



**SNPTEE
SEMINÁRIO NACIONAL
DE PRODUÇÃO E
TRANSMISSÃO DE
ENERGIA ELÉTRICA**

GMI – 25
16 a 21 Outubro de 2005
Curitiba - Paraná

**GRUPO XII
ASPECTOS TÉCNICOS E GERENCIAIS DE MANUTENÇÃO EM INSTALAÇÕES ELÉTRICAS -GMI**

USO DE SISTEMAS INTELIGENTES NA PRIORIZAÇÃO DA MANUTENÇÃO DE LINHAS DE TRANSMISSÃO

Taciana Nascimento Lima*
Analista de Manutenção

Roberto Ângelo Fernandes Santos
Gerente de Projetos

Irany Pascoal de Macedo
Gestor de Manutenção

RESUMO

O planejamento da manutenção das linhas de transmissão da Celpe era feito apenas com o conhecimento *a priori* dos especialistas. Com o objetivo de sugerir ações para corrigir eventuais deficiências, focar investimentos necessários, dimensionar o impacto em pontos de decisão e analisar as variáveis quanto à priorização da manutenção, foram utilizadas técnicas de Inteligência Artificial com algoritmos *a priori* para indução de regras e redes neurais (MPL) para estimação do risco de falhas a partir dos defeitos encontrados, uma vez que as correções de defeitos encontrados, exceto as emergências, devem ser programadas para datas posteriores.

PALAVRAS-CHAVE

Manutenção, sistemas inteligentes, defeito, falha, priorização.

1. INTRODUÇÃO

Os sistemas de informação estão cada vez mais difundidos nas empresas como ferramenta na gestão da manutenção. Um dos aspectos mais importantes é a disponibilidade do sistema que, por sua vez, depende tanto da tecnologia dos materiais e equipamentos utilizados quanto da adequada coleta de dados.

O presente trabalho foi motivado pelo interesse da Celpe – Companhia Energética de Pernambuco – do Grupo Neo Energia em promover uma consulta aos dados de manutenção de linhas de transmissão de energia elétrica entre 69kV e 138kV e analisá-los, gerando um diagnóstico de falhas. Devem ser priorizados os defeitos encontrados no sistema de acordo com seu risco de falha.

Para análise, interpretação e decisão foram utilizadas técnicas de mineração de dados (data mining), através da ferramenta computacional NeuralScorer Development, desenvolvida pela empresa NeuroTech. Essa ferramenta foi utilizada basicamente para descobrir as principais causas e características dos defeitos através de indução de regras e para priorizar a manutenção dos defeitos através de uma rede neural.

2. CARACTERIZAÇÃO DO PROBLEMA

Para uma análise adequada dos dados coletados pelas inspeções nas linhas de transmissão é necessário conhecer alguns aspectos do planejamento da manutenção de linhas da Celpe.

Com o intuito de diminuir as interrupções do fornecimento de energia devido a falhas intempestivas, através de ações preventivas, o planejamento da manutenção de linhas de transmissão (LT) é constituído de dois planos, baseados em metodologia, critério e periodicidade específica, definidos pelo órgão gestor:

Plano de inspeção de linhas - Inspeções nas linhas de transmissão (LT) por tipo. Inspeções minuciosas, expeditas e termográficas realizadas trianualmente de maneira intercalada e inspeções de patrulhamento realizadas anualmente. Desta forma a LT é percorrida duas vezes pela equipe de inspeção por ano. Este plano está limitado ao homem x hora disponível nas unidades executoras de manutenção.

Plano de correção de defeitos – Manutenções corretivas programadas para substituição de peças e equipamentos defeituosos ou fora do padrão admitido, onde os defeitos nas LT's são identificados através das inspeções do ano anterior, salvo aqueles que são caracterizados como risco eminente ao sistema (emergências) que são corrigidas de imediato. Este plano está limitado aos recursos financeiros anuais disponibilizados para manutenção.

Desta forma, periodicamente são realizadas inspeções nas linhas de transmissão para detecção de defeitos potenciais de falhas, que são catalogados em banco de dados para posterior análise. A partir da análise dos dados de defeitos gerados pela inspeção, é possível fazer uma programação para correção desses defeitos de modo a minimizar os pontos potenciais de falha das linhas de transmissão de energia elétrica.

A questão principal é a dificuldade de análise desses dados a fim de gerar uma priorização dos defeitos de modo a reduzir a probabilidade de falhas com os recursos financeiros disponíveis.

Uma vez que as manutenções realizadas são limitadas ao custeio, é necessário selecionar, baseado na confiabilidade (1), os defeitos nas LT's que causem maior risco de falha ao sistema. Para aplicação dessa metodologia, o analista humano sem a ferramenta adequada, teria dificuldade em analisar uma grande massa de dados de modo a garantir a qualidade das decisões.

A solução é então utilizar técnicas de inteligência computacional de modo a sistematizar o tratamento e uso dos dados para que eles possam gerar subsídios necessários a uma tomada de decisão com um maior nível de certeza. A partir da utilização da mineração dos dados, torna-se possível a extração das informações implícitas e potencialmente úteis para os especialistas (2). Desta forma, é possível identificar os defeitos que causam as falhas nas linhas de transmissão e otimizar as correções dos defeitos mais críticos.

3. MASSA DE DADOS

A massa de dados utilizada consiste de 2122 registros de defeitos encontrados pela manutenção da Celpe, obtidos no ano de 2003. Algumas características das massas foram suprimidas por questões de confidencialidade. Os dados foram extraídos do banco gerado pelo sistema SAP/R3 e por inspeções realizadas por terceiros.

Foi selecionado um conjunto das variáveis mais relevantes, segundo a análise dos especialistas de manutenção da Celpe. Algumas dessas variáveis sofreram transformações para adequação às técnicas utilizadas. Foram divididas em variáveis de entrada (*a priori*) e variáveis de saída (*a posteriori*) ou alvo. São consideradas variáveis de entrada aquelas que estão disponíveis no momento da priorização e de saída os defeitos que levaram à saída da LT, ou seja, os dados de ocorrências de interrupção do fornecimento de energia elétrica de janeiro de 2003 a fevereiro de 2004. Para simplificação da análise consideramos apenas uma variável de saída (alvo) que indica se ocorreu falha (maus) ou não (bons).

Bons são aqueles defeitos cuja relação entre variáveis tende a não gerar falha. *Maus* são aqueles defeitos cuja relação entre variáveis tende a gerar falha.

A relação de variáveis utilizadas na análise do problema estão relacionadas na Tabela 1 a seguir.

TABELA 1- Variáveis Utilizadas na Análise do Problema

TÍTULO	TIPO	DESCRIÇÃO
NOTA-ITM	ENTRADA	SERIAL DO DEFEITO. INDICA SE A INSPEÇÃO FOI EXTRAÍDA DO SAP-R3 OU FEITA POR TERCEIROS
KM LT	ENTRADA	EXTENSÃO DA LT, AGRUPADA EM FAIXAS DE KM
CAPACID(A) LT	ENTRADA	CAPACIDADE DA LT, AGRUPADA EM FAIXAS
TIPO CABO	ENTRADA	TIPO DE CABO, COLOCADA EM FAIXAS DE BITOLA
LT ESPEC	ENTRADA	SE A LT É CONSIDERADA ESPECIAL
MACRO-DEFEITOS	ENTRADA	MACRO DEFEITOS
DEFEITOS	ENTRADA	DEFEITOS
CUSTO MANUT	ENTRADA	CUSTO DA MANUTENÇÃO
CAUSA DA FALHA	ENTRADA	CAUSA DA FALHA PROVENIENTE DO DEFEITO
INSPEÇÃO	ENTRADA	TIPO INSPEÇÃO
DEF FALHOU LT?	ALVO	DEFEITO QUE PROVOCOU FALHA NA LT

Dois processos de natureza e objetivos diferentes foram realizados: a codificação dos campos categóricos em numéricos para compatibilizar com a entrada das redes neurais e a categorização das variáveis numéricas em faixas para compatibilizar com o indutor de regras.

Para se chegar à massa final de registros foi feita uma padronização de atributos, formatos, convenções de nomes, além de remoção de inconsistências. As linhas foram caracterizadas pelos especialistas de manutenção como especiais ou não, baseados nas regras e consensos existentes na Celpe. Os defeitos foram agrupados por faixa de acordo com sua importância. Foram eliminados os registros que continham informações incompletas ou de preenchimento errôneo. Para cada registro de defeito, foram estimados os custos individuais com material e mão-de-obra.

A variável *KM LT* foi agrupada em faixas de distância em quilômetros de forma a agrupar as LT's em dificuldade de inspeção de acordo com os parâmetros de homem x hora utilizados na Celpe.

Algumas variáveis tais como as que indicam a codificação das linhas de transmissão e seus circuitos correspondentes não foram utilizadas, pois deixaria os dados tendenciosos e facilitariam sua identificação pelas redes neurais. Essas variáveis, contudo, podem ser usadas no futuro quando tivermos defeitos de vários anos, uma vez que a massa de dados foi colhida apenas através das inspeções de 2003. Por esse motivo, algumas variáveis foram categorizadas de maneira a não identificar diretamente a LT.

O processamento dos dados resultou em 2 (duas) massas de dados, uma para aplicação na rede neural e outra extração de regras através do algoritmo *a priori*.

4. TÉCNICAS UTILIZADAS

4.1. Algoritmo *a priori* e indução de regras

Diversas técnicas de indução de regras poderiam ter sido utilizadas neste trabalho. Utilizamos uma variação do algoritmo *a priori* (3) para indução, conjugada com a técnica de refinamento de Major e Mangano (4), que é realizada em 3 etapas.

Consiste em selecionar os subconjuntos mais frequentes, a partir dos de mais baixa ordem. Desta forma são:

- 1º Induzidas regras que satisfazem a uma cobertura mínima e que têm uma complexidade limitada;
- 2º Selecionadas regras que atendem a níveis de cobertura mais amplos que os induzidos originalmente e a níveis de confiança de interesse;
- 3º Selecionadas as regras que apresentam mais interesse sob os pontos de vista do especialista e da eliminação de redundância nas mesmas.

4.2. Redes neurais e estimação de falha

Para fazer a avaliação do risco de falha, optamos pelo modelo de rede neural Multilayer perceptron (MLP) treinada com o algoritmo de backpropagation (retropropagação) (5, 6), pois esse modelo tem sido bastante

utilizado com sucesso em aplicações de problemas de classificação de padrões. Dentre as características mais atrativas deste tipo de rede neural é possível destacar a excelente capacidade de generalização, a simplicidade de operação da rede e o fato de a mesma ser um aproximador universal de funções (7).

As redes neurais artificiais apresentam um modelo inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência. A rede MLP é formada por conjuntos de unidades de processamento, usualmente chamados de neurônios artificiais, agrupadas em camadas de entrada, intermediárias (ou escondidas) e de saída.

Em uma rede MLP típica estas camadas são dispostas em seqüência, uma após a outra. Os neurônios de uma camada são completamente conectados aos neurônios da camada antecessora. A cada conexão entre os neurônios é associado um peso. São os valores destes pesos aplicados aos padrões de entrada da rede neural que determinam a magnitude dos estímulos apresentados a cada neurônio da rede. Estes estímulos são usados como argumentos para as funções de ativação contidas nos neurônios artificiais. Os resultados da aplicação destas funções de ativação são passados adiante através das camadas seguintes até que, na última camada, o resultado final da rede neural é produzido.

Normalmente vários ciclos são necessários até que seja possível aos pesos da rede MLP modelar adequadamente os dados de treinamento fornecidos para o problema.

4.3. Treinamento da rede neural

Uma propriedade importante das redes neurais é a habilidade de aprender e esse aprendizado é feito através de treinamento, onde há o ajuste dos pesos das conexões entre neurônios.

O treinamento de uma rede neural com backpropagation possui dois passos distintos que formam um ciclo de treinamento. No primeiro passo, um padrão de entrada é apresentado para a camada de entrada e as unidades de processamento computam, camada por camada, os resultados das funções de ativação até que um conjunto de estímulos é finalmente fornecido à camada de saída que o utiliza para determinar a saída final da rede neural. O segundo passo do treinamento começa com o cálculo do erro observado entre os valores de saída da rede neural e os valores desejados de saída (inerentes ao problema a ser resolvido). Os erros são então propagados retroativamente a partir da camada de saída até a camada de entrada e os pesos das conexões entre as camadas são ajustados utilizando uma taxa de aprendizagem em função do gradiente da função objetivo. Este processo ocorre de tal forma que é minimizada a diferença entre as saídas calculadas pela rede neural e as saídas almejadas.

As redes neurais utilizadas nos experimentos deste trabalho foram redes MLP com 35 neurônios na camada de entrada, 3 na camada intermediária e 1 neurônio na camada de saída (1 para defeito sem falha e 0 para falha). Todas as redes neurais foram treinadas com o algoritmo de retropropagação de erro padrão com taxa de aprendizagem igual a 0,01 e momentum 0,1. O critério adotado para parada do treinamento foi o monitoramento do poder de generalização da rede neural utilizando a validação cruzada (cross validation) (8) com MSE (Mean Square Error) como métrica.

Todos os dados numéricos foram normalizados ou codificados de maneira que assumissem valores entre 0 e 1 (intervalo de valores necessários para entrada na rede). Essa massa de dados codificada foi separada em 2 lotes para uso da técnica de redes neurais, um com 75% dos casos para treinamento e validação da rede neural e outro de 25% para teste. O treinamento deve ser interrompido quando for encontrado o ponto ótimo de parada com erro mínimo e capacidade de generalização máxima, definidos pelo especialista.

Para o lote de treinamento houve uma replicação dos registros de maneira a igualar a quantidade de defeitos com ou sem falha. Os lotes de teste e validação foram mantidos com as suas características originais.

5. AVALIAÇÃO DE DESEMPENHO

Com os dados tratados e analisados foi possível sugerir ações para corrigir eventuais deficiências, focar investimentos necessários, dimensionar o impacto em pontos de decisão e analisar as variáveis para tomada de decisão quanto à priorização da manutenção em linhas de transmissão.

Para efeito de análise é importante saber a distribuição das falhas na massa de dados para medir o efeito da aplicação do score no escalonamento e avaliar a qualidade das regras geradas. Observa-se que 4,24% dos dados são maus e 95,76% são bons, conforme representado na Figura 1 a seguir.

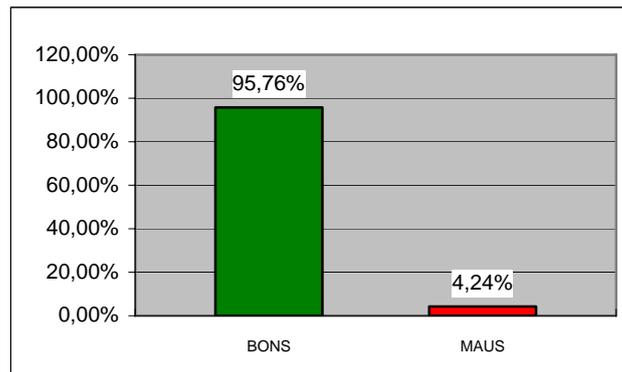


FIGURA 1 - Histograma da situação dos dados

5.1. Análise pelo Score

A transformação da saída da rede MLP em uma resposta contínua para o problema binário da decisão (score) demanda um limiar (ponto de corte) para definir a separação das classes. Para a análise dos dados de manutenção foi utilizada a metodologia baseada em escores. É definida uma pontuação de 0 a 100 de modo que as variáveis de entrada converjam através de um algoritmo decisor para uma variável de saída unidimensional (score).

Utilizando o conjunto de teste e o score gerado a partir dele (com o risco de ocorrer uma falha) foi escolhido um limiar para decidir quais os defeitos que seriam primeiro consertados. Considerando que a priorização inicial fosse dada para aproximadamente 10% mais críticos (com risco maior) determinou-se um ponto de corte em pouco mais do score 18. Com esse limiar, apenas 0,2% falhas do conjunto de teste não teriam sido cobertas, Como mostra a Tabela 2 abaixo.

TABELA 2 - Matriz confusão em percentual de defeitos

	FALHA PREVISTA (%)	NORMAL PREVISTA (%)	TOTAL (%)
FALHA REAL (%)	4	6,4	10,4
NORMAL REAL (%)	0,2	89,4	89,6
TOTAL (%)	4,2	95,8	100

Através do gráfico da FDP – Função Distribuição de Probabilidade - de Bons e Maus pelo Score, verificamos aproximadamente que entre as pontuações de 0 a 19 a probabilidade de obter resultado mau é superior a de se obter resultado bom, requerendo maior atenção para esta faixa. Para escores maiores que 85 a probabilidade maior é de se obter resultados bons. Esta situação pode ser observada na Figura 2 a seguir.



FIGURA 2 - Gráfico da Função de Distribuição de Probabilidade de Bons e Maus pelo Score

Pode-se aumentar ou diminuir o valor do escore adotado (pouco maior que 18) de acordo com a limitação dos recursos financeiros disponíveis.

6. ANÁLISE DESCRITIVA DOS DADOS

Utilizando-se o algoritmo de indução de regras foram geradas diversas regras, tendo como alvo as ocorrências ou não de falhas na LT. Para efeito desse estudo, os dados considerados como relevantes foram aqueles com representatividade maior que 5% da massa e tomando como base os 4,24% de maus da massa de dados completa (ver Figura 1). Podemos considerar como regras ruins àquelas que são superiores a esse limiar e regras boas as que ficam abaixo dele.

Caracterizando o percentual de falhas e o alcance em termos da parcela dos defeitos, pode-se observar algumas condições de maior relevância, conforme descrito a seguir.

BONS:

- ✓ Se uma LT tem entre 20-39 km e sua capacidade de condução de corrente é de até 250 A, então existe uma probabilidade de falha de 0,08%, representando 15.53% da massa de dados.
- ✓ Linhas de Transmissão que não são consideradas como especiais que foram inspecionadas através de patrulhamento anual, apresentam uma probabilidade de falha de 0,44% em 10,74% da massa.

MAUS:

- ✓ Linhas de Transmissão que possuem cabos com bitolas maior que 4/0 AWG e que estão localizadas em áreas propensas a queimadas ou incêndio provocado possuem uma probabilidade de falha de 21,16%, representando 06.30% da massa.
- ✓ Em grandes linhas de transmissão (acima de 40km) com aterramento defeituoso, existe uma probabilidade de falha de 50,49% em cerca de 5% dos defeitos registrados.

7. RESULTADOS E CONSIDERAÇÕES

Com a ordenação dos defeitos nas linhas de transmissão por escore, foi possível fazer uma seleção iniciando pela pontuação mais baixa, ou seja, com maior risco de falha (maus) até o limite dos recursos financeiros disponíveis (ver Figura 2).

Diante da análise descritiva dos dados, puderam ser ainda observados alguns aspectos interessantes referentes ao planejamento da manutenção:

- ✓ A qualidade do serviço das empresas contratadas na realização de inspeções minuciosas e termográficas é comparável a das equipes próprias da empresa. A contratação externa fica dependendo apenas do tipo de estratégia a ser adotada.
- ✓ Foi possível verificar que em linhas que possuem grandes carregamentos, após a realização de inspeções minuciosas, praticamente todos os defeitos com potencial de falha são corrigidos em caráter emergencial, uma vez que não foram registrados tais defeitos no banco de dados do SAP/R3. Isso se deve ao fato dos órgãos executores da manutenção temerem uma falha intempestiva, o que causaria uma perda devido a faturamento cessante considerável.
- ✓ A probabilidade de ocorrer falha em linhas que não são consideradas como especiais é menor que 1% em 45% dos registros de defeitos. Por esse motivo é possível propor a priorização dos recursos financeiros na correção dos defeitos encontrado em linhas especiais.
- ✓ De acordo com as estatísticas verificadas, cerca de 5% das LT's com capacidade de condução de corrente elevada estão sem acesso e têm 32% de chance de falhar. Isto significa que se houver uma ocorrência nessas condições, o tempo para restabelecimento do sistema provavelmente será alto devido à falta de acesso.

Algumas melhorias ainda podem ser realizadas para auxiliar na resolução do problema em questão, tais como a padronização de valores, bem como a captação de novas variáveis que contemplem melhor a realidade em campo.

8. CONCLUSÃO

A mineração de dados é uma forte aliada da manutenção centrada na confiabilidade, uma vez que fornece subsídios para importantes tomadas de decisões e direcionamento dos recursos financeiros para a solução de defeitos que tenham um potencial maior de falhas. A utilização de técnicas de inteligência computacional no diagnóstico de falhas nas linhas de transmissão vem sendo experimentada como uma boa alternativa para a gestão de manutenção com excelentes possibilidades de otimização dos custos e garantia da disponibilidade do sistema da Celpe.

9. REFERÊNCIAS BIBLIOGRÁFICAS

- (1) ALMEIDA, T. A.; SOUZA, F. M. C. (organizadores), "Gestão da Manutenção – Na direção da competitividade". Pernambuco: UFPE, 2001.
- (2) U.M.Fayyad, G.Piatetsky-Shapiro e P.Smyth, "From Data Mining to Knowledge Discovery: An Overview", Menlo Park, AAAI Press, 1996.
- (3) Witten, I. H., and E. Frank, "Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations", Morgan Kaufmann, California, USA, 2000
- (4) Major J., and Mangano J. "Selecting among Rules Induced from a Hurricane Database". In Proc of 1993 KDD Workshop, pp. 28--47, Washington, DC, July, 1993.
- (5) Haykin, S., "Neural Networks: a comprehensive foundation" , 2nd ed., Prentice Hall Press PTR, 1998.
- (6) D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation", MIT Press Computational Models of Cognition and Perception Series, Vol. 1, 1986, pp. 318-362.
- (7) K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators", Neural Network, 1989, Vol. 2 (5), pp. 359-366
- (8) L. Prechelt, "Proben1 – A Set of Neural Network Benchmark Problems and Benchmarking Rules" Technical Report 21/94, Fakultät für Informatik, Universität Karlsruhe, Germany, 1994