

# UM MODELO HÍBRIDO BASEADO EM OTIMIZAÇÃO POR COLÔNIAS DE FORMIGAS E REDES NEURAIAS PARA PREVISÃO DE VAZÕES

*Mêuser Valença<sup>1</sup>; Ivna Valença<sup>2</sup>*

**RESUMO** --- A determinação de valores adequados dos parâmetros de uma rede neural é de extrema importância para a obtenção de uma boa capacidade de generalização da mesma. Este artigo apresenta um novo modelo para otimização de parâmetros de uma rede neural construtiva, treinada via *backpropagation* com momento, baseado no *MAX-MIN Ant System*, um algoritmo de Otimização por Colônias de Formigas. Os parâmetros otimizados foram: a taxa de aprendizagem e a constante de momento. A rede neural foi modelada para previsão de vazões. Os resultados foram comparados com os obtidos através de um processo tradicional de busca exaustiva e se mostraram satisfatórios.

**ABSTRACT** --- The determination of appropriate values of neural networks parameters is extremely important to reach a good capacity of generalization. This paper presents a novel MAX-MIN Ant System-based model to optimization of Neural Networks parameters in a Constructive Neural Network trained with the back propagation algorithm. The optimized parameters include learning rate and momentum. The Neural Network was modeled to inflow forecasting. Results were compared with traditional process of exhaustive search, and have shown promising.

**Palavras-chave:** previsão de vazões, redes neurais, otimização por colônia de formigas.

---

1) Professor da Universidade de Pernambuco (UPE) – Departamento de Sistemas e Computação. Recife/PE. E-mail: meuserv@yahoo.com.br.

2) Mestranda em Ciências da Computação da Universidade Federal de Pernambuco. Recife/PE. E-mail: [icbv@cin.ufpe.br](mailto:icbv@cin.ufpe.br).

# 1 - INTRODUÇÃO

A utilização de técnicas de inteligência computacional na área de recursos hídricos e meio ambiente teve um grande avanço nos últimos anos. Uma das técnicas cuja utilização está se tornando bastante corriqueira corresponde a utilização das redes neurais artificiais principalmente para a previsão de vazões. Portanto, em função da grande aplicabilidade das redes neurais artificiais (RNA), tentativas de automação de algumas variáveis vem sendo realizadas. Dentre essas variáveis, destaca-se a dificuldade em encontrar a arquitetura ideal para a rede bem como os parâmetros ideais a serem utilizados no algoritmo de treinamento. Em alguns modelos, como é o caso da rede *Multilayer Perceptron* (MLP) treinada com o algoritmo *backpropagation*, que é um dos modelos de rede mais largamente utilizado faz-se necessário determinar no mínimo três variáveis ver, por exemplo, Valença (2009). Neste modelo de rede neural a arquitetura da rede é definida pelas variáveis de entrada e saída e pelo estabelecimento do número de neurônios na(s) camada(s) escondida(s) que de maneira geral é feito através de um processo de busca em geral por tentativa e erro. Além do número de neurônios na camada escondida para o treinamento com o algoritmo *backpropagation* dois outros parâmetros precisam ser definidos, a taxa de aprendizagem e a constante de momento. Este trabalho apresenta uma nova abordagem de otimização dos parâmetros do algoritmo *backpropagation* para uma rede neural construtiva chamada NSRBN, Valença (2005), chamada ACONN (*Ant Colony Optimization* para Redes Neurais), a qual é baseada no *MAX-MIN Ant System*, um algoritmo de Otimização por Colônia de Formigas (ACO). A rede neural construtiva, NSRBN, aqui utilizada tem como vantagem definir de forma automática a quantidade ótima de neurônios da camada escondida, de tal forma que nos resta para treinamento à definição dos dois parâmetros do algoritmo de treinamento.

Esta nova abordagem foi aplicada na previsão de vazões médias diárias com o horizonte de sete dias à frente. O restante deste trabalho é organizado como segue. A seção Metodologia introduz respectivamente a definição do *MAX-MIN Ant System*, a apresentação dos parâmetros a serem otimizados e por fim a abordagem proposta. Na seção Estudo de Caso a metodologia é aplicada a previsão de vazões para três usinas hidroelétricas e finalmente na seção Discussão e Conclusões analisamos os resultados obtidos e apresentamos nossas conclusões.

## 2 – METODOLOGIA

O modelo proposto neste trabalho é baseado na abordagem Otimização por Colônias de Formigas. Os algoritmos baseados no comportamento das formigas reais foram inicialmente propostos por Dorigo e colegas, Dorigo *et al* (1991) e também como parte da tese de doutorado de

Marco Dorigo, Dorigo (1992) como um método para resolver problemas complexos de otimização combinatorial (exemplo: O problema do caixeiro viajante).

Desta forma para se aplicar os algoritmos ACO é necessário modelar o problema a ser trabalhado como um problema de busca por menor caminho em grafos. Isto é necessário, pois o ACO é baseado em um processo no qual um conjunto de formigas buscam o menor caminho entre o ninho e a fonte de alimento. Portanto, o problema de determinação dos parâmetros do algoritmo *backpropagation* para treinamento de uma rede neural pode ser visto como um problema de otimização de rotas, no qual os parâmetros são adaptativamente ajustados em relação ao desempenho da rede neural com relação com relação as métricas utilizadas para o problema a ser resolvido.

Na literatura encontram-se propostos vários algoritmos baseados em colônias de formigas que foram desenvolvidos principalmente com o objetivo de resolver problemas de otimização combinatorial complexos. Neste artigo nós apresentaremos o *Ant System* (AS) que foi o primeiro algoritmo baseado em colônia de formigas proposto na literatura e em seguida o *MAX-MIN Ant System* que será utilizado neste trabalho.

## 2.1 – *Ant System* (AS)

O *Ant System* foi o primeiro algoritmo a ser proposto na literatura, Dorigo (1992); Dorigo *et al* (1991); Dorigo *et al* (1996). Uma vez que estes algoritmos foram desenvolvidos para resolver problemas de otimização combinatorial e testados para o clássico problema do caixeiro viajante a sua descrição e detalhamento de implementação serão baseados neste clássico problema. Este algoritmo utiliza  $m$  formigas que são inicialmente dispostas aleatoriamente entre as cidades (que são os nós do grafo) e em cada passo estas vão se deslocando de cidade a cidade, utilizando uma regra de decisão (os trechos entre cidades são os arcos do grafo) até que após  $n$  passos que chamaremos de uma iteração elas construam suas rotas (*tour*). A principal característica deste algoritmo é que o ajuste do feromônio depositado é realizado após todas as formigas terem realizado suas rotas. Desta forma após as formigas terem realizado suas rotas (*tour*) faz-se primeiramente a evaporação do feromônio e posteriormente seu depósito de acordo com a seguinte expressão:

$$\tau_{ij} = (1 - \rho)\tau_{ij} + \sum_{k=1}^m \Delta\tau_{ij}^k \quad (1)$$

onde:  $\tau_{ij}$  é o feromônio ajustado para o trecho que liga as cidades i e j (que são os nós do grafo),  $\rho$  representa a taxa de evaporação, m é o número de formigas e  $\Delta\tau_{ij}^k$  é a quantidade de feromônio depositada no trecho (i,j) pela formiga k.

A quantidade de feromônio  $\Delta\tau_{ij}^k$  é calculada por  $\Delta\tau_{ij}^k = 1/L_k$  se o trecho (i,j) é percorrido pela formiga k ou  $\Delta\tau_{ij}^k = 0$  se a formiga k não andou por este trecho. O termo  $L_k$  é o comprimento da rota (*tour*) realizado pela formiga k.

As formigas artificiais ao se deslocarem enquanto constroem suas rotas (tours) têm que tomar uma decisão em cada local para definir o próximo local a ser visitado o que é feito através de uma regra probabilística em cada local.

A regra probabilística  $p_{ij}^k$  utilizada por este algoritmo para definir o local a ser visitado pela formiga k que se encontra no local i é dada por:

$$p_{ij}^k = \frac{\tau_{ij}^\alpha \cdot \eta_{ij}^\beta}{\sum_{l \in N_i^k} \tau_{il}^\alpha \cdot \eta_{il}^\beta} \quad (2)$$

onde:  $N_i^k$  é o conjunto de cidades vizinhas a i que ainda não foram visitadas pela formiga k,  $\tau_{ij}$  é o feromônio depositado no trecho (i,j) na iteração t,  $\eta_{ij}$  é um valor heurístico chamado de visibilidade e no caso do caixeiro viajante é  $\eta_{ij} = 1/d_{ij}$ ,  $d_{ij}$  é o comprimento do trecho (i,j) e  $\alpha$  e  $\beta$  são os parâmetros que controlam a importância do feromônio depositado versus a informação heurística (visibilidade).

## 2.2 – Max-Min Ant System (MMAS)

O algoritmo *MAX-MIN Ant System* é um sistema derivado do algoritmo original Ant System pela introdução de 2 mudanças. Ele foi proposto por *Stützle and Hoos* (2000) que introduziram as seguintes modificações:

- O ajuste de depósito de feromônio é realizado apenas nos trechos da rota (tour) que é realizada pela formiga que apresenta o melhor desempenho;
- Os valores de depósito de feromônio em cada trecho são limitados dentro de um intervalo que possui um valor mínimo e um valor máximo.

Logo, a função de ajuste do feromônio depositado passa a ser:

$$\tau_{ij} = (1 - \rho)\tau_{ij} + \Delta\tau_{ij}^{\text{ótimo}} \quad (3)$$

onde:  $\Delta\tau_{ij}^{\text{ótimo}}$  é o valor do ajuste de feromônio a ser realizado apenas nos trechos percorridos pela formiga que apresenta o melhor desempenho (distância mínima), de tal forma que seu valor é  $\Delta\tau_{ij}^{\text{ótimo}} = 1/L_{\text{ótimo}}$  e  $L_{\text{ótimo}}$  é o comprimento total do percurso realizado pela formiga de melhor desempenho. Os trechos não percorridos pela formiga que apresenta o melhor desempenho não têm depósito de feromônio, isto é,  $\Delta\tau_{ij}^{\text{ótimo}} = 0$ .

O valor  $L_{\text{ótimo}}$  pode estar relacionado ao melhor valor encontrado na iteração atual ou ao melhor valor encontrado entre todas as iterações (a melhor solução desde o início do algoritmo até a iteração atual. Algumas implementações também foram feitas utilizando-se uma combinação, ou seja, durante algumas iterações utiliza-se o melhor valor encontrado em cada iteração (melhor valor local) e durante outro conjunto de iteração o melhor valor dentre todas as iterações (melhor valor global).

Quanto aos valores máximos e mínimos de feromônio apesar de Stützle and Hoos (2000) apresentarem em seu trabalho expressões analíticas para estabelecimento destes valores, eles sugerem que a depender do problema que se está solucionando a escolha de forma experimental é ainda a mais eficiente.

## 2.2 – Modelo proposto ACONN (*Ant Colony Optimization* para Redes Neurais)

O nosso problema consiste em determinar por otimização os dois parâmetros: taxa de aprendizagem ( $\alpha$ ) e momento ( $\beta$ ). A taxa de aprendizagem é responsável pela velocidade do aprendizado, ou seja, a taxa de mudança dos pesos de tal forma que valores muito baixos exige um maior tempo computacional além de aumentar a possibilidade de o algoritmo ficar preso a mínimos locais. Por outro lado, valores grandes podem provocar perda de capacidade de generalização e instabilidade numérica. Os possíveis valores deste parâmetro estão no intervalo  $[0,1]$ . O termo momento consiste em um parâmetro adicionado ao algoritmo *backpropagation* de forma empírica com o objetivo de acelerar o processo de convergência sem aumentar a instabilidade numérica de forma a permitir que o algoritmo tenha mais chances de escapar de mínimos locais. O valor deste parâmetro também se encontra no intervalo  $[0,1]$ .

Logo, para que possamos aplicar o algoritmo baseado em colônia de formigas precisamos modelar nosso problema como um problema de encontrar menores caminhos em grafos. Neste contexto, foi criado um grafo a partir possíveis valores de cada parâmetro do algoritmo *backpropagation*, como mostra a figura 1.

Inicialmente todas as formigas são colocadas no ponto A e a quantidade de feromônio em cada trilha é considerada como o valor máximo, ou seja,  $\tau_{ij} = \tau_{ij}^{Máx}$ . As formigas então iram se deslocar entre os pontos A, B e C da figura 1.

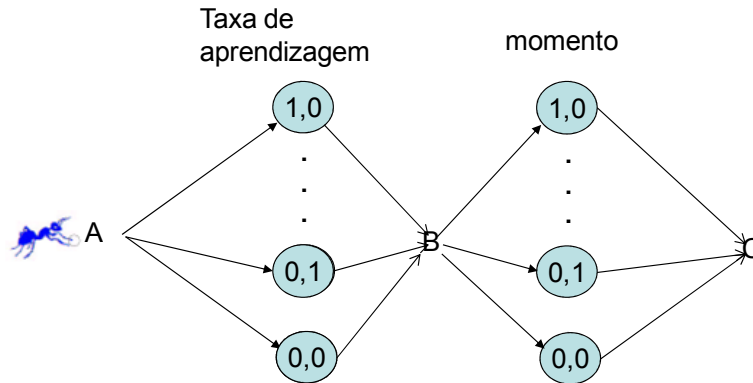


Figura 1 – Rota percorrida pelas formigas

O cálculo da probabilidade de uma formiga dada formiga  $k$  utilizar uma determinada trilha entre os estados  $i$  e  $j$  é dado pela Equação (2), sendo que não é utilizado o termo visibilidade, uma vez que não há o conceito de comprimento da trilha.

A atualização da quantidade de feromônio depositado na trilha entre os estados  $i$  e  $j$  é realizada de acordo com a Equação (3) onde a quantidade de feromônio a ser depositada  $\Delta\tau_{ij}^{ótimo}$  está relacionada com uma dada métrica obtida após treinamento com validação cruzada da rede neural. Neste artigo se utiliza o coeficiente de determinação, de tal forma que

$$\Delta\tau_{ij}^{ótimo} = r^2 \quad (4)$$

A figura 2 mostra uma possível rota realizada por uma formiga  $k$ . Nesta figura, a configuração da rede é  $\alpha = 0,1$  e  $\beta = 0,1$ , que são respectivamente os valores da taxa de aprendizagem e momento, que serão utilizados pela rede neural a fim de verificar seu desempenho obtido com tais parâmetros. Depois de realizado o cálculo, isto é, após todas as formigas terem concluído seu percurso, elas são novamente colocadas no ponto inicial A, de tal forma que este processo é repetido até que todas as formigas utilizem um mesmo caminho, ou seja, o sistema chegue para a solução ótima ou próxima dela.

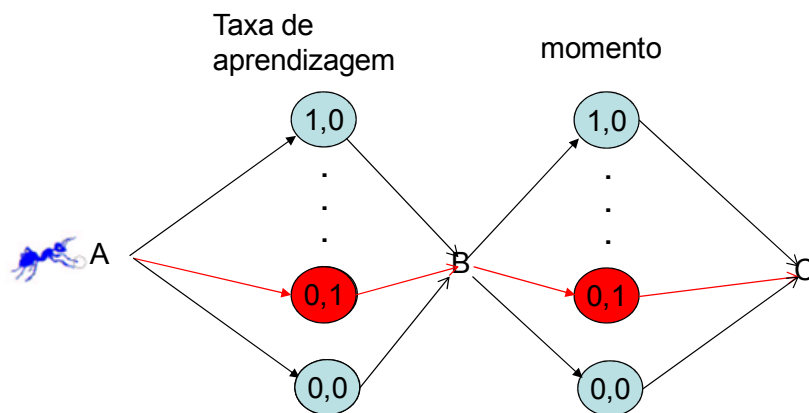


Figura 2 – Rota percorrida pela formiga k com  $\alpha = 0,1$  e  $\beta = 0,1$

### 3 – ESTUDO DE CASO

Para ilustrar o trabalho sobre o uso da técnica aqui proposta se utiliza aqui os dados das séries de vazões médias diárias de três barragens: Tucuruí (Norte), Três Marias (Nordeste), e Foz do Areia (Sul), situadas no Brasil.

O reservatório de Tucuruí esta situado na bacia hidrográfica do rio Tocantins que abrange as superfícies dos estados de Goiás, Tocantins, Mato Grosso, Pará, Maranhão e Distrito Federal. A área de drenagem da bacia totaliza 757.577 km<sup>2</sup> até a Usina Hidroelétrica Tucuruí, dos quais 384.460 km<sup>2</sup> pertencem à bacia do rio Araguaia. A barragem de Três Marias fica no rio São Francisco, que tem uma extensão de 2.700 km, possui uma área de drenagem de, aproximadamente, 631.000 km<sup>2</sup> até a foz, que representa 7,5% do território nacional; estando 83% da área nos estados de Minas Gerais e Bahia, 16% nos estados de Pernambuco, Alagoas e Sergipe, e o restante 1% no estado de Goiás e Distrito Federal. A barragem de Foz do Areia esta localizada no rio Iguaçu. O rio Iguaçu é o principal rio do estado do Paraná. A sua bacia drena uma área de 56.977 Km<sup>2</sup> até o aproveitamento de Salto Caxias.

O horizonte semanal foi adotado uma vez que ele é utilizado para o planejamento do setor elétrico nacional. As séries de dados de vazões diárias foram divididas em três conjuntos de dados: um conjunto para ajuste dos pesos (50%), um conjunto para validação cruzada (25%) e um terceiro conjunto para avaliar o desempenho da metodologia proposta. O histórico de vazões médias diárias aqui utilizadas corresponde ao período de 1968 a 2004. As redes para Tucuruí, Foz do Areia e Três Marias tiveram 14 neurônios de entrada (14 dias anteriores) e 7 na camada de saída.

A comparação que será utilizada consiste em comparar os resultados obtidos quando se determina os parâmetros do algoritmo *backpropagation* através da busca exaustiva (tentativa e erro) com aqueles obtidos de forma automática com a metodologia proposta. A métrica para comparação

utilizada foi o erro percentual médio absoluto dado por:

$$EPMA = \frac{1}{n} \sum_{i=1}^n \frac{(Q_{o,i} - Q_{s,i})}{Q_{o,i}} * 100 \quad (5)$$

O quadro 1 mostra os resultados obtidos para as três usinas hidroelétricas consideradas neste trabalho. Neste quadro 1 se pode verificar que a metodologias proposta fornece resultados até melhores do que aqueles obtidos através da tentativa e erro. O processo de tentativa e erro consistiu em rodar pelo menos 30 simulações com variação dos parâmetros taxa de aprendizagem e momento.

Quadro 1 – Resumo dos resultados obtidos

		Três Marias	Tucuruí	Foz do Areia
Busca exaustiva	Média	21,5%	8,8%	25,4%
	D. Padrão	18,2%	8,1%	9,2%
Modelo proposto	Média	20,8%	8,6%	20,1%
	D. Padrão	11,1%	7,2%	6,7%

## 4 – DISCUSSÃO E CONCLUSÕES

De maneira geral o procedimento utilizado para escolha de valores ótimos dos parâmetros para o algoritmo *backpropagation* utilizado para treinamento de uma rede neural é obtido por tentativa e erro, o que pode ser muito enfadonho e dispendioso quando o usuário não tem grande familiaridade com a utilização da técnica de redes neurais.

Em função disso, este artigo propõe um novo modelo, denominado ACONN, que é capaz de selecionar os valores destes parâmetros de forma automatizada, com intuito de melhorar a eficiência da previsão de vazões. Os resultados mostraram que o modelo proposto, ACONN, foi capaz de encontrar valores ideais para os parâmetros, taxa de aprendizagem e momento de uma rede neural treinada via *backpropagation*. O ACONN obteve resultados tão bons e até melhores, em relação aqueles obtidos por meio de um processo exaustivo de busca por tentativa e erro.

## BIBLIOGRAFIA

- DORIGO, M.; MANIEZZO, V.; COLORNI, A. (1991). "Positive feedback as a search strategy," *Technical Report n. 91-016*, Politecnico di Milano, 1991.
- DORIGO, M.; MANIEZZO, V.; COLORNI, A. (1996). Ant System: Optimization by a colony of cooperating agents. *IEEE Transactions on System, Man, and Cybernetics – Part B*, 26(1):29-41.



- DORIGO, M. (1992). *Optimization, Learning and Natural Algorithms*, Ph.D. Thesis, Dip. Elettronica e Informazione, Politecnico di Milano, Italy.
- DORIGO, M. ; STÜTZLE, T. (2004). *Ant Colony Optimization*. MIT Press, Cambridge, MA.
- STÜTZLE, T. ; HOOS, H., H. (2000). MAX-MIN Ant System. *Future Generation Computer Systems*, 16(8):889–914.
- VALENÇA, M.J.S. (2005). *Aplicando Redes Neurais: um guia completo*. Livro Rápido, Olinda – PE, 264 p.
- VALENÇA, M.J.S. (2007). *Fundamentos das Redes Neurais: exemplos em java*. Livro Rápido, Olinda – PE, 382p.