

MODELAGEM DE COMPORTAMENTO DINÂMICO DE UM SISTEMA TÉRMICO BASEADA EM REDE NEURAL COM TREINAMENTO USANDO FILTRO DE KALMAN EVOLUTIVO

Leandro dos Santos Coelho

Laboratório de Automação e Sistemas, LAS

Programa de Graduação em Engenharia de Produção e Sistemas, PPGEPS

Pontifícia Universidade Católica do Paraná – PUCPR

Rua Imaculada Conceição, 1155, CEP 80215-901, Curitiba, PR, Brasil

leandro.coelho@pucpr.br

Resumo. As redes neurais função de base radial (RNs-RBF) são modelos de aproximação de funções que podem ser treinados por exemplos para implementar um mapeamento (interpolação) de entrada e saída desejado. O desempenho das RNs-RBF depende do número e centros das funções radiais de base, suas formas, e o método utilizado para aprender o mapeamento de entrada-saída. Este artigo apresenta uma abordagem de RN-RBF usando funções Gaussianas e treinamento com filtro de Kalman baseado em otimização através de evolução diferencial. Os resultados de simulação indicam as potencialidades desta abordagem de aprendizado de RN-RBF para identificação de curto prazo (um passo à frente) da densidade de fluxo de calor (saída do sistema térmico) através de duas camadas de parede (tijolo e camada de isolamento). As entradas do sistema são a temperatura interna e a temperatura externa da parede.

Palavras chave: sistema térmico, identificação de sistemas, rede neural, filtro de Kalman, evolução diferencial.

1. Introdução

A identificação de sistemas é o estudo de procedimentos que permitem desenvolver modelos matemáticos a partir de dados e sinais observados (Aguirre, 2000). A área de identificação de sistemas está consolidada para modelos lineares. Os modelos lineares desempenham importante papel no desenvolvimento das técnicas de identificação. Devido à sua simplicidade e à facilidade de obtenção, os modelos lineares são utilizados para aproximar o comportamento de sistemas em determinadas faixas de operação. Os modelos lineares desempenham um importante papel na modelagem de sistemas dinâmicos, quando é possível limitar a operação a regiões de linearidade. Na realidade, todos os sistemas físicos são não-lineares em alguma região de operação. Assim, a presença de comportamento não-linear conduziu ao desenvolvimento de ferramentas que consigam modelar e analisar este comportamento.

Neste contexto, a descrição matemática de sistemas dinâmicos nem sempre é uma tarefa simples na qual princípios básicos podem ser aplicados. Para sistemas complexos, modelagens baseadas em leis elementares para determinar o seu comportamento dinâmico nem sempre são possíveis. Uma alternativa interessante para solucionar este problema é a abordagem de identificação de sistemas. Por esta abordagem, é possível determinar um modelo matemático baseado nas entradas e saídas do sistema. A identificação de sistema é um procedimento que determina um modelo de um sistema dinâmico, com a intenção de prever e/ou compreender seu comportamento dinâmico. Quando existe pouco conhecimento *a priori* do sistema, é comum utilizar-se do modelo do tipo caixa-preta. Este modelo, que é uma estrutura flexível, podendo ser utilizado para aproximar diferentes sistemas não-lineares (Guerra, 2006).

As redes neurais artificiais provaram ser eficientes estruturas não-lineares do tipo caixa-preta em muitas aplicações (Haykin, 1996), e podem ser treinadas usando os dados de entrada e saída observados do próprio sistema. A principal potencialidade na estrutura das redes neurais consiste na sua habilidade de adaptação e aprendizagem. A habilidade de adaptação e aprendizagem pelo ambiente significa que modelos de redes neurais podem lidar com dados imprecisos e situações complexas. Uma rede neural treinada adequadamente tem a habilidade de generalizar quando são apresentadas as entradas que não estão presentes em dados já conhecidos por ela.

Um tipo de rede neural que pode ser aplicado à identificação de sistemas é a rede neural de base radial (RN-RBF, *Radial basis function neural network*). Esta rede compõe uma classe de redes neurais cuja camada oculta é definida por um conjunto de funções de base radial, das quais esta deriva seu nome. As unidades ocultas fornecem um conjunto de funções não-lineares que constituem uma base arbitrária para os padrões de entrada, quando eles são expandidos sobre o espaço oculto. Essas funções são chamadas de funções de base radial. Na estrutura da rede neural de base radial, aplicam-se os algoritmos de agrupamento para encontrar os centros de cada função de base radial, com a intenção de descobrir padrões no espaço de entrada.

Grande parte do interesse pelas RNs-RBF é proveniente do fato delas possuírem fundamentação estatística e são consideradas aproximadores universais de funções com dada precisão (Chen *et al.*, 1991), realizando um mapeamento das entradas de rede de dimensão vetorial m , em uma saída de dimensão vetorial n de qualquer função vetorial contínua.

As RNs-RBF apresentam potencialidades de generalização local, em que a interferência no aprendizado é minimizada e o aprendizado é relativamente rápido, devido ao número mínimo de pesos ajustados a cada vetor de treinamento apresentado à rede neural. Na literatura, diferentes abordagens de treinamento de redes neurais usando filtro de Kalman têm sido propostas, tais como Puskorius e Feldkamp (1994), Birgmeier (1995), Simon (2002), Xiong e Zhou (2003), Bai e Guo (2004) e Zhang e Luh (2005).

Na área de ciências térmicas, a literatura da aplicação de métodos de identificação em sistemas térmicos possui alguns trabalhos relevantes, a citar Cunningham (2001) e Haghghat *et al.* (1998). Entretanto, as potencialidades da utilização de redes neurais é ainda pouco explorada e poucos trabalhos têm sido apresentados, como por exemplo os trabalhos de Chen e Chen (2000), Leung *et al.* (2001), Conesa *et al.* (2004), Torrecilla *et al.* (2004), Turias *et al.* (2005), Aththariyakul e Leephakpreeda (2005), Abassi e Bahar (2005), Kumar *et al.* (2006), Aquino e Brigham (2006) e Parlak *et al.* (2006).

A contribuição deste trabalho é avaliar as potencialidades de usar uma abordagem de treinamento de RNs-RBF baseada em filtro de Kalman estendido (EKF, *Extended Kalman Filter*) com otimização usando evolução diferencial (Storn e Price, 1995; Storn, 1997) para identificação do comportamento dinâmico de um sistema térmico. Neste contexto, o estudo de caso testado tem por objetivo a estimativa da densidade de fluxo de calor através de uma parede de duas camadas (camada de tijolos e isolamento).

O artigo é organizado da seguinte forma. A descrição e a abordagem de treinamento da RN-RBF usando evolução diferencial é detalhada na seção 2. Uma breve descrição do sistema térmico e a análise dos resultados obtidos são detalhadas na seção 3. Finalizando, a conclusão e comentários em relação aos resultados obtidos são apresentados na seção 4.

2. Rede neural de base radial

Em relação ao cérebro humano, tem-se em torno de 10^{11} neurônios, cada um podendo estar conectado a milhares de outros neurônios, o que facilita o estabelecimento de processamentos paralelos de informações e a formação de redes muito complexas. Assim, as redes neurais artificiais tentam reproduzir as funções das redes biológicas, buscando implementar seu comportamento básico e sua dinâmica (Braga *et al.*, 1998).

As redes neurais artificiais consistem em elementos de processamento altamente interconectados denominados neurônios (artificiais), cada qual possuindo um número de entradas e uma saída. A saída de cada neurônio é determinada como uma função não-linear de uma soma ponderada das entradas, embora operações matemáticas complexas possam ser incluídas. Os neurônios se interconectam através de pesos, os quais são ajustados durante o período de treinamento.

As redes neurais são especificadas pela topologia, características dos neurônios e regras de treinamento. Estas regras de treinamento especificam como o conjunto de pesos iniciais deve ser adaptado para a rede aprender o comportamento de um sistema. Os algoritmos de treinamento (ou aprendizado) de uma rede neurais ajustam iterativamente os pesos das conexões entre os neurônios até que os pares desejados de informações de entrada(s) e saída(s) sejam obtidos e as relações de causa e efeito possam ser estabelecidas. Se as condições mudarem de tal modo que o desempenho do modelo não seja mais adequado, pode-se submeter à rede neural a mais um treinamento sob estas novas condições de entrada(s) e saída(s) para corrigir seu desempenho. Assim sendo, podem-se apresentar dados à rede neural a uma atualização periódica, resultando num modelo auto-ajustável em linha.

De uma forma geral, as principais características das redes neurais artificiais são: (i) aprendizado através de exemplos, (ii) adaptabilidade, (iii) capacidade de generalização, (iv) tolerância a falhas e (v) rápida implementação.

A RN-RBF pode ser considerada como um projeto de rede neural para um problema de ajuste de curva (aproximação) em um espaço de alta dimensionalidade. Desta maneira, aprender é equivalente a encontrar uma superfície, em um espaço multidimensional, que forneça o melhor ajuste para os dados de treinamento, com o critério de “melhor ajuste” sendo medido em um sentido estatístico. Então o processo de generalização equivale a utilizar esta superfície multidimensional para interpolar outros pontos que não pertençam aos dados de treinamento, mas estejam em sua vizinhança.

Para uma RN-RBF, as unidades ocultas fornecem um conjunto de “funções” que consistem em uma “base” arbitrária para os padrões (vetores) de entrada, quando eles são expandidos sobre o espaço oculto: estas funções são denominadas de funções de base radial, daí quais a rede deriva seu nome. Cada uma dessas funções é centrada em uma particular coordenada do espaço multidimensional dos pontos que compõem o espaço de dados de entrada. Cada uma destas coordenadas particulares caracteriza-se por definir o centro de uma (entre várias possíveis) região de maior aglomeração de pontos ou grupos (*clusters*), do espaço de dados de entrada (Haykin, 1996).

A arquitetura da RN-RBF, em sua estrutura mais básica, é constituída de três camadas, com propósitos totalmente diferentes:

- *Camada de entrada*: é composta por nós de fonte (unidades sensoriais) que conectam a rede ao seu ambiente;
- *Camada oculta*: é única, aplica uma transformação não-linear do espaço de entrada para o espaço oculto (na maioria das aplicações, o espaço oculto é de alta dimensionalidade);

- *Camada de saída:* é linear, fornecendo a resposta da rede ao padrão (sinal) de ativação aplicada à camada de entrada.

O processamento da entrada de uma rede neural se realiza de acordo com sua topologia e o tipo da ativação dos neurônios. Em uma RN-RBF clássica, este aspecto consiste na alimentação adiante (*feedforward network*) de uma entrada de dimensão m em uma camada oculta, que possui c neurônios ocultos que implementam uma função de ativação baseada em uma medida de similaridade entre os vetores de entrada e os protótipos. Cada neurônio da camada oculta possui um protótipo. As saídas da camada oculta são interligadas às unidades da camada de saída por ramos com pesos sinápticos que ponderam cada saída dos neurônios presentes na camada oculta. Cada unidade da camada de saída implementa uma soma ponderada das saídas da camada oculta. A resposta das camadas de saída forma os elementos do vetor de resposta da RN-RBF.

Uma função de base radial, ϕ , apresenta uma saída simétrica em torno de um centro associado, μ_c , tal que $\phi_c = \phi(\|x - \mu_c\|)$, onde $\|\cdot\|$ é a norma vetorial. Uma condição suficiente para uma função ser candidata à função de base radial é a de ser radialmente simétrica e possuir primeira derivada monotônica. A função Gaussiana é o tipo de função de base radial utilizada neste artigo. As funções Gaussianas são caracterizadas por um parâmetro de escala (ou comprimento), σ , e isto é verdade para muitas outras classes de funções de base radial. Tais classes podem ser representadas como:

$$\phi_c(x) = \phi(\|x - \mu_c\|; \sigma) \quad (1)$$

Um conjunto de funções de base radial pode servir como uma base para representação de uma diversificada classe de funções, que são expressas como combinações lineares de funções de base radial, tal que:

$$y(x) = \sum_{j=1}^M \omega_j \phi(\|x - \mu_c\|) \quad (2)$$

Para identificação de sistemas, as RN-RBFs modelam uma função implícita de um fenômeno físico. O modelo neural de identificação é paramétrico do tipo “caixa-preta”, pois a RN-RBF contém parâmetros livres sem significado físico a serem ajustados para que a rede consiga representar a função implícita do processo.

Os vetores de protótipos e pesos sinápticos são os parâmetros da estrutura da RN-RBF. Deve-se ressaltar que os valores desses parâmetros não acrescentam nada de conhecimento a respeito do sistema, porém devem cumprir o determinado: aproximar a resposta do sistema a uma determinada entrada. Para tanto a estimação dos parâmetros das redes neurais em geral, incluindo as RNs-RBF pode ser concebido como um problema de aprendizado com uma “possível” interpretação biológica e também com equivalência matemática (problema de otimização não-linear na estimação dos parâmetros do modelo segundo um critério de otimalidade) (Haykin, 1996).

Alguns pesquisadores têm treinado as RNs-RBF, escolhendo os centros das unidades ocultas aleatoriamente no conjunto de treinamento. Outros têm utilizado métodos não-supervisionados (por exemplo, agrupamento pelo método de k -médias e fuzzy c -médias), como também algoritmos supervisionados para a seleção dos centros (Simon, 2002). Os métodos de treinamento separam as tarefas de determinação dos protótipos e a otimização dos pesos. A vantagem dessa abordagem é a rapidez no treinamento da rede como resultado da separação das tarefas e aplicação de um método híbrido de treinamento em alternativa ao treinamento global e a otimização não-linear explícita da descida de encosta (incluindo o algoritmo de retropropagação do erro) (Chen *et al.*, 1991).

A separação das tarefas no treinamento da rede é possível devido a estrutura da RN-RBF e seu tipo de função de ativação que age localmente no espaço de entrada. Foi dada uma considerável atenção na literatura das RNs-RBF em como obter um método de treinamento baseado em técnicas lineares explorando características estruturas das RNs-RBF. Observa-se na figura 1 que se os protótipos são fixados e a função da camada escondida $g(\cdot)$ conhecida, a rede é linear em seus pesos sinápticos w . Assim, a estrutura da RN-RBF possibilita a separação de tarefas no processo de aprendizado, com a vantagem que na estimação dos pesos w o treinamento da rede se torna uma tarefa rápida de ser resolvida através de métodos do tipo mínimos quadrados.

Segundo Chen *et al.* (1991), “o problema do mal-condicionamento nas estratégias híbridas pode, por exemplo, ocorrer devido aos centros estarem próximos, resultando em uma dependência linear entre ambos os centros”. Assim o mal-condicionamento numérico da rede pode ser contornado pelos métodos de agrupamento (*clustering*) que faz a escolha dos protótipos como sendo representantes de populações encontradas nos dados de entrada por aprendizado não supervisionado. Observa-se nos métodos que exploram a localidade da RN-RBF que somente as entradas similares aos protótipos serão capazes de contribuir de forma efetiva na resposta da rede e também não usam toda informação fornecida dos exemplos de entrada(s) e saída do conjunto de treinamento na fase de estimação dos parâmetros, acarretando em uma pobre generalização no aprendizado da rede. Neste contexto, o algoritmo do filtro de Kalman é descrito na subseção apresentada a seguir.

2.1. Filtro de Kalman estendido

Criado por Rudolph Emil Kalman em 1960, o filtro de Kalman foi desenvolvido inicialmente como uma solução recursiva para filtragem linear de dados discretos (Kalman, 1960). Para isto, utiliza equações matemáticas que implementam um estimador preditivo de estados, buscando corrigir iterativamente a resposta de um determinado sistema através de múltiplas variáveis relacionadas a ele. Suas áreas de aplicação são muito diversificadas, tais como: processamento de imagem, supervisores de eventos discretos, processamento de sinais, sistemas de inferência, entre outros (Côrrea, 2005).

O filtro de Kalman original (Kalman, 1960; Kalman e Bucy, 1961) foi concebido para ser aplicado em sistemas com dinâmica linear. Para um sistema dinâmico linear com ruído branco de processo de medição, o filtro de Kalman é um estimador ótimo dos estados do sistema. A fim de tornar a aplicação do filtro de Kalman em sistemas não-lineares, pesquisadores propuseram o filtro de Kalman estendido (Kuchner, 1967; Gordon *et al.*, 1993), o que se mostrou eficiente para muitos sistemas cuja não-linearidade é pouco significativa. Para sistemas não-lineares com ruído colorido, o filtro de Kalman pode ser estendido pela linearização da matriz de transição não-linear da descrição em variáveis de estado do sistema em torno dos parâmetros estimados correntes (Haykin, 2001). Neste contexto, o filtro de Kalman estendido atualiza os parâmetros de forma a ser consistente com todos os dados medidos do processo e geralmente converge em poucas interações (Simon, 2002; Welch e Bishop, 1995).

A formulação dos filtros de Kalman está fundamentada na representação em espaço de estados de um sistema dinâmico linear discreto de dimensão finita, ou seja, o filtro de Kalman é descrito em termos de equações a diferenças, porém se o sistema for não-linear, pode-se estender o uso do filtro de Kalman com um procedimento de linearização de um sistema da forma

$$\begin{aligned}\theta_{k+1} &= f(\theta_k) + \omega_k \\ y_k &= h(\theta_k) + v_k\end{aligned}\quad (3)$$

onde o vetor θ_k é o estado do sistema no instante de tempo K , ω_k é o ruído provocado pela própria dinâmica do processo, y_k é o vetor de observação, v_k é o ruído de observação (ou de medida), e $f(\cdot)$ e $h(\cdot)$ são funções vetoriais não-lineares e representam esse sistema possivelmente variante no tempo.

A idéia básica do filtro de Kalman estendido consiste em linearizar a equação (3) a cada instante de tempo em torno do estado estimado recente. Se as não-linearidades são suficientemente suaves, pode-se fazê-lo aplicando a aproximação de Taylor de primeira ordem nas funções não-lineares v e $h(\cdot)$ em torno do estado estimado $\hat{\theta}$ para obter

$$f(\theta_k) = f(\hat{\theta}_k) + F_k(\theta_k - \hat{\theta}_k) + \text{termos de ordem elevada} \quad (4)$$

$$h(\theta_k) = h(\hat{\theta}_k) + H_k^T(\theta_k - \hat{\theta}_k) + \text{termos de ordem elevada} \quad (5)$$

sendo adotada a seguinte notação para as derivadas parciais:

$$F_k = \frac{\partial f(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}_k} \quad (6)$$

$$H_k^T = \frac{\partial h(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}_k} \quad (7)$$

Desconsiderando os termos de ordem mais elevada na equação (5), um sistema linear aproximado é obtido de (4), onde

$$\theta_{k+1} = F(\theta_k) + \phi_k \quad (8)$$

$$y_k = H_k^T(\theta_k) + \varphi_k \quad (9)$$

Neste caso, ϕ_k e φ_k são definidos por

$$\phi_k = f(\hat{\theta}_k) - F_k(\hat{\theta}_k) \quad (10)$$

$$\varphi_k = h(\hat{\theta}_k) - H_k^T(\hat{\theta}_k) \quad (11)$$

A obtenção da aproximação linear do sistema representado pela equação (4) é o primeiro passo para utilização do filtro de Kalman. O problema do filtro de Kalman consiste em resolver simultaneamente e de maneira ótima as equações de processo e de medida para um estado desconhecido.

Pode-se formular o problema de filtragem de um sistema linear achando um valor ótimo de $\hat{\theta}_k$, tal que o sistema deve satisfazer as restrições impostas pelos teoremas do estimador médio condicional e o do princípio da ortogonalidade. Finalmente, pode-se mostrar que a estimação desejada $\hat{\theta}_k$ pode ser obtida recursivamente pelas equações do filtro de Kalman:

$$\hat{\theta}_k = f(\hat{\theta}_{k-1}) + K_k[y_k - h(\hat{\theta}_{k-1})] \quad (12)$$

$$K_k = P_k H_k (R + H_k^T P_k H_k)^{-1} \quad (13)$$

$$P_{k+1} = F_k (P_k - K_k H_k^T P_k) F_k^T + Q \quad (14)$$

onde K_k é o ganho de Kalman e P_k a matriz de covariância do erro de estimação do estado θ_k . As matrizes Q e R são as matrizes de covariância dos ruídos artificiais ω_k e v_k , respectivamente. Como o sistema descrito nas equações (8) e (9) é uma aproximação de um sistema não-linear, o filtro não é ótimo e a estimação é somente uma aproximação da média condicional.

O problema de otimização da matriz de pesos sinápticos W e as coordenadas dos vetores protótipos v_j sob o ponto de vista da estimação ótima de parâmetros com filtro de Kalman de uma RN-RBF pode ser tratado como um problema de minimização do tipo mínimos quadrados ponderados. Neste caso, para aplicar o EKF, faz-se com que os elementos da matriz de pesos W e os elementos dos vetores de protótipos v_j constituam o estado do sistema discreto não-linear em que o filtro de Kalman é aplicado.

2.2. Evolução diferencial

A ED é um paradigma da computação evolutiva (ou evolucionária) desenvolvido por Rainer Storn e Kenneth Price (Storn e Price, 1995; Storn, 1997) para problemas de otimização não-linear contínua. Basicamente, a ED realiza mutações nos vetores pela adição ponderada de diferenças aleatórias entre eles. A escolha da ED para problemas de otimização é baseada nas características de que a ED:

- é um algoritmo de busca estocástica que é motivado pelos mecanismos de seleção natural;
- é menos susceptível a mínimos (ou máximos) locais, pois busca a solução ótima global pela manipulação de uma população de soluções candidatas, ou seja, busca um número de diferentes áreas simultaneamente no espaço de busca;
- não requer informação de derivadas para o cálculo da função de aptidão;
- manipula diretamente os números de ponto flutuante (fenótipo), diferente dos algoritmos genéticos canônicos ou binários que manipulam cadeias de bits (genótipo), o que diminui o custo computacional necessário para transformação de uma representação genotípica em fenotípica;
- geralmente não necessita de populações grandes para funcionar eficientemente.

Na ED clássica, cada variável (indivíduo) é representada por um valor real (ponto flutuante). A variante implementada neste trabalho foi a ED/rand/1/bin.

3. Descrição do sistema térmico e análise dos resultados de identificação usando RN-RBF

A seguir são apresentados os detalhes do sistema térmico a ser identificado e após são mostrados e comentados os resultados de identificação usando RN-RBF com filtro de Kalman e evolução diferencial.

3.1. Descrição do sistema térmico

A obtenção de uma estimativa apropriada para o comportamento dinâmico da densidade de fluxo de calor através de uma parede de duas camadas (camada de tijolos e isolamento) é o objetivo do procedimento de identificação baseado

em modelo neural deste trabalho. As entradas do sistema são a temperatura interna ($u_1(t)$) e a temperatura externa ($u_2(t)$) da parede. O sinal de saída $y(t)$ consiste da densidade de fluxo de calor através de uma parede (Bloem, 1994). Neste caso, os dados (1680 amostras) usados no procedimento de identificação estão disponíveis na base de dados DaISy (De Moor, 2006). É importante mencionar que o sistema térmico, avaliado neste trabalho, é um sistema teste para métodos de identificação. Detalhes do sistema térmico podem ser encontrados em Bloem (1994) e De Moor (2006).

3.2. Análise dos resultados de identificação

O procedimento de análise da identificação através de RN-RBF é dividido nas seguintes etapas: (i) escolha da estrutura da RN-RBF para representar o sistema a ser identificado, (ii) otimização da RN-RBF usando agrupamento de dados e pseudo-inversa; (iii) estimação dos parâmetros do modelo matemático (*fase de estimação*), e (iv) a validação do modelo matemático (*fase de validação*).

Na procura de um modelo “ideal”, devem ser observadas as seguintes questões: se o modelo encontrado é flexível o suficiente e se o modelo é desnecessariamente complexo, ou em outras palavras, se o modelo é confiável para o propósito desejado. Existem diversas maneiras de se realizar a validação de um modelo. Neste trabalho, o índice de desempenho utilizado é o erro médio quadrático (*Mean Squared Error*, MSE).

Para identificação foi escolhido um modelo matemático para a representação do processo. A estrutura de modelo testada é o modelo *NARX (Nonlinear AutoRegressive with eXogenous inputs)* em concepção série-paralela com 2 entradas ($[u_1(t-1); u_2(t-1)]$) e 1 saída estimada, $\hat{y}(t)$, para a densidade de fluxo de calor.

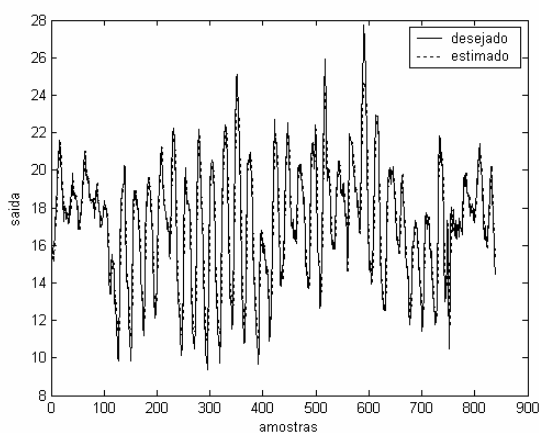
Na tabela 1 é apresentado o resultado para o projeto do filtro de Kalman (executado apenas em uma época de aprendizado) para o treinamento da RN-RBF. O espaço de busca usado na ED foi entre 0,1 e 2000 para as variáveis da diagonal principal de P , Q e R . Neste caso a função de aptidão (*fitness*) usada na ED foi o *MSE*. Foram usadas as amostras de 1 a 840 na fase de estimação da RN-RBF, e as amostras de 841 a 1680 na fase de validação.

Os valores dos termos da diagonal principal das matrizes P , Q e R foram obtidos usando uma abordagem de ED com uma população de 5 e critério de parada de 20 gerações. O melhor resultado obtido pela ED para o projeto do filtro de Kalman usado na RN-RBF foi $P=238,2820$, $Q=125,3487$ e $R=15,2179$ usando uma RN-RBF com 3 funções Gaussianas (grupos) na camada oculta.

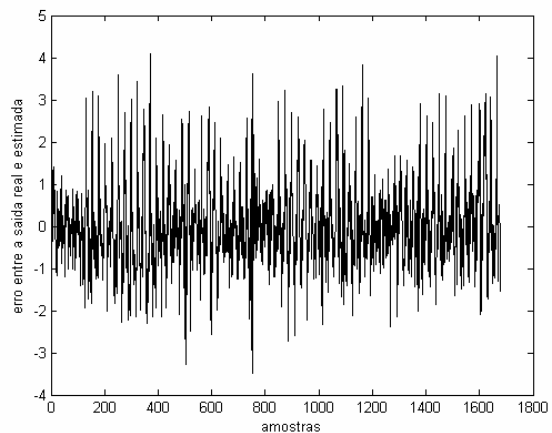
Tabela 1. Resultados das simulações para previsão densidade de fluxo de calor através de uma parede de duas camadas.

Gaussianas na camada oculta (nº de grupos)	MSE ($\times 10^{-3}$) (fase de estimação)	MSE ($\times 10^{-3}$) (fase de validação)
2	0,4139	0,0238
3	0,3983	0,0154
4	0,3993	0,0147
5	0,5110	0,0975

Na figura 1 é apresentado o sinal de erro e saída estimada para a densidade de fluxo de calor através de uma parede de duas camadas, respectivamente, para o melhor resultados apresentado na tabela 1 obtidos pela RN-RBF com 4 funções Gaussianas na camada oculta. Nota-se que a RN-RBF apresentou um índice de desempenho apropriado para modelar o comportamento dinâmico da densidade de fluxo de calor com MSE ($\times 10^{-3}$)= 0,3983 (fase de estimação) e MSE ($\times 10^{-3}$)=0,1054 (fase de validação).



(a) densidade de fluxo de calor



(b) sinal de erro de estimativa

Figura 1. Saída estimada para a densidade de fluxo de calor usando RN-RBF (4 funções Gaussianas na camada oculta).

4. Conclusão e futura pesquisa

As RNs-RBF têm como uma das principais características agirem como se fossem aproximadores universais de funções, podendo ser utilizadas com êxito para identificar sistemas complexos. Este trabalho apresentou uma abordagem de RN-RBF baseada em treinamento por filtro de Kalman híbrido com evolução diferencial. Os resultados de simulação indicam as potencialidades desta abordagem de rede neural para identificação de curto prazo (um passo à frente) da densidade de fluxo de calor em uma parede.

O filtro de Kalman possibilitou lidar com informações ruidosas do sistema dinâmico, utilizando informações matemáticas para estimar o estado atual do ambiente no qual está inserido. Uma dificuldade no projeto do filtro de Kalman foi o projeto das matrizes P , Q e R . Neste contexto, neste trabalho, utilizou-se a técnica de evolução diferencial para otimizar os termos da diagonal principal das matrizes de P , Q e R . Os resultados de empregar filtro de Kalman com evolução diferencial no projeto da RN-RBF obtidos foram satisfatórios em termos de MSE.

Em pesquisa futura, deseja-se realizar uma análise estatística de diferentes abordagens de algoritmos de otimização para o projeto do filtro de Kalman em aplicações de previsão de séries temporais não-lineares de sistemas térmicos com características complexas de atraso de transporte e não-linearidades.

5. Referências

- Abbassi, A., Bahar, L., 2005, "Application of Neural Network for the Modeling and Control of Evaporative Condenser Cooling Load", Applied Thermal Engineering, Vol. 25, No. 17-18, pp. 3176-3186.
- Aquino, W., Brigham, J. C., 2006, "Self-learning Finite Elements for Inverse Estimation of Thermal Constitutive Models", International Journal of Heat and Mass Transfer, Vol. 49, No. 15-16, pp. 2466-2478.
- Aguirre, L. A., 2000, "Introdução à Identificação de Sistemas – Técnicas Lineares e Não-lineares", Editora UFMG, elo Horizonte, MG.
- Atthajariyakul, S., Leephakpreeda, T., 2005, "Neural Computing Thermal Comfort Index for HVAC Systems", Energy Conversion and Management, Vol. 46, No. 15-16, pp. 2553-2565.
- Bai, L.; Guo, P., 2004, "A Novel Approach to Stellar Recognition by Combining EKF and RBF-net", Advances in Neural Networks, International Symposium on Neural Networks, Dalian, China, Lecture Notes in Computer Science, vol. 3173/2004, Springer, pp. 1002-1007, 2004.
- Birgmeier, M., 1995, "A Fully Kalman-trained Radial Basis Function Network for Nonlinear Speech Modeling", Proceedings of IEEE International Conference on Neural Networks, Vol. 1, pp. 259-264.
- Bloem, J. (ed.), 1994, "System Identification Competition, Benchmark Tests for Estimation Methods of Thermal Characteristics of Buildings and Building Components", Joint Research Centre, Ispra, Italy.
- Braga, A. P., Ludermir, T. B., Carvalho, A. C. P. L. F., 1998, Fundamentos de Rede Neurais Artificiais, 11ª Escola de Computação, SBC.
- Chen, S.; Cowan, C.; Grant, P., 1991, "Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks", IEEE Transactions on Neural Networks, Vol. 2, No. 2, pp. 302-309, 1991.
- Chen, Y., Chen, Z., 2000, "A Neural-Network-based Experimental Technique for Determining z-transfer Function Coefficients of a Building Envelope", Building and Environment, Vol. 35, pp. 181-189.
- Conesa, J. A., Caballero, J. A., Reyes-Labarta, J. A., 2004, "Artificial Neural Network for Modelling Thermal Decompositions", Journal of Analytical and Applied Pyrolysis, Vol. 71, No. 1, pp. 343-352.
- Côrrea, L. G. C., 2005, "Inferência da Qualidade de Produtos de Destilação utilizando Redes Neurais Artificiais e Filtro de Kalman Estendido", Dissertação de mestrado, Programa de Pós-Graduação em Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, RJ.
- Cunningham, M. J., 2001, "Inferring Ventilation and Moisture Release Rates from Field Psychrometric Data only System Identification Techniques", Building and Environment, Vol. 36, pp. 129-138.
- De Moor B.L.R. (ed.), 2006, "DaISy: Database for the Identification of Systems", Department of Electrical Engineering, ESAT/SISTA, K.U.Leuven, Belgium, URL: <http://homes.esat.kuleuven.be/~smc/daisy/>, 10/06/2006. [Used dataset: Heat flow density through a two layer wall, section *Thermic Systems*, code 96-011]
- Gordon, N. J., Salmond, D. J., Smith, A. F. M., 1993, "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation", IEE Proceedings-F, vol. 140, pp. 107-113.
- Guerra, F. A., 2006, "Estudo e Análise de Métodos de Agrupamento para o Treinamento de Redes Neurais de Base Radial em Aplicações de Identificação de Sistemas Não-lineares", Dissertação de mestrado, PPGEPS, Pontifícia Universidade Católica do Paraná, Curitiba, PR.
- Haghighat, F., Fazio, P., Zmeureanu, R. A., 1998, "A Systematic Approach for Derivation of Transfer Function Coefficients of Buildings from Experimental Data", Energy and Buildings, Vol. 12, pp. 101-111.
- Haykin, S., 1996, Neural Networks, 2nd edition, Prentice-Hall, Upper Saddle River, New Jersey, 1996.
- Haykin, S., 2001, Kalman Filtering and Neural Networks, John Wiley & Sons, Inc, New York, 2001.
- Kalman, R. E., 1960, "A New Approach to Linear Filtering and Prediction Problems", Transactions of the ASME – Journal of Basic Engineering, Vol. 83, pp. 35-45.

- Krauss, G.K., Kindangen, J. I., Depecker, P., 1997, "Using Artificial Neural Networks to Predict Interior Velocity Coefficients", *Building and Environment*, Vol. 32, No. 4, pp. 295-303.
- Kumar, R., Kaushik, S. C., Garg, S. N., 2006, "Heating and Cooling Potential of an Earth-to-air Heat Exchanger Using Artificial Neural Network", *Renewable Energy*, Vol. 31, No. 8, pp. 1139-1155.
- Kuchner, H. J., 1967, "Approximations to Optimal Nonlinear Filters", *IEEE Transactions on Automatic Control*, Vol. 12, No. 5, pp. 546-556.
- Leung, A. W. T., Tam, C. M., Liu, D. K., 2001, "Comparative Study of Artificial Neural Networks and Multiple Regression Analysis for Predicting Hoisting Times of Tower Cranes", *Building and Environment*, Vol. 36, pp. 457-467.
- Parlak, A., Islamoglu, Y., Yasar, H., Egrisogut, A., 2006, "Application of Artificial Neural Network to Predict Specific Fuel Consumption and Exhaust Temperature for a Diesel Engine", *Applied Thermal Engineering*, Vol. 26, No. 8-9, pp. 824-828.
- Puskorius, G.V.; Feldkamp, L.A., 1994, "Neurocontrol of Nonlinear Dynamical Systems with Kalman Filter Trained Recurrent Networks", *IEEE Transactions on Neural Networks*, Vol. 5, No. 2, pp. 279-297, 1994.
- Rössler, O. E., 1976, "An Equation for Continuous Chaos", *Physical Letters*, Vol. 35A, pp. 397-398, 1976.
- Simon, D., 2002, "Training Radial Basis Neural Networks with the Extended Kalman Filter", *Neurocomputing*, Vol. 48, No. 1-4, pp. 455-475, 2002.
- Storn, R., 1997, "Differential Evolution — A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces", *Journal of Global Optimization*, Vol. 11, No. 4, pp. 341-359.
- Storn, R.; Price, K., 1995, "Differential Evolution: A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces", Technical Report TR-95-012, International Computer Science Institute, Berkeley.
- Torrecilla, J. S., Otero, L., Sanz P. D., 2004, "A Neural Network Approach for Thermal/pressure Food Processing", *Journal of Food Engineering*, Vol. 62, No. 1, pp. 89-95.
- Turias, I. J., Gutiérrez, J. M., Galindo, P. L., 2005, "Modelling the Effective Thermal Conductivity of an Unidirectional Composite by the Use of Artificial Neural Networks", *Composites Science and Technology*, Vol. 65, No. 3-4, pp. 609-619.
- Welch, G., Bishop, G. An Introduction to the Kalman Filter, Technical Report TR- 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, NC, USA, 1995.
- Xiong, S.-S.; Zhou, Z.-Y., 2003, "Neural Filtering of Colored Noise based on Kalman Filter Structure", *IEEE Transactions on Instrumentation and Measurement*, Vol. 52, No. 3, pp. 742-747.
- Zhang, L.; Luh, P.B., 2005, "Neural Network-based Market Clearing Price Prediction and Confidence Interval Estimation with an Improved Extended Kalman Filter Method", *IEEE Transactions on Power Systems*, Vol. 20, No. 1, pp. 59-66.

DYNAMIC BEHAVIOR MODELING OF A THERMAL SYSTEM BASED ON NEURAL NETWORKS WITH TRAINING USING EVOLUTIONARY-KALMAN FILTER

Leandro dos Santos Coelho
Laboratório de Automação e Sistemas, LAS
Programa de Graduação em Engenharia de Produção e Sistemas, PPGEPS
Pontifícia Universidade Católica do Paraná – PUCPR
Rua Imaculada Conceição, 1155, CEP 80215-901, Curitiba, PR, Brasil
leandro.coelho@pucpr.br

Abstract

Radial basis function neural networks (RBF-NNs) are function approximation models that can be trained by examples to implement a desired input-output mapping (interpolation). The performance of RBF-NNs depends on the number and centers of the radial basis functions, their shapes, and the method used for learning the input-output mapping. This paper presents a RBF-NN model using Gaussian functions and Kalman filter training based on differential evolution optimization. The simulation results indicate the potentialities of this RBF-NN learning approach to one step ahead identification of heat flow density (output of thermal system) through a two layer wall (brick and insulation layer). The system inputs are the internal and external temperature of the wall.

Keywords: thermal system, systems identification, neural network, Kalman filter, differential evolution