

1º Congresso Brasileiro de Redes Neurais

Escola Federal de Engenharia de Itajuba
Itajuba, 24 a 27 de outubro de 1994

Redes RAM Radial

ANNE MAGÁLY DE PAULA CANUTO¹
EDSON COSTA DE BARROS CARVALHO FILHO¹

¹UFPE-Universidade Federal de Pernambuco
Departamento de Informática
Cx 7851, 50.732-970, Recife, PE, Brasil
ampc@di.ufpe.br e ecdbcf@di.ufpe.br

Sumário. O modelo de redes neurais conhecido como RAM-Based é capaz de computar qualquer função booleana com um dado número de entradas. Neste artigo, é proposto e investigado o modelo RAM Radial, que é uma generalização do RAM original, diferenciando no estado de uso. Neste estado, quando uma entrada é apresentada ao neurônio, não apenas o conteúdo endereçado, conhecido como ponto central, é acessado, mas sim toda uma região radial. A análise do desempenho da rede mostra que o RAM Radial obteve melhores resultados, sendo sempre *mais confiável* que o RAM original. As implicações destes resultados vão além da área de redes neurais, aplicando-se na área de reconhecimento de padrões com uma generalização da técnica de n-tupla.

1 Introdução

Os modelos de redes neurais são inspiradas no sistema nervoso biológico e são compostos de elementos (neurônios) que funcionalmente são análogos a modelos simplificados de neurônios biológicos. Existe uma classe de redes neurais, conhecida como RAM devido aos seus neurônios terem uma estrutura similar a de uma memória de acesso aleatório, que é capaz de computar qualquer função booleana com um dado número de entradas [3, 5].

Neste artigo, será apresentado o modelo de rede RAM Radial, que é uma generalização do RAM original diferindo na forma de acesso aos conteúdos. O modelo proposto não acessa apenas um endereço para calcular a saída, mas sim toda uma região radial deste endereço. A proposta do modelo RAM Radial emerge de estudos no modelo Goal Seeking Neuron (GSN) [1, 2].

Este trabalho está dividido em quatro partes: a primeira parte será realizada uma breve descrição do modelo RAM radial; a segunda parte será mostrada a arquitetura da rede; a terceira parte será mostrada uma análise do desempenho dos experimentos realizados com a rede; e a quarta e última parte serão feitas as considerações finais sobre o modelo.

2 Neurônio RAM Radial

O modelo RAM [3] pode ser visto como uma técnica de implementação de um processo de reconhecimento primeiramente descrito por Bledsoe e Browning em 1959 [9]. Este processo é conhecido como o método da n-tupla. O termo **n-tupla** vem do fato que cada unidade (neurônio) recebe n entradas como uma tupla.

O neurônio RAM radial se assemelha muito com

o neurônio RAM no fato que seus neurônios recebem valores binários e são capazes de computar qualquer função booleana com um dado número de entradas.

A diferença entre os modelos RAM Radial, que está sendo analisado neste trabalho, e o RAM está no estado de uso. Neste estado, quando uma entrada é apresentada ao neurônio, não apenas o conteúdo endereçado, conhecido como ponto central, é acessado, mas sim toda uma região radial (região de acesso). Esta região é composta pelos conteúdos cujo endereço está até uma determinada distância do ponto central, conhecidos como conjunto endereçáveis.

Se os pesos forem rotulados como 2^k onde $0 \leq k \leq c_i - 1$, então o ponto central, a_{ik}^c , acessado por uma entrada é definido pela equação 1.

$$a_{ik}^c = \sum_{j=1}^{j=c_i} x_{ij} z_{ij} \quad (1)$$

onde x_{ij} é o valor da entrada do j-ésimo terminal de entrada e z_{ij} é o peso do j-ésimo terminal.

O cálculo de todos os endereços do conjunto endereçável pode ser visto na equação 2.

$$A_{im} = \{a_{im} = \sum_{j=1}^{j=c_i} x_{ij} z_{ij}, \forall a_{im} \mid D(a_{im}, a_{ik}^c) \leq R\} \quad (2)$$

onde :

- . R é o valor da distancia máxima;
- . c_i é a conectividade do neurônio;
- . a_{ik}^c é o endereço do ponto central da região;
- . a_{im} são os endereços dos conteúdos da região;

Este conjunto endereçável pode ser genérico, como todos os pontos de uma esfera ou cubo, ou analisando a distribuição de probabilidade do conjunto de

dados. Uma vez que os endereços estão distribuídos ao longo de uma distância escolhida para a formação da região, utiliza-se uma função de ponderação para diferenciar a contribuição dos conteúdos acessados na saída do neurônio. Vários tipos de funções podem ser utilizadas, tais como: a função linear, a função gaussiana, a função hiperbólica, etc. que encontram presupostos biológicos e são utilizadas nos modelos de Linsker e Kohonen [7, 8].

A função de ativação é computada em cima do conjunto endereçável com seus respectivos pesos. Uma forma genérica que pode ser útil no cálculo da saída do neurônio radial é dada na equação 3.

$$o_i = F(s_i) \quad (3)$$

onde :

$$s_i = \sum_{\forall a_{im} \in A_i} G(C[a_{im}], D(a_{im}, a_{ik}^c))$$

e :

- o_i é a saída do neurônio i ;
- $C[a_{im}]$ é o conteúdo do endereço a_{im} ;
- F é a função de ativação;
- G é a função de ponderação;

A função G pondera valores com o objetivo de valorizar alguns conteúdos ($C[a_{im}]$) da região, dependendo da distância (D) do ponto central.

3 Uma Arquitetura para Reconhecimento

Uma arquitetura que pode ser utilizada com o modelo RAM Radial é uma rede de discriminadores. Esta rede é formada por apenas uma camada de neurônios. A escolha desta rede é porque ela se assemelha com o método da n-tupla. A utilização do modelo RAM Radial nesta arquitetura representa uma modificação na técnica de n-tupla original.

Na arquitetura de discriminadores de uma camada, o processo de aprendizagem é supervisionado, associando cada discriminador com uma classe de padrão [6]. A figura 1 ilustra a arquitetura utilizada na rede.

Neste tipo de arquitetura, cada discriminador possui um número fixo de neurônios e é associado com uma classe para aprendê-la, e os neurônios possuem uma conectividade fixa [5, 6].

3.1 A Fase de Aprendizagem

No início, todos os neurônios de todos os discriminadores são zerados. Quando a rede entra na fase de aprendizagem, ela associa um discriminador com uma classe de padrão. Nesta fase, associa-se todos os padrões de treinamento de uma classe a um discriminador e ensina a este discriminador aquela classe, este ensino é feito colocando 1 nos conteúdos acessados pelas entradas em todos os neurônios do discriminador selecionado, ou seja, apenas o discriminador selecionado aprende a classe.

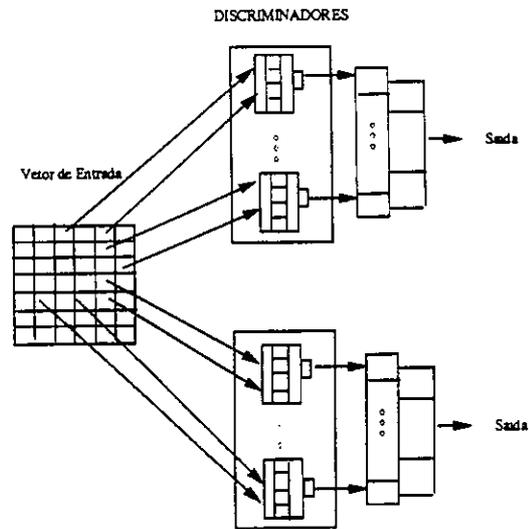


Figura 1: A Arquitetura de discriminadores com RAM Radial

3.2 A fase de teste

Quando um padrão de teste é apresentado a rede, todos os neurônios de todos os discriminadores são ativados e produzem uma saída, esta saída indica o quanto similar a região de acesso é da entrada. A saída do discriminador é calculada somando todas as saídas dos seus neurônios. O discriminador que possui a maior saída é aquele que melhor representa o padrão de entrada.

4 Análise do desempenho

Para investigar as propriedades fundamentais e habilidades da rede RAM Radial, um conjunto de dados foi selecionado para ser utilizado nos experimentos. O conjunto de dados escolhido é formado por caracteres digitalizados alfanuméricos de 0 a 9 podendo ser estendido para todas as letras e caracteres manuscritos. Cada classe contém 300 padrões digitalizados em 384 pontos brancos e pretos numa matriz de 24 X 16.

Nestes experimentos, a função de ativação utilizada foi a função sigmoide e a função de ponderação foi a gaussiana. A função de ponderação valoriza os pontos que estão próximos do ponto central. As funções podem ser vista nas equações 4 e 5 respectivamente.

$$F(s_i) = \frac{1}{1 + \exp(-s_i/\theta)} \quad (4)$$

$$G(x, y) = \begin{cases} \exp(\frac{-y^2}{2\sigma^2}) & \text{se } C_{ij} \text{ for } 1 \\ -\exp(\frac{-y^2}{2\sigma^2}) & \text{se } C_{ij} \text{ for } 0 \end{cases} \quad (5)$$

onde:

θ estabelece a largura da sigmoide

δ é o sigma da gaussiana e estabelece a largura da gaussiana.

O desempenho da rede será analisado em termos de respostas corretas, erradas e rejeitadas, similaridade e confiabilidade. A similaridade é calculada através do valor do desempenho de respostas corretas do discriminador vencedor de todos os padrões. A confiabilidade é calculada pela diferença de desempenho do discriminador vencedor e o segundo maior discriminador.

As próximas seções demonstrarão a análise comparativa do desempenho das redes com a variação do conjunto de treinamento, o sigma da gaussiana, o número de neurônios por discriminador, a conectividade do neurônio e o peso dos conteúdos radiais.

O objetivo desta análise é verificar se ocorreu uma melhora no nível de reconhecimento utilizando o modelo RAM Radial que representa uma contribuição no modelo RAM e n-tupla.

4.1 Número de neurônios por discriminadores

Com o objetivo de analisar o desempenho da rede variando a quantidade de neurônios por discriminador, foram implementadas várias redes, todas elas possuíam 10 discriminadores onde cada discriminador possuía 10, 50, 100, 150 e 200 neurônios por discriminador com conectividade 4. Estas redes foram treinadas com 64 padrões por classe e recuperadas com 100 padrões por classe. Os padrões foram apresentados na ordem de 0 a 9, apresentando todos os padrões da classe 0, depois todos da classe 1 e vai até todos da classe 9. O sigma da gaussiana utilizada para cálculo dos pesos foi 0.5.

A figura 2 ilustra o desempenho de reconhecimento em termos de respostas corretas, erradas e rejeitadas. A simbologia RB e RBR utilizada a partir desta figura significa RAM e RAM Radial respectivamente.

Os resultados obtidos mostram que com o número de neurônios igual a 10, a diferença de desempenho das duas redes praticamente não existe, isto é devido a pequena área coberta, 40/384. Com o número de neurônios igual a 50 esta diferença aumenta devido a um melhor desempenho de respostas corretas da RAM Radial. No entanto, com número de neurônios maior que 50 esta diferença vai diminuindo até ficar constante.

Agora, será visto o desempenho da rede em relação a similaridade e a confiabilidade. A figura 3 ilustra o desempenho das redes.

Os resultados obtidos nesta análise mostram que as similaridades da RAM Radial está quase sempre abaixo da RAM, apenas na similaridade de rejeição que a rede é maior com 10, 50, abaixando logo em seguida. Isto ocorre pois para uma neurônio responder

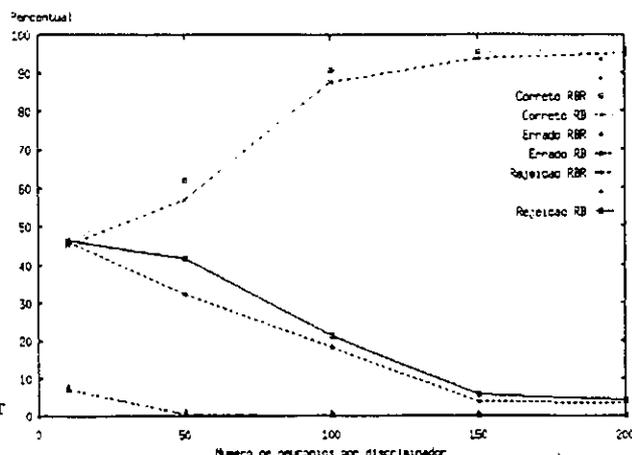


Figura 2: O Desempenho de respostas corretas erradas e rejeitadas com variação do número de neurônios

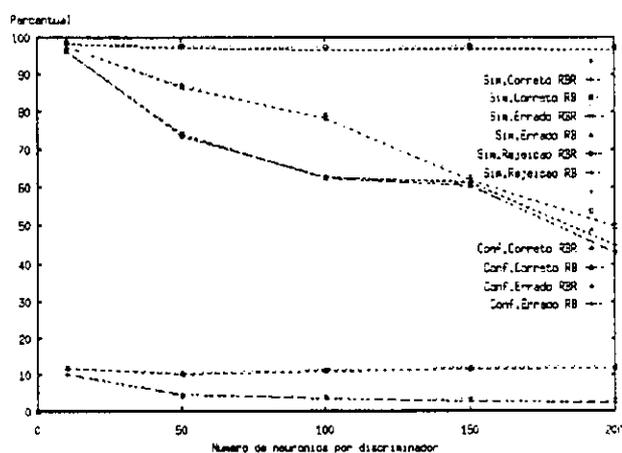


Figura 3: O Desempenho de similaridade e Confiabilidade com variação do número de neurônios

1. todos os endereços da região devem conter 1 e isso dificulta a resposta 1 dos neurônios, geralmente a saída fica entre 0.8 e 0.9, já o RAM normal responde 1 se o seu conteúdo possuir 1. Agora, a confiabilidade do RAM radial é sempre melhor que o RAM-based, esta diferença se torna maior com o aumento do número de neurônios.

A primeira vista, se pode concluir que o melhor desempenho da RAM Radial é com o número de neurônios igual a 50, no entanto, a área coberta desta rede é pequena, 200/384, e apenas pouco mais da metade do padrão está sendo pesquisado pelos neurônios. Diante disso, pode-se notar que, em relação ao número de neurônios, o melhor desempenho da RAM Radial é com um número de neurônios igual a 100.

4.2 O Conjunto de Treinamento

Para verificar o desempenho da rede com diferentes tamanhos conjuntos de treinamento, a rede descrita

anteriormente foi treinada com 1, 2, 4, 8, 16, 32 e 64 padrões por classe. Ao verificar os resultados, figura 4, pode-se ver que com o aumento do conjunto de treinamento, a diferença de desempenho, favorecendo o RAM Radial, vai aumentando. O aumento dessa diferença ocorre devido ao número crescente de 1's nos conteúdos dos endereços da região de acesso, dos discriminadores treinados com aquele padrão, com o aumento do conjunto de treinamento. Ao verificar o desempenho de respostas erradas e rejeitadas, pode-se notar que está havendo uma troca, as respostas que eram erradas, com o aumento do conjunto de treinamento, estão sendo respostas rejeitadas.

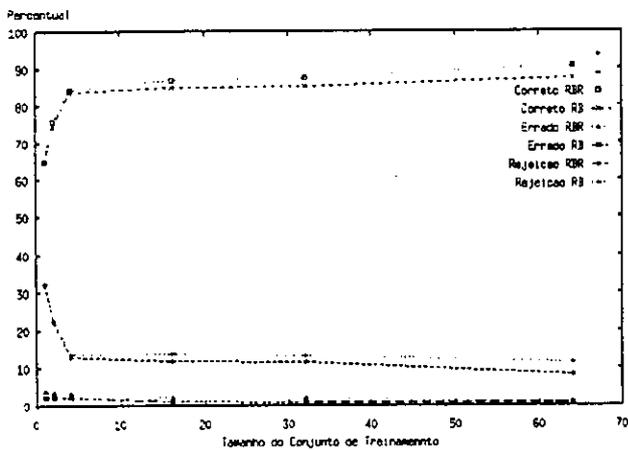


Figura 4: O desempenho de respostas corretas erradas e rejeitadas com variação do conjunto de treinamento

A figura 5 ilustra o desempenho da rede em relação a similaridade e confiabilidade. Os resultados mostram que com o aumento do conjunto de treinamento a diferença de desempenho, no início o RAM é melhor, vai diminuindo até praticamente não existir diferença. Agora, ao verificar a confiabilidade pode-se notar que com os conjuntos de treinamento pequenos, a RAM Radial é menos confiável que o RAM, mas com o aumento do conjunto de treinamento o RAM Radial passa a ser mais confiável.

Após toda essa análise, pode-se concluir que a rede que obteve os melhores resultados foi a rede treinada com 64 padrões por classe pois tem o maior desempenho de respostas corretas e é a mais confiável.

4.3 A Conectividade

Para analisar o desempenho da rede com conectividades diferentes, foram treinadas várias redes, semelhantes à descrita anteriormente, com conectividade 2, 4, 6 e 8, com área coberta de 200/384, 400/384, 600/384 e 800/384 respectivamente.

Os resultados, figura 6, mostram que no início o RAM Radial tinha melhor percentagem de respostas

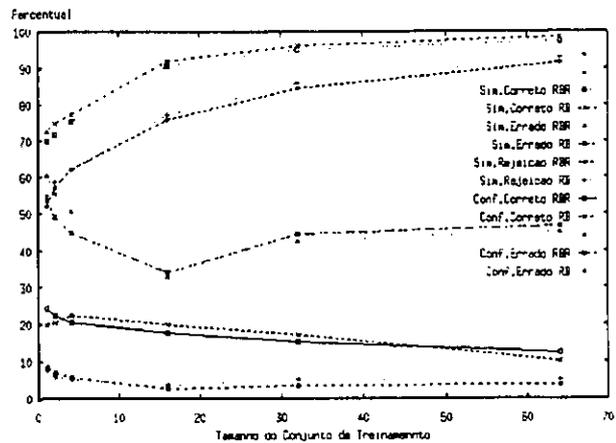


Figura 5: O Desempenho de similaridade e Confiabilidade com variação do conjunto de treinamento

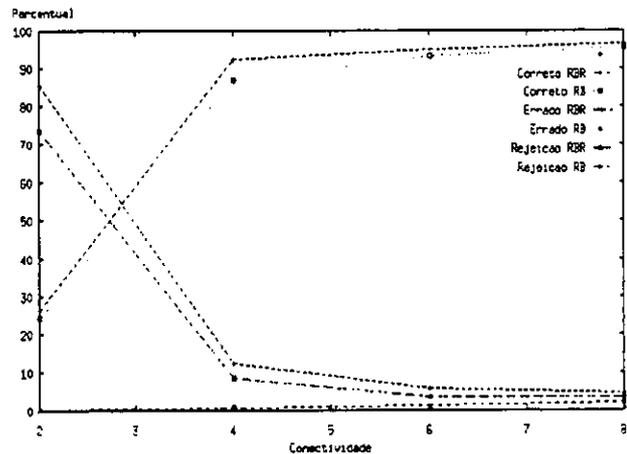


Figura 6: O Desempenho de respostas corretas erradas e rejeitadas com variação da conectividade

corretas e com o aumento da conectividade o RAM vai se aproximando do RAM Radial até praticamente não existir diferença com conectividade 8. Esta diminuição ocorre pois com o aumento da conectividade há um aumento exponencial nos endereços dos neurônios e como o conjunto de treinamento e a quantidade de endereços acessados continuam fixos então ocorre um esparsamento de endereços acessados dentro do neurônio, é o mesmo processo que ocorre com conjuntos de treinamento pequenos.

A figura 7 ilustra o comportamento da rede em relação a similaridade e confiabilidade. Os resultados obtidos mostram que todas as similaridades e confiabilidades não tem diferença entre as duas redes com conectividade baixa e essas diferenças são diretamente proporcionais ao aumento da conectividade, sendo o RAM-Based Radial mais confiável e menos similar.

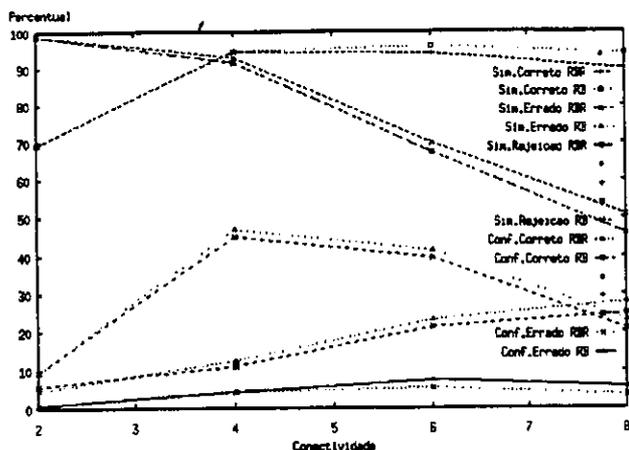


Figura 7: O Desempenho de similaridade e Confiabilidade com variação da conectividade

Com esses resultados, pode-se concluir que apesar de ser menos confiável, a rede com conectividade 4 é a melhor configuração pois as similaridades estão mais próximas do RAM e com uma conectividade maior há um maior consumo de tempo e memória.

4.4 O Peso dos conteúdos radiais

A variação dos pesos dos conteúdos radiais é calculada através da variação do sigma da gaussiana. Com o objetivo de analisar o desempenho da rede variando o sigma da gaussiana, foi implementada a rede descrita anteriormente variando o sigma de 0.1 a 1.0 com passos de 0.1. Com o sigma de 0.1, a gaussiana é bem fechada e o pesos para os endereços com distância de Hamming igual ou maior que 1 é 0, que é o mesmo comportamento da RAM. A figura 8 ilustra o comportamento da rede.

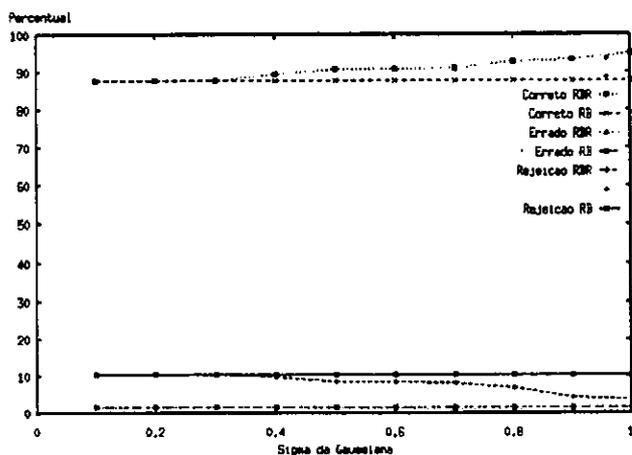


Figura 8: O Desempenho de respostas corretas, erradas e rejeitadas com variação dos pesos

Os resultados obtidos mostram que com o aumento do sigma, aumenta a diferença de desempenho, com

o RAM Radial tendo melhor desempenho de respostas corretas. Com esse aumento, houve uma diminuição de respostas erradas e rejeitadas.

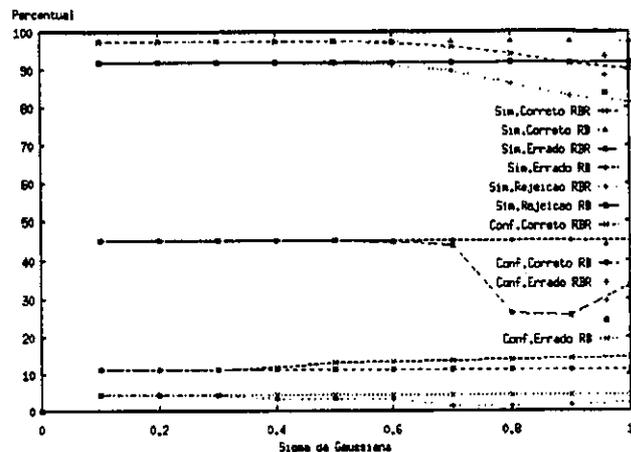


Figura 9: O Desempenho da similaridade e confiabilidade com variação dos pesos

A primeira vista, o melhor desempenho seria a rede com sigma igual a 1.0, mas ao observar o desempenho de similaridade e confiabilidade, figura 9, pode-se notar que a confiabilidade diminui com a diminuição do sigma mas a similaridade do RAM Radial vai diminuindo com o aumento do sigma, tendo uma diferença razoável com sigma igual a 1.0, em favor do RAM. Com isso, é mais adequado escolher um sigma menor pois tem uma melhor similaridade e também é confiável.

4.5 O Raio da Gaussiana

O raio da gaussiana define a esfera de distância de endereço da região de acesso. Com o objetivo de analisar o desempenho da rede variando o raio da gaussiana, foi implementada a rede descrita anteriormente variando o raio de 1, 2 e 3. Para avaliar estes experimentos, foi utilizado um sigma de 1.0 pois se o sigma for menor, o peso para distância de Hamming igual a 3 seria muito pequeno.

Ao verificar os resultados, figura 10, pode-se notar que o desempenho da rede é inversamente proporcional ao raio, ou seja, o melhor desempenho é obtida com um raio da gaussiana de 1. Isto ocorre pois com o raio igual a 1, a região de acesso tem um tamanho. Com o aumento do raio, há um aumento considerável no tamanho da região de acesso e ocorre a inclusão de endereços não muito similares a entrada, e isso dificulta a resposta dos neurônios.

Ao observar o desempenho da rede em relação a similaridade e confiabilidade, figura 11, pode-se ver que os melhores resultados são obtidos com raio igual a 1 pois a rede é mais confiável e possui as similaridades

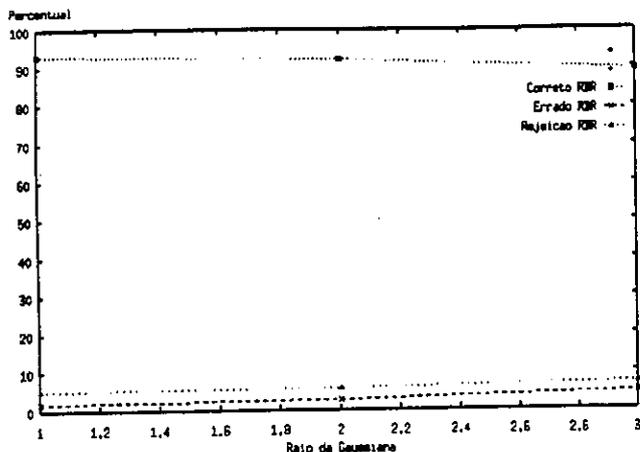


Figura 10: O desempenho de respostas corretas erradas e rejeitadas com variação do raio da gaussiana

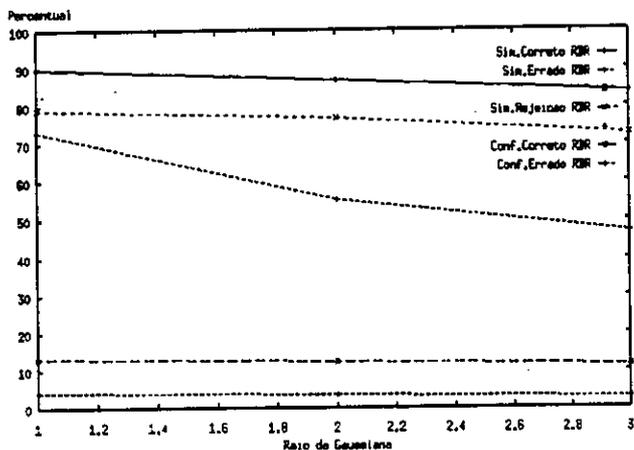


Figura 11: O Desempenho da similaridade e confiabilidade com variação do raio da gaussiana

mais próximas das da RAM. Novamente, esta queda ocorre devido a um aumento da região de acesso.

Naturalmente, é observado neste modelo que o uso de conectividades altas, tais como 10, 30 e 50, e conjuntos de treinamento alto, as redes com raios maiores teriam melhores resultados pois com o aumento exponencial na quantidade de endereços há uma necessidade de aumentar a região de acesso e o conjunto de treinamento. Com raios maiores, essa região aumenta e se terá melhores resultados.

5 Conclusão

O modelo proposto neste artigo é uma generalização do RAM original. Os modelos diferem na forma de acesso ao conteúdo. O modelo RAM Radial não acessa apenas um endereço para calcular a saída, mas sim toda uma região radial deste endereço

Para os experimentos realizados, a arquitetura utilizada é uma rede de discriminadores formada por

apenas uma camada de neurônios.

A análise de desempenho mostrou que em todos os parâmetros, o RAM Radial obteve melhores reconhecimento de resposta corretas que o RAM. O RAM Radial possui uma menor similaridade mas **sempre mais confiável** que o RAM. Um dos parâmetros que ocorreu uma melhora sensível no RAM Radial foi o conjunto de treinamento, com o conjunto de treinamento igual a 64, o desempenho da RAM Radial foi cerca de 93% de respostas corretas, enquanto o RAM original foi cerca de 89%. Estes resultados são interessantes para a construção de sistemas de reconhecimento de padrões, pois muitos sistemas que são baseados no modelo RAM e o método da n-tupla tem sido utilizado na prática e o modelo proposto comporta-se como uma versão otimizada de tais técnicas..

Referências

- [1] E C B C Filho. *Investigation of Boolean Neural Network Based on a Novel Goal-seeking Neuron*. Doctor's thesis 1990.
- [2] E C B C Filho, M C FairHurst and D L Bisset. *Adaptive Recognition using Goal Seeking Neurons*, Pattern Recognition Letters, 1991, 12, pp.131-138.
- [3] I Aleksander. *Adaptive Systems of Logic Network and Binary Memories*, Proc. Spring Computer Conference. 1967.
- [4] I Aleksander and H. Morton. *An Introduction to Neural Computing*. primeira edição. 1990.
- [5] I Aleksander, T J Stoneham. *A Guide to Pattern Recognition Using Random Access Memories*. IEE Proc on Computer and Digital Techniques, vol.2, pp 29-36. 1979.
- [6] I Aleksander, W V Thomas and P A Bowden. *Wisard - A Radical step forward in image recognition*. Sensor review, 4(3), pp 120-124, 1984.
- [7] T Kohonen and H Ritter. *Self-organizing Semantic Maps*, Journal of Biological Cybernetics, Vol.61, pp. 241-254. 1989.
- [8] R Linsker. *Self-organization in Perceptual Networks*. IEEE Journal of computer, Vol.61. number 3, pp.105-117, 1988
- [9] W.W. Bledsoe and I. Browning. *Pattern recognition by machine*. The Eastern joint computer conference. pp. 225-232. 1959.