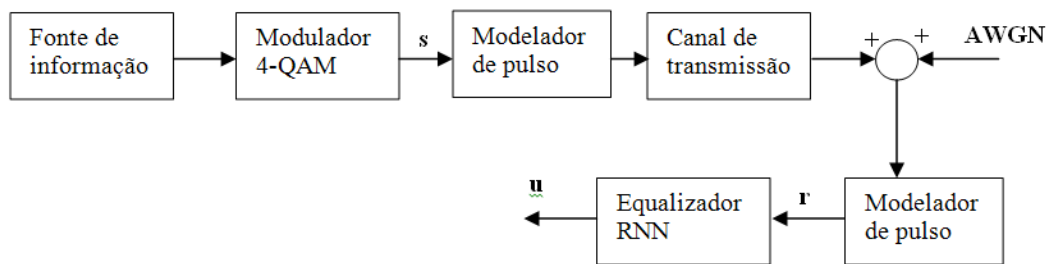
Fonte : Elaborado pelo autor.

O equalizador neural recursivo proposto é equivalente a um Equalizador com Decisão Realimentada (DFE - *Decision Feedback Equalizer*) baseado em neurônios e implementado com parâmetros $m=3$, $n=2$ e $d=2$ (CHOI; LIMA; HAYKIN, 2005). O DFE é um tipo de equalizador não linear que utiliza um filtro realimentado com n entradas em conjunto com um filtro direto com m entradas. Isso significa que a decisão do símbolo anterior é levada em

consideração na estimação do símbolo atual, com o tempo de atraso d entre o símbolo atual e o último símbolo estimado. Um DFE é geralmente aplicado em canais com interferência intersimbólica (ISI) elevada.

Em sistemas de telecomunicações, o filtro casado é utilizado para minimizar as distorções causadas pelo ruído térmico (PROAKIS, 1995). Tal sistema é instalado no transmissor e no receptor do canal de comunicações, consistindo em um par de filtros que possuem a mesma forma e trabalham em torno da frequência de modulação do equipamento. Além de eliminar o ruído fora da banda de interesse, esse aparato aplica correlação do sinal transmitido no receptor, atenuando o ruído também dentro da banda de transmissão. O modelador de pulso, que desempenha as mesmas funções do filtro casado em banda-base, é utilizado nesse trabalho, pois aqui não foi realizada modulação em frequência do sinal. Todo o sistema é mostrado na figura 26.

Figura 26- Modelo do sistema utilizado.

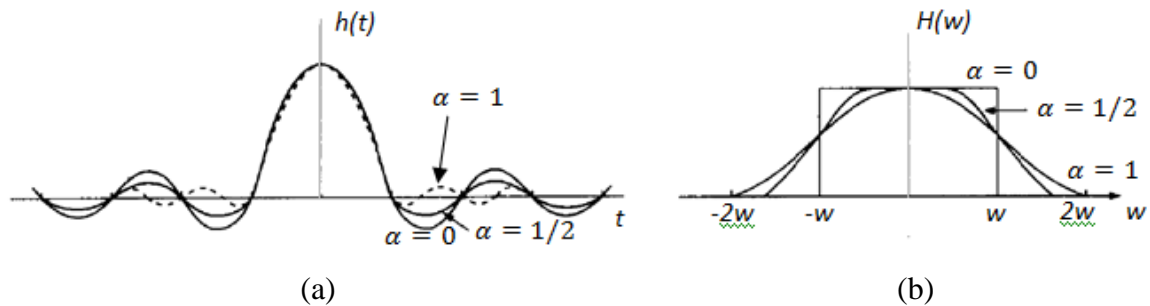


Fonte : Elaborado pelo autor

O modelador de pulso utiliza o filtro do cosseno elevado, que limita as componentes espectrais do pulso transmitido, tendo sua resposta em frequência modelada matematicamente pela equação 5.2 e a curva gerada pela função pode ser visualizada na figura 27. O fator de *roll-off*, representado por α na equação 5.2, que determina a largura de banda do sinal transmitido, foi utilizado com valor igual a 0,2, escolhido por meio de teste de simulação, de forma a prover o melhor desempenho ao sistema. A utilização de duas amostras por símbolo permite que o modelador funcione adequadamente disponibilizando a maior taxa de transmissão possível.

$$H(w) = \begin{cases} 1 & \text{para } w < w_c(1 - \alpha) \\ 0 & \text{para } w > w_c(1 + \alpha) \\ \frac{1 + \cos\left(\frac{\pi(w - w_c(1 - \alpha))}{2\alpha w_c}\right)}{2} & \text{para } w_c(1 - \alpha) > w > w_c(1 + \alpha) \end{cases} \quad (5.2)$$

Figura 27- (a) resposta no tempo e (b) pulso modelado pelo filtro do cosseno elevado.



Fonte : (PROAKIS, 1995), modificada pelo autor.

O algoritmo de Gauss-Newton, utilizado por Corral, Ludwig e Lima (2010) para atualização dos pesos sinápticos, conforme descrito na equação 5.3, é um algoritmo de otimização classificado como Quasi-Newton que não possui garantia de convergência. Nesse trabalho foi utilizada uma modificação do Gauss-Newton, proposta por Levenberg-Marquardt, que insere o termo λI , conforme descrito na equação 5.4.

$$w(k+1) = w(k) - [J(k)^T J(k)]^{-1} J(k)^T e(k) \quad (5.3)$$

$$w(k+1) = w(k) - [J(k)^T J(k) + \lambda I]^{-1} J(k)^T e(k) \quad (5.4)$$

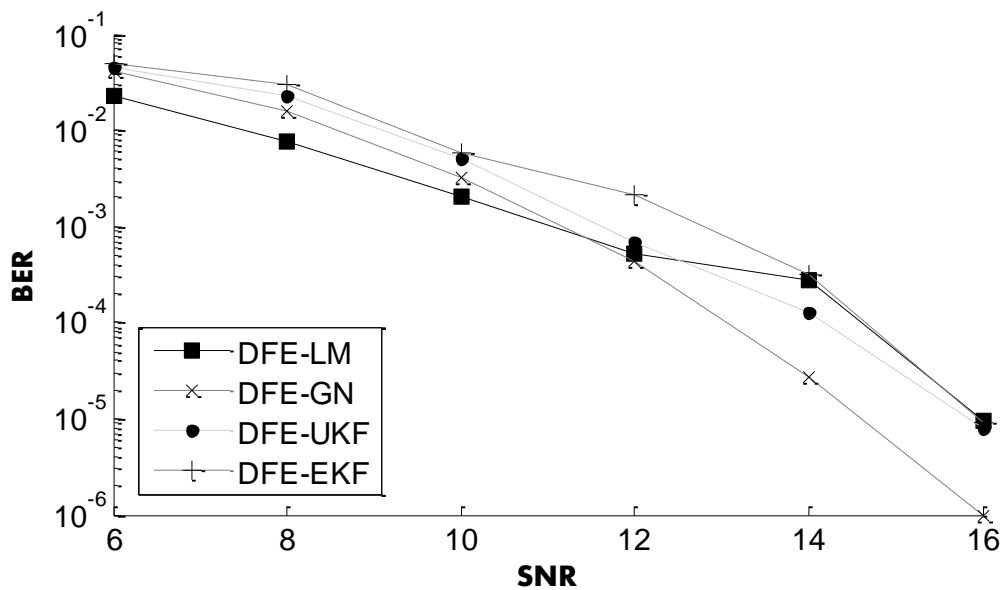
A função do fator λ é rastrear a convergência da função erro de forma que um aumento do erro implique num aumento do fator e a redução do erro gere redução do mesmo. Um valor muito alto de λ conduz função descrita pela equação 5.4 a se comportar como o algoritmo do Gradiente que possui garantia de convergência. Uma redução acentuada do fator λ leva a equação 5.4 a se aproximar do método de Gauss-Newton, descrito pela equação 5.3, que tem como característica convergência mais rápida. Dessa forma, a convergência rápida do algoritmo de Gauss-Newton é aproveitada, para os casos em que a minimização da função erro converge, e a garantia de convergência do algoritmo do Gradiente é inserida no processo de atualização dos pesos sinápticos. O conjunto obtido da união da estrutura proposta com o algoritmo de otimização de Levenberg-Marquardt foi chamada de DFE-LM (*Decision Feedback Equalizer – Levenberg-Marquardt*).

5.1 AVALIAÇÃO DE DESEMPENHO

A figura 28 mostra o desempenho do BER versus relação sinal-ruído (SNR) para a

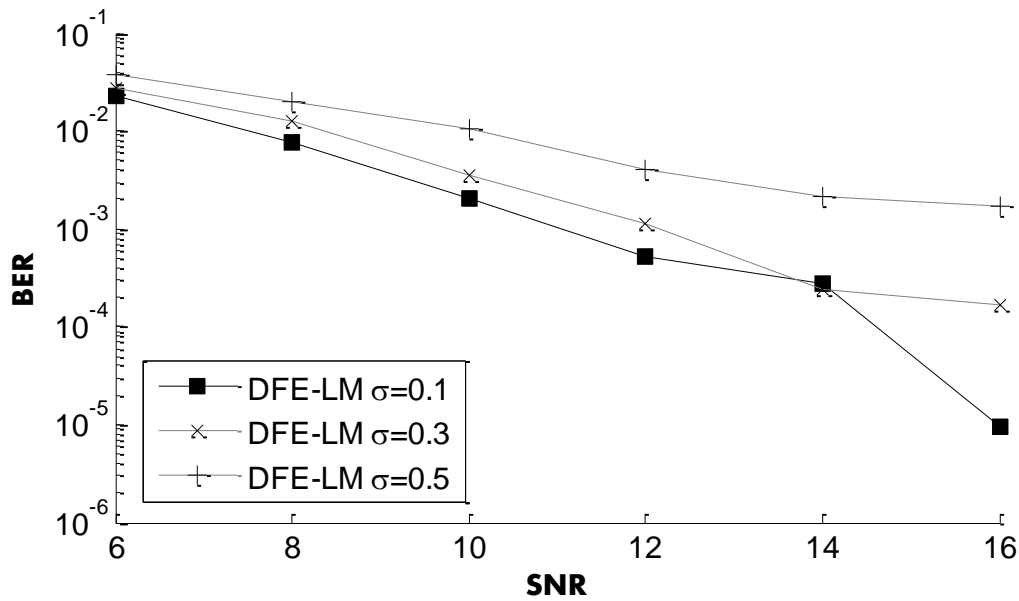
estrutura DFE-LM proposta e mais três equalizadores. Todos foram avaliados utilizando-se o modelo de canal representado pela equação 5.1 com desvio padrão dos coeficientes (σ) iguais 0,1. Para o DFE-LM foram usados 400 símbolos para treinamento e 400.000 para teste, mantendo-se a relação de 1 para 1.000, utilizada em trabalhos correlatos. Cada curva é resultado da média de 10 simulações estatisticamente independentes, mantendo-se os pesos sinápticos congelados durante a fase de teste. O DFE-LM apresentou melhor desempenho para valores mais baixos de SNR, no entanto o DFE-GN e o DFE-UKF o superaram para valores maiores de SNR, o que poderia sugerir uma falha do algoritmo utilizado no DFE-LM em lidar com essa situação. Contudo, esse comportamento pode ser explicado levando-se em consideração a existência do modelador de pulso no receptor e transmissor, que proporcionam maior redução da potência de ruído quando a relação sinal-ruído é menor.

Figura 28- Comparação de desempenho de quatro equalizadores com decisão realimentada (DFE).



Fonte : Elaborado pelo autor.

Figura 29- Desempenho do equalizador proposto sob diferentes condições do canal.



Fonte : Elaborado pelo autor.

Podem ser questionáveis as comparações entre sistemas com estruturas diferentes. O modelador de pulso poderia melhorar os resultados dos DFE-EKF e DFE-UKF, no entanto, seu alto custo computacional o torna pouco recomendável para aplicações em tempo real, nas quais a interferência intersimbólica (ISI) em canais sujeitos a desvanecimento severo precisa ser reduzida. Além disso, a questão da convergência do algoritmo GN não pode ser solucionada com a adição do modelador de pulso.

Finalmente, a figura 29 mostra o desempenho do equalizador DFE-LM utilizando-se o modelo de canal (5.1) com desvio padrão dos coeficientes iguais a 0,1, 0,3 e 0,5, mantendo as mesmas condições de simulação descritas anteriormente. Os valores de BER são aceitáveis mesmo para canais com desvio padrão de 0,5 e relação sinal ruído (SNR) de 6dB. Esse resultado mostra que o sistema de equalização proposto é uma alternativa promissora para lidar com interferência intersimbólica em canais com desvanecimento severo.

Os resultados expostos nesse capítulo foram publicados no *IJCNS: International Journal of Computer Science and Network* (MOTA; LEAL; LIMA, 2014).

6 CONCLUSÃO / TRABALHOS FUTUROS

Nesse trabalho foi proposto um equalizador inspirado em redes neurais para prover a solução, não encontrada na literatura científica, para o problema da equalização de canais sujeitos a desvanecimento severo, além de interferência intersimbólica e ruído térmico. Para atingir tal objetivo foi necessário o desenvolvimento de metodologia, também não encontrada na literatura, que auxiliou na escolha da melhor estrutura para ser aplicada ao problema proposto.

No capítulo 4, foi desenvolvido a metodologia cujo objetivo foi avaliar o desempenho de equalizadores inspirados em redes neurais, independente do algoritmo de otimização utilizado para atualização dos pesos sinápticos. Para comparar a nova abordagem a uma já empregada na literatura, foram utilizadas três estruturas neurais, aplicadas por Patra et al. (1999) para tratar as distorções produzidas por um canal não-linear num sinal 4-QAM. A comparação das abordagens conduziu a algumas conclusões importantes, tais como a de que o algoritmo *backpropagation*, proposto por Patra et al. (1999), está mais bem ajustado à estrutura FLANN e que as estruturas MLP e FLANN são capazes de atingir resultados muito próximos, em geral superiores aos da estrutura PPN. Foi possível observar também que algumas não-linearidades utilizadas por Patra et al. (1999), tal como a *tanh*, podem facilitar o trabalho do equalizador, já que separam algumas zonas de indecisão. A aplicação do método proposto ao problema de equalização do canal recomendado pela UIT mostrou-se coerente com os resultados da literatura, evidenciando que arquiteturas que apresentam a característica da recursividade têm melhor desempenho para tratar o problema e que as estruturas mais simples, com menor número de neurônios, levam nítida vantagem sobre as demais.

No capítulo 5, um par de neurônios recursivos baseados em estrutura DFE, escolhida com o auxílio da metodologia desenvolvida no capítulo 4, foi proposta para mitigar as distorções causadas em sinal 4-QAM que atravessa canal com desvanecimento severo. Um modelador de pulso foi adicionado de forma a superar as limitações do equalizador nessa tarefa. Os pesos sinápticos foram atualizados durante a fase de treinamento pelo algoritmo de Levenberg-Marquardt. A abordagem foi avaliada em um canal tradicional variante no tempo, mencionado na literatura, e o sinal foi submetido a diferentes níveis de desvanecimento. A estrutura proposta foi capaz de lidar com condições severas de variação e superou trabalhos anteriores (CHOI; LIMA; HAYKIN, 2005) e (CORRAL; LUDWIG; LIMA, 2010), na situação de baixo SNR.

Nesse trabalho, foram utilizados canais sujeitos a multipercurso discreto para testar os equalizadores propostos. Esse tipo de canal, apesar de ter sido recomendado pela UIT e TIA/EIA para simulações de sistemas móveis, são muito mais simples que canais sujeitos a multipercurso contínuo, tais como Rayleigh e Ricean.

Uma sugestão para trabalhos futuros seria a utilização da metodologia desenvolvida para encontrar a melhor estrutura, capaz de lidar com canais sujeitos a multipercurso contínuo. A substituição de estruturas inspirada em redes neurais por outras, tais como células HTM (*Hierarchical Temporal Memory*), conforme descritos por Hawkins e Blakeslee (2004), introduziria um novo paradigma para a estrutura de equalizadores. A utilização de modulação de alta ordem, tal qual 16-QAM ou 64-QAM, elevaria a complexidade do problema. Além disso, a substituição do algoritmo de atualização dos pesos neurais pode gerar ganhos de desempenho.

Apesar da metodologia desenvolvida para a comparação de estruturas neurais ter sido utilizada em problemas de equalização, ela pode ser aplicada a qualquer problema que utilize redes neurais, ou até mesmo filtros adaptativos, em suas soluções.

A metodologia proposta no capítulo 4 poderia ainda ser expandida para que não se tivesse apenas a melhor estrutura entre algumas indicadas, mas que o resultado obtido fosse mais genérico, indicando a melhor estrutura entre alguns tipos de estrutura indicadas, por exemplo, MLP, RNN, RBF, FLANN e PPN. Para isso, os seguintes procedimentos poderiam ser desenvolvidos:

- Seriam escolhidos alguns tipos de arquiteturas mais utilizadas para se resolver o problema proposto, ou até mesmo algum novo tipo de arquitetura que se queira testar, num caso hipotético poderiam ser MLP, RNN, RBF, FLANN e PPN;
- Uma arquitetura inicial (número de camadas, quantidade de neurônios, pesos sinápticos iniciais, etc.) para cada tipo de rede seria indicada, gerando, portanto, uma arquitetura MLP inicial, uma RBF inicial, uma RNN inicial, uma FLANN inicial e uma PPN inicial;
- Em cada uma das arquiteturas iniciais seria aplicado o chamado *Pruning and growing algorithm* (GUANG-BIN; SARATCHANDRA; SUNDARARAJAN, 2005), responsável por reduzir ou aumentar o tamanho de cada estrutura (número de camadas, quantidade de neurônios, etc.), com o objetivo de encontrar a melhor arquitetura dentro de cada tipo;

- A melhor estrutura de cada tipo seria então submetida à metodologia desenvolvida nesse trabalho, ou seja, o algoritmo genético seria utilizado para indicar os melhores pesos sinápticos de cada uma;
- As estruturas com seus pesos “ideais” seriam então aplicadas ao problema proposto e, de acordo com os resultados obtidos, a melhor estrutura seria eleita.

Referências Bibliográficas

- ALBU, F. et al. **Adaptive channel equalization using neural network**, Telecommunications Symposium, ITS '98 Proceedings, SBT/IEEE International, vol. 2, p. 438-441, 1998.
- BACK, A. D.; TSOI A. C. **FIR and IIR synapses, a new neural network architecture for time series modeling**, Neural Comput., vol. 3, n. 3, p. 337–350, 1991.
- BARAKAT et al. **Advanced Growing Neural Network approach for the fault diagnosis of large scale industrial plants**, 16th IEEE Mediterranean Electrotechnical Conference (MELECON), p.532-535, doi: 10.1109/MELCON.2012.6196489, 2012.
- BARUAH, B. B.; SARMA K. K. **Equalization in Faded Wireless Channels Using ANN and Coding**, International Conference on Emerging Trends in Networks in Computer Communications (ETNCC), p. 374-378, 2011.
- BENVENUTO, N.; PIAZZA, F. **On the Complex Backpropagation Algorithm**, IEEE Trans. on Signal Processing, vol. 40, n. 4, p. 967 – 969, 1992.
- BHUYAN, M.; SARMA K. K. **Nonlinear Model based Prediction of Time Varying SISO-MIMO Channels using FANN-DFE Combination**, International Conference on Emerging Trends and Applications in Computer Science (ICETACS), p. 104-107, 2013.
- BURSE, K.; YADAV, R. N.; SHRIVASTAVA, S. C. **Channel Equalization Using Neural Networks: A Review**, IEEE Trans. on Systems, Man, and Cybernetics, vol. 40, n. 3, Maio 2010.
- CHATURVEDI, S; TITRE, R.N; SONDHIIYA, N. **Review of Handwritten Pattern Recognition of Digits and Special Characters Using Feed Forward Neural Network and Izhikevich Neural Model**, International Conference on Electronic Systems, Signal Processing and Computing Technologies (ICESC), p. 425-428, doi: 10.1109/ICESC.2014.83 2014.
- CHEN, S.; GIBSON, G. J.; COWAN, C. F. N. **Adaptive channel equalization using a polynomial perceptron structure**, *Proc. Inst. Elect.Eng.*, vol. 137, n. 1, p. 257–264, Out. 1990.
- CHOI, J.; LIMA, A. C. de C.; HAYKIN, S., **Kalman filter-trained recurrent neural equalizers for time-varying channels**, IEEE Trans. Commun., vol. 3, n. 3, p. 472–480, Mar. 2005.
- CORRAL, P.; LUDWIG, O.; LIMA, A. C. de C. **Time-varying channel neural equalization using Gauss-Newton algorithm**, Electronics Letters, vol. 46, n. 15, Jul. 2010.
- DING, F.; CHEN, T. **Identification of Hammerstein nonlinear ARMAX systems**, Automatica, Elsevier, 2005.
- DING, T.; HIROSE, A. **Fading Channel Prediction Based on Combination of Complex-Valued Neural Networks and Chirp Z-Transform**, IEEE Transactions on Neural Networks and Learning Systems, vol. PP, n. 99, 2014.

DING, T.; HIROSE, A. **Fading Channel Prediction Based on Complex-Valued Neural Networks in Frequency Domain**, International Symposium on Electromagnetic Theory (EMTS), p. 640-643, 2013.

GUANG-BIN HUANG; SARATCHANDRA, P.; SUNDARARAJAN, N., **A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation**, IEEE Transaction on Neural Network, vol. 16, N. 1, p. 57-67, doi: 10.1109/TNN.2004.836241, 2005.

GUIMARÃES, A. G.; PINTO, E. L., **Técnicas de Simulação de Canais Rádio Móvel**, Revista Científica Periódica – Telecomunicações, ISSN 1516-2338, Vol. 4, N. 1, p. 24-34 Maio 2001.

HAO LI et al, **Application of Artificial Neural Networks in predicting abrasion resistance of solution polymerized styrene-butadiene rubber based composites**, IEEE Workshop on Electronics, Computer and Applications, p. 581-584, doi: 10.1109/IWECA.2014.6845687, 2014.

HAWKINS, J.; BLAKESLEE, S., **On Intelligence**, New York: Times Books, 2004.

HAYKIN, S. **Adaptive Filter Theory**, Englewood Cliffs, NJ, Prentice-Hall, 3 edition, 1996.

HAYKIN, S, **Communication Systems**, 4 Ed., NY , 2004.

HAYKIN, S., **Digital Communications**, Wiley, 1988.

HAYKIN, S. **Neural networks: a comprehensive foundation**, Ontario, Canadá: Prentice Hall, 1994.

HOLTZ, G. C da C. **Traçado automático de envoltórias de esforços em estruturas planas utilizando um algoritmo evolucionário**, Dissertação de Mestrado em Engenharia Civil. Rio de Janeiro: PUC-Rio, 2005.

ITU-R M.1225 Recommendation, **Guideline for Evaluation of Radio Transmission Technologies for IMT-2000**, International Telecommunication Union, 1997.

JERUCHIM, M. C.; BALABAN, P.; SHANMUGAN, K. S., **Simulation of Communication Systems**, New York, Kluwer Academic/Plenum, Second edition, 2000.

JUAN, T.J.; CARLOS, R.-R.; JORGE, R., **Automatic Behavior Generation in a Multi-Agent System through Evolutionary Programming**, Robotics Symposium. LARS '06. IEEE 3rd Latin American, doi: 10.1109/LARS.2006.334316, 2006.

KECHRIOTIS, G.; ZERVAS, E.; MANOLAKOS, E. S. **Using Recurrent Neural Networks for Adaptive Communication Channel Equalizations**, IEEE Trans. Neural Net., Vol. 5, n. 2, p. 267–278, 1994.

KIM, T.; ADALI, T. **Fully Complex Multi-Layer Perceptron Network For Nonlinear Signal Processing**, The Journal of VLSI Signal Processing, Springer, 2002.

LAWRENCE, S.; GILES, C. L.; TSOI, A. C. **Lessons in Neural Network Training: Overfitting May be Harder than Expected**, Proceedings of the Fourteenth National Conference on Artificial Intelligence, AAAI-97, AAAI Press, Menlo Park, California, p. 540–545, 1997.

LEUNG, H.; HAYKIN, S. **The Complex Backpropagation Algorithm**, IEEE Trans. on Signal Proc., vol. 3, n. 9, p. 2101 – 2104, 1991.

LING, F.; PROAKIS, J. G. **Adaptive Lattice Decision-Feedback Equalizer - Their Performance and Application to Time-Variant Multipath Channels**, IEEE Trans. Commun. vol. COM-33, p.348–356, Abr. 1985.

MAHDAVIANI, K. et al. **A Method to Resolve the Overfitting Problem in Recurrent Neural Networks for Prediction of Complex Systems' Behavior**, International Joint Conference on Neural Networks, p. 3723-3728, 2008.

MEYER, M.; PFEIFFER, G. **Multilayer perceptron based decision feedback equalisers for channels with intersymbol interference**, Communications, Speech and Vision, IEEE Proceedings I, vol. 140, n. 6, p. 420-424, 1993.

MOTA, T.; LEAL, J.; LIMA, A., **Fast Fading Channel Neural Equalization Using Levenberg-Marquardt Training Algorithm and Pulse Shaping Filters**, Int'l J. of Communications, Network and System Sciences, vol. 7, N. 2, p. 71-74, doi: 10.4236/ijcns.2014.72008, 2014.

NING GAO; CAI-YUN GAO, **Combining the genetic algorithms with BP Neural Network for GPS height Conversion**, International Conference on Computer Design and Applications (ICCD), vol. 2, p. V2-404 – V2-408, doi: 10.1109/ICCD.2010.5541393 2010, 2010.

PATRA, J. C. et al. **Nonlinear channel equalization for QAM signal constellation using artificial neural networks**, IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 29, n. 2, p. 262–271, Abr. 1999.

PAPOULIS, A.; PILLAI, S., **Probability, Random Variables and Stochastic Processes**, McGraw Hill, 4 edition, 2001.

PEIXIAN LI et al., **Time Series Prediction of Mining Subsidence Based on Genetic Algorithm Neural Network**, Symposium on Computer Science and Society (ISCCS), 2011 International, p. 83-86, doi: 10.1109/ISCCS.2011.30, 2011.

PENG, M.; NIKIAS, C.L.; PROAKIS, J.G., **Adaptive equalization with neural networks: new multi-layer perceptron structures and their evaluation**, IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 2, p. 301-304, 1992.

PROAKIS, J. G. **Digital Communications**, New York: McGraw Hill, 1995.

SCHAFFER, J. D.; WHITLEY, D.; ESHELMAN, L.J., **Combinations of Genetic Algorithms and Neural Networks: A Survey of the State of the Art**, . International

- Workshop on Combinations of Genetic Algorithms and Neural Networks, 1992., COGANN-92, doi: 10.1109/COGANN.1992.273950, p. 1-37, 1992.
- SHAFI, A.; ZERGUINE, A.; BETTAYEB, M. **Neural network-based decision feedback equalizer with lattice structure**, *Electronics Letters*, p. 1705-1707, 1999.
- SHEEJA, K.L. ; SAHU, P.K. **Decision Feedback Equalization Using RBF and MLP Networks**, *International Conference on Information Technology for Real World Problems*, p. 96-101, doi: 10.1109/VCON.2010.27, 2010.
- SIGTIA, S.; DIXON, S. **Improved music feature learning with deep neural networks**, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, p. 6959 a 6963, doi: 10.1109/ICASSP.2014.6854949, 2014.
- SIU, S.; GIBSON, G.J.; COWAN, C.F.N. **Decision Feedback Equalisation Using Neural Network Structures and Performance Comparison with Standard architecture**, *Communications, Speech and Vision, IEE Proceedings I*, vol. 137, n. 4, p. 221-225, 1990.
- SIVAKUMAR, P. et al. **Performance Evaluation of Channel Equalization by using MNN Rayleigh Fading Channel**, *International Conference Electronics Computer Technology (ICECT)*, vol. 4, p. 47-51, 2011.
- SOARES, G. L. **Algoritmos Genéticos: Estudo, Novas Técnicas e Aplicações**, Dissertação de Mestrado em Engenharia Elétrica. Minas Gerais: UFMG, 1997.
- TIA/EIA Standard 136-270-B, **TDMA Third Generation Wireless – Mobile Station Minimum Performance**, *Telecommunications Industry Association*, Mar 2000.
- XIAOHUI HUANG et al. **Application of Multilayer Feedforward Neural Networks in predicting tree height and forest stock volume of Chinese fir**, *IEEE Workshop on Electronics, Computer and Applications*, p. 610-613 doi: 10.1109/IWECA.2014.6845693, 2014.
- XINGCHENG PU; YANQIN LIN; PENGFEI SUN. **A Pruned Cooperative Co-evolutionary Genetic Neural Network and its Application on Stock Market Forecast**, *The 26th Chinese Control and Decision Conference (2014 CCDC)*, p. 2344-2349, doi: 10.1109/CCDC.2014.6852564, 2014.
- YUNDUAN CUI ; TAKAHASHI, K. ; HASHIMOTO, M. **Remarks on robot controller application of Clifford multi-layer neural networks**, *IEEE 13th International Workshop on Advanced Motion Control (AMC)*, p. 410-415, doi: 10.1109/AMC.2014.6823317, 2014.



**TERMO DE AUTORIZAÇÃO PARA DISPONIBILIZAÇÃO
DAS TESES E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFBA**

Eu, TIAGO ANDRADE MOTA
 Nacionalidade: BRASILEIRO, estado civil: CASADO, profissão: ENG. ELETRICISTA, residente e domiciliado na (endereço): RUA PLINIO MOSCOSO, Nº 627, AP. 902 A, cidade: SALVADOR, estado: BAHIA, portador do documento de identidade nº 08321178-03, CPF nº 798167.545-68, na qualidade de titular dos direitos morais e patrimoniais de autor da OBRA (título) SELEÇÃO DE ESTRUTURA NEURAL PARA O PROBLEMA DE EQUALIZAÇÃO UTILIZANDO ALGORITMO GENÉTICO dissertação de (nível) MESTRADO ou tese de (nível) _____, apresentada na Universidade Federal da Bahia – UFBA em (data) _____, com base no disposto da *Lei Federal nº 9.610, de 19 de fevereiro de 1998*:

DECLARO, dessa forma, que **cedo, em caráter gratuito e por tempo indeterminado**, minha tese ou dissertação acima identificada, cuja cópia, segue em anexo, para que seja reproduzida ou divulgada, total ou parcialmente no site: <http://www.bdttd.ufba.br/>

DECLARO, que sou o **único e exclusivo responsável**, pelo conteúdo da mencionada tese ou dissertação acima referida.

a- **AUTORIZO** a Universidade Federal da Bahia/ Sistema de Bibliotecas a disponibilizar, gratuitamente, sem ressarcimento dos direitos autorais, o texto integral da publicação acima citada, de minha autoria, no site: <http://www.bdttd.ufba.br/> e em forma de arquivo PDF, para fins de leitura e/ou impressão pela Internet, a título de divulgação da produção científica gerada pela Universidade, a partir de 01/12/2014.

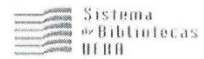
b- **AUTORIZO**, ainda, a Universidade Federal da Bahia/ Sistema de Bibliotecas a **remover** a tese ou dissertação referida acima do site: <http://www.bdttd.ufba.br/> a **qualquer tempo e independentemente de motivo e/ou notificação prévia** à minha pessoa.

Salvador, 02/12/2014
Local Data

Tiago Andrade Mota
Ass. do (a) Autor (a)/Representante legal.

Preenc

her em três vias. A primeira via deste formulário deve ser encaminhada ao Sistema de Bibliotecas da UFBA/Biblioteca Central; a segunda deve ser enviada para a Biblioteca de sua Unidade, juntamente com o arquivo contendo o documento; a terceira via deve permanecer no Programa de Pós-Graduação para o registro do certificado de conclusão do Curso.



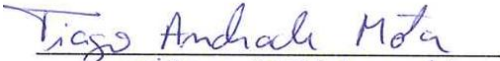
CADASTRO DE INFORMAÇÕES PARA PUBLICAÇÃO DIGITAL
NA BIBLIOTECA DIGITAL DA UFBA

1. Identificação do tipo de material	
Tese () Dissertação (X) Monografia () Trabalho de Conclusão de Curso ()	
2. Colegiado do Curso de Pós-Graduação:	
Título: SELEÇÃO DE ESTRUTURA NEURAL PARA O PROBLEMA DE EQUALIZAÇÃO UTILIZANDO ALGORITMO GENÉTICO	
Autor(a): TIAGO ANDRADE MOTA	
CPF: 798.167.545-68	E-mail: TANDRADEMOTA@GMAIL.COM
Orientador(a):	
Nome: ANTONIO CEZAR DE CASTRO LIMA	
CPF: 268.879.915-00	E-mail: ACDCL@UFBA.BR
Co-Orientadores	
Nome:	
CPF:	E-mail
Membros da Banca	
Nome: KARCIUS DAY ROSÁRIO ASSIS	
CPF: 884.346.305-59	E-mail: KARCIUS.ASSIS@UFBA.BR
Nome: ÂNGELO AMÂNCIO DUARTE	
CPF: 365.265.415-34	E-mail: ANGELODUARTE@ECOMP.UFES.BR
Nome: LEIZER SCHNITMAN	
CPF: 386.828.695-00	E-mail: LEIZER@UFBA.BR
Nome:	
CPF:	E-mail
Data de Homologação Pós Graduação:	
Financiadores:	
Data:	
Assinatura:	

Salvador, 30/10/2014

DECLARAÇÃO

Declaro para os devidos fins que o texto final apresentado para a conclusão do meu curso de Mestrado em Mecatrônica da Universidade Federal da Bahia é de minha autoria. Declaro também que quaisquer informações utilizadas neste texto, mas que sejam provenientes de outros trabalhos tem fonte claramente expressa e, quando for o caso, foram devidamente autorizadas pelo(s) respectivo(s) autor(es).


Nome: TIAGO ANDRADE MOTA
CPF: 798.167.545-68