

## **DETECÇÃO E ISOLAÇÃO DE FALHAS BASEADO NA TÉCNICA DE AGRUPAMENTO POSSIBILÍSTICO**

**Edgar Jhonny Amaya, eamaya@unb.br**

**Alberto José Álvares, alvares@AlvaresTech.com**

Universidade de Brasília, Departamento de Engenharia Mecânica e Mecatrônica, Grupo de Inovação em Automação Industrial (GIAI), CEP 70910-900, Brasília, DF, Brasil

**Resumo.** Neste artigo é apresentada uma abordagem alternativa para detecção e isolação de falhas baseada na técnica de agrupamento possibilístico. O algoritmo proposto é aplicado para classificação e agrupamento de padrões de falhas de máquinas. O agrupamento é realizado a partir do histórico de falhas e suas variáveis associadas obtidas do banco de dados de um sistema de monitoramento. A técnica de agrupamento possibilístico objetiva superar as desvantagens do algoritmo fuzzy c-means (FCM). As restrições probabilísticas impostas pelo algoritmo FCM são permitidas no agrupamento possibilístico. Sem essas restrições a abordagem de agrupamento possibilístico apresentada neste artigo fornecerá melhores resultados nas tarefas de detecção e isolação de falhas. Esta técnica de agrupamento é usada para determinar regiões de operação normal e de falhas de uma máquina, e localizar os centróides de todas as regiões. O comportamento defeituoso de uma máquina é quantificado calculando a distância euclidiana entre um dado atual e os centróides dos grupos de falhas. O estudo de caso do algoritmo proposto é realizado no sistema de resfriamento do gerador elétrico da usina hidrelétrica de Balbina. Uma das vantagens deste método é que não requer o modelo matemático do sistema diminuindo a complexidade e o custo de implementação. Os resultados esperados são o agrupamento das variáveis adquiridas em regiões de operação normal e de falhas, permitindo a detecção e isolação de falhas. Posteriormente, o algoritmo será implementado no módulo de diagnóstico e fornecerá informações para o módulo de prognóstico de um sistema de manutenção preditiva.

**Palavras-chave:** Detecção de falhas, isolação de falhas, c-means, agrupamento possibilístico, classificação de padrões.

### **1. INTRODUÇÃO**

A frequência de ocorrência de falhas nas plantas de processos e máquinas acrescenta o custo de produção, afetando a qualidade do produto, reduzindo a velocidade de produção e limitando a flexibilidade do sistema (Tay e Shen, 2003). Conseqüentemente é muito importante entender como as falhas ocorrem em uma planta de processos. A detecção a tempo ajudará localizar e controlar as conseqüências com flexibilidade. Também permitirá maximizar a qualidade do produto e estabilizar a velocidade de produção (Tay, 1999). Em uma planta de processos, centenas de variáveis são monitoradas e armazenadas a cada segundo. Estas medições fornecem informações do estado de saúde das máquinas da planta. Os métodos baseados em dados históricos tentam extrair o máximo de informação dos dados armazenados e requerem informações mínimas do modelo físico da planta. Os métodos baseados em dados históricos freqüentemente são os preferidos em lugar dos métodos baseados em modelos, quando não se dispõe o modelo acurado da planta e sim os dados armazenados pelos sistemas de supervisão e controle.

O método de detecção e isolação de falhas baseado em dados históricos explora os dados armazenados e identifica os padrões que indicam ocorrência de falhas (Detroja *et al.*, 2006). Tipicamente, as falhas paramétricas e as influencias dos sensores e atuadores geram diferentes padrões nas variáveis de processo. Estes padrões ou sinais podem ser classificados em diferentes grupos que representem operação normal ou anormal. Subseqüentemente, o modo de operação da planta pode ser classificado em termos das propriedades ou valores de pertinência dos dados aos grupos conhecidos. Deste modo, a tarefa de monitoração é freqüentemente relacionada à capacidade de classificar os sinais disponíveis da planta (variáveis medidas e manipuladas) em modos de operação normal ou com falhas. Isto pode ser efetivamente realizado por varias técnicas de agrupamento e reconhecimento de padrões.

As técnicas de agrupamento fuzzy c-means (FCM) (Duda *et al.*, 2003) e fuzzy Gustafson-Kessel (FGK) proposto por Gustafson e Kessel 1979 são populares em análise de imagens e classificação de padrões. Na área de detecção e isolação de falhas (FDI – *Fault Detection and Isolation*) existem esforços no uso de diferentes tipos de algoritmos. A técnica de agrupamento k-means é uma técnica robusta de agrupamento e classificação de padrões. Esta técnica de agrupamento k-

means foi aplicada por Oliveira *et al.* (2009) na detecção de falhas de um sistema de helicóptero de três graus de liberdade. Para avaliar o desempenho dos algoritmos desenvolvidos, Oliveira (2010) aplica o método das curvas de característica de operação de receptor (ROC – *Receiver Operating Characteristic*) e a área sob a curva (AUC – *Area Under Curve*). Uma metodologia de três etapas que inclui o algoritmo k-means junto com a análise dos componentes principais (PCA – *Principal Components Analysis*) e a análise discriminante de Fisher (FDA – *Fisher Discriminant Analysis*) é proposto por Peter He *et al.* 2005. Esta metodologia é aplicada na tarefa de detecção e diagnóstico de falhas (FDD – *Fault Detection and Diagnosis*). Teppola e Minkkinen 1999 usaram FCM adaptativo para monitoração do processo de uma planta de águas residuais. Choi *et al.* 2003 aplicaram o algoritmo de agrupamento de confiabilidade baseado na abordagem de monitoração de processos. Porém, os métodos acima são propostos para processamento de diferentes cenários presentes em dados armazenados e não podem lidar com cenários de novos tipos de falhas, quando os dados armazenados não contem informações para novas condições de falha.

Uma das limitações dos algoritmos de agrupamento existentes é que os sinais resultantes de falhas similares, mas com diferente intensidade poderia confundir ao algoritmo na etapa de isolamento de falhas. Também é importante destacar que os dados armazenados não necessariamente incluem todos os possíveis cenários ou tipos de falhas. Idealmente seria desejável construir um método de monitoração que é baseado em dados históricos existentes e atualizar a base de conhecimento assim que uma nova falha aparecer durante o processo de monitoração em tempo real. Além disso, o algoritmo FDD precisa ter a capacidade de aprendizagem. Durante a execução em tempo real, o algoritmo deveria ser capaz de identificar a ocorrência de novas falhas e estabelecer características ou padrões relevantes. Uma importante característica relevante é assegurar que uma nova situação de falha não seja mal-classificada como um caso de falha conhecida e conseqüentemente gerar um diagnóstico inconsistente.

Neste artigo, a proposta é superar a dificuldades citadas anteriormente usando o algoritmo de agrupamento possibilístico (Krishnapuram e Keller, 1993). Agrupar os dados em região de operação normal e regiões de falhas. Os dados analisados são variáveis de temperatura de ar frio e ar quente dos radiadores do sistema de resfriamento do gerador elétrico da usina hidrelétrica de Balbina.

## 2. REVISÃO DE ALGORITMOS DE AGRUPAMENTO

O objetivo de classificação ou agrupamento é gerar partições a partir de um conjunto de  $N$  dados baseado em algumas similaridades métricas. Os dados de um grupo devem ser similares uns aos outros. A seguir apresenta-se uma revisão de técnicas de agrupamento não hierárquico. Os algoritmos de agrupamento não hierárquico têm como objetivo gerar um número de grupos pré-especificado. O agrupamento começa a partir de uma condição inicial do conjunto de dados e continua um processamento iterativo.

O problema de otimização pode envolver muitas variáveis inteiras, podendo levar muito tempo o processamento. No entanto, teoricamente pode ser comprovada a convergência a um ótimo global. Para superar o problema das variáveis inteiras é usada uma variante ao algoritmo de agrupamento não hierárquico, conhecido como c-means. No agrupamento c-means, os dados são designados aos grupos com centróides mais próximos, e os centróides dos novos grupos são calculados a cada iteração durante o processo de otimização, até que o valor da coordenadas do centróide do grupo consiga convergir a um valor constante. Este método iterativo algumas vezes pode produzir soluções não ótimas. O algoritmo de agrupamento c-means é também conhecido como uma técnica de agrupamento robusta, no qual um dado é membro de um grupo só. Algumas vezes é difícil ou mesmo incorreto designar alguns dados a um grupo só. No algoritmo FCM esta condição é permitida e cada dado pode ser membro de mais de um grupo.

### 2.1 Agrupamento *fuzzy* c-means

Na técnica de agrupamento FCM, os dados podem ser membros de mais de um grupo com diferentes graus de pertinência. O valor de pertinência do dado  $j$ th ao grupo  $i$ th é  $\mu_{ij}$ . Para o conjunto de dados  $\{X|x_1, \dots, x_N \in X\}$ , que possui  $c$  grupos, o algoritmo FCM minimiza a função objetivo Eq. (1) sujeito às restrições da Eq. (2), Eq. (3) e Eq. (4).

$$\min J \sum_{i=1}^c \sum_{j=1}^N (\mu_{ij})^m d_{ij}^2, \quad (1)$$

$$\sum_{i=1}^c \mu_{ij} = 1, \quad (2)$$

$$0 < \sum_{j=1}^N \mu_{ij} < N, \quad (3)$$

$$\mu_{ij} \in [0, 1], \quad (4)$$

A onde,  $c$  é o número de grupos,  $N$  é o número total de dados,  $m$  é o expoente *fuzzy* e  $d_{ij}$  é a distância entre o dado  $x_i$  e o centro do grupo  $v_i$ . A restrição imposta pela Eq. (2), também conhecida como restrição probabilística, no qual o somatório dos valores de pertinência de todos os dados aos grupos deve ser a unidade.

O algoritmo FCM inicia como algumas suposições para cada matriz *fuzzy* formado pelo centróide dos grupos e itera até a convergência (Duda *et al.*, 2003). A convergência do algoritmo FCM é garantida (Bezdek, 1980), podendo ele convergir a um mínimo local.

O valor de pertinência de um dado a um grupo não depende só da distância do dado ao centróide do grupo, também depende da distância do dado aos centróides dos grupos. Esta restrição pode causar que o algoritmo assinale diferentes valores de pertinência aos dados com a mesma distância ao centróide de um grupo, sendo que a distância desses dados aos centróides dos grupos sejam diferentes. Este problema de inconsistência aparece devido às restrições probabilísticas descritas na Eq. (2) e pode ser descrito como (i) os pontos equidistantes desde o centróide pode ter valores de pertinências diferentes dependendo da localização dos outros grupos, e (ii) os pontos que são equidistantes de todos os centróides terão os mesmos valores de pertinência. Para superar essa desvantagem, Krishnapuram e Keller 1993 propõem uma nova técnica de agrupamento chamado agrupamento possibilístico, permitindo as restrições probabilísticas.

## 2.2 Agrupamento possibilístico

No agrupamento possibilístico, a restrição probabilística (Eq. (2)) na função objetivo (Eq. 1) é permitida para obter os valores de pertinência que representem "grau típico" de pertinência a um grupo. A aplicação das restrições probabilísticas produz uma solução trivial, por exemplo, a função objetivo é minimizada assinando todos os valores de pertinência a zero. Então a função objetivo (Eq. (1)) é modificada, resultando a Eq. (5).

$$\min J \sum_{i=1}^c \sum_{j=1}^N (\mu_{ij})^m d_{ij}^2 + \sum_{i=1}^c \eta_i \sum_{j=1}^N (1 - \mu_{ij})^m, \quad (5)$$

O primeiro termo da Eq. (5) minimiza a distância dos dados aos centros dos grupos, enquanto o segundo termo força aos valores de pertinência ser o maior possível. O valor de  $m$  determina a nebulosidade da partição possibilística final. O valor do parâmetro  $\eta_i$  determina quando o valor de pertinência de um dado a um grupo chega a ser 0.5. O parâmetro  $\eta_i$  deve ser escolhido dependendo do ancho de banda da função possibilidade para cada grupo (Detroja *et al.*, 2006). Na prática, no entanto, pode ser usada a definição da Eq. (6) (Krishnapuram e Keller, 1993):

$$\eta_i = \frac{\sum_{j=1}^N (\mu_{ij})^m d_{ij}^2}{\sum_{j=1}^N (\mu_{ij})^m} \quad (6)$$

A atualização dos valores de pertinência depende do tipo de cálculo da distância medida. Se a medida de distância escolhida é Euclidiano ou Mahalanobis, o algoritmo fornecerá valores de pertinência Possibilístico c-means (PCM). Porém, se a medida de distância é baseada na distância Mahalanobis em escala e a matriz de covariância *fuzzy*, o resultado do algoritmo serão valores de pertinência Possibilístico Gustafson-Kessel (PGK). A solução da função objetivo (Eq. (5)) serão valores de pertinência assim como descrito pela Eq. (7).

$$\mu_{ij} = \frac{1}{1 + \left(\frac{d_{ij}^2}{\eta_i}\right)^{1/(m-1)}} \quad (7)$$

O algoritmo para agrupamento possibilístico é similar ao algoritmo FCM, exceto para os parâmetros adicionais  $\eta_i$  o qual poderia ser estimado desde a matriz de agrupamento inicial. Porém,  $\eta_i$  precisa ser calculado a cada iteração.

Desde que o parâmetro  $\eta_i$  seja independente da localização relativa dos grupos, o valor de pertinência  $\mu_{ij}$  dependerá unicamente da distância de um ponto ao centróide do grupo. Portanto, a diferença do caso probabilístico, o valor de pertinência de um ponto a um grupo é determinada unicamente pela distância do ponto ao centróide e não tem relação com a localização respeito aos outros grupos.

As vantagens dos algoritmos PCM e PGK estão em encontrar grupos significantes definidos pela densidade da região. Isto acontece porque cada grupo é independente dos outros grupos nos algoritmos PCM e PGK. Portanto, a função objetivo que corresponde ao grupo  $i$  pode ser calculada usando a Eq. (8).

$$\min J_i = \sum_{j=1}^N (\mu_{ij})^m d_{ij}^2 + \eta_i \sum_{j=1}^N (1 - \mu_{ij})^m. \quad (8)$$

### 3. ABORDAGEM DE DIAGNÓSTICO BASEADO EM DADOS HISTÓRICOS

As abordagens baseadas em agrupamento são focadas em classificar os dados históricos em um determinado número de grupos, por exemplo, grupos de operação normal e dos diferentes tipos de falhas. Dependendo do valor de pertinência dos dados aos diferentes grupos, a operação da planta é declarada normal ou defeituosa. O deslocamento de um dado que encontra-se em um grupo de operação normal a qualquer um dos grupos de falha não é instantâneo e a resposta transiente depende da dinâmica do sistema. O algoritmo FCM poderá assinar diferentes valores de pertinência a pontos iguais durante as transientes, controlados probabilisticamente segundo a Eq. (2).

#### 3.1 Aquisição e compressão de dados

A capacidade de uma abordagem estatística para detectar e isolar falhas depende da disponibilidade e da riqueza de dados históricos. Estes dados contêm informações dos modos de operação normal e defeituosa da máquina monitorada. Idealmente, o conjunto de dados para o treinamento da técnica baseado em agrupamento deveria conter informações que representem todos os possíveis cenários de falha. Na prática, no entanto, não é possível ter todos os conjuntos de dados e o algoritmo deve ser capaz de aprender com a capacidade de aprendizagem.

Em uma planta de processo é monitorada uma grande quantidade de variáveis correlacionadas. Em uma planta de grande porte pode ser muitas vezes difícil executar agrupamentos de todos os conjuntos de dados. Então é importante comprimir os dados mantendo informações relevantes do conteúdo. Técnicas de compressão de dados multi-variável como PCA podem ser usados nestes casos para obter uma aproximação simplificada dos dados. A representação simplificada dos dados deve representar adequadamente o comportamento da planta e discriminar entre uma operação normal e diferentes tipos de falhas. Então, o agrupamento pode ser executado neste espaço dimensional simplificado ao invés do espaço total.

#### 3.2 Agrupamento dos dados históricos

Em geral, os dados históricos contêm informações das variáveis medidas e manipuladas a cada período de amostragem ou a representação dimensional simplificada. A abordagem de agrupamento poderia diretamente processar as medidas ou gerar um vetor de características que represente as medidas simplificadas. No primeiro caso, a classificação é realizada diretamente no espaço das medidas. No segundo caso, o agrupamento é realizado no espaço definido pelo vetor de características das medidas. Esta última abordagem é aplicada neste artigo devido à alta quantidade de dados. O agrupamento e análise é realizado usando os três componentes principais de um total de nove variáveis.

O algoritmo de agrupamento pode ser aplicado na matriz que contém os dados das variáveis associadas aos modos de operação. A especificação do número exato de grupos presentes no conjunto de dados é obrigatório na abordagem proposta. O algoritmo de agrupamento possibilístico procura os melhores  $c$  grupos. Neste artigo o número de grupos considerado é quatro, um grupo de operação normal e três de diferentes tipos de falhas.

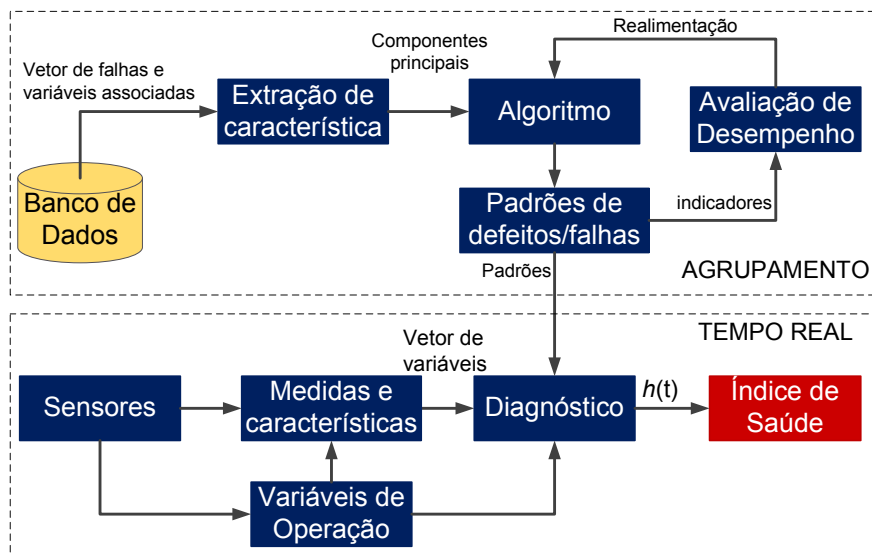


Figura 1. Arquitetura para detecção e isolamento de falhas.

### 3.3 Detecção e isolamento de falhas

Na etapa de detecção e isolamento de falhas é importante destacar que os valores de pertinência aos grupos de falhas podem ser influenciados pela dinâmica inerente das máquinas. O perfil de pertinência também pode mudar pela presença de transientes e por medidas ruidosas. Deste modo, é importante confirmar a ocorrência de uma falha depois de ser detectada. Para a confirmação e isolamento das falhas propõe-se usar uma janela de  $M$  amostras consecutivas no qual o valor de pertinência é analisado.

Tabela 1. **Dados históricos de temperatura.**

Data e Hora	GAF1	GAF2	GAF3	GAF4	GAF5	GAF6	GAF7	GAF8	GAQ1
5/3/08 16:31	35,8262	35,4873	35,9343	35,8266	34,6353	35,8874	35,8565	35,8789	49,2054
5/3/08 18:47	35,5193	35,1754	35,5709	35,4927	34,3896	35,5344	35,4844	35,5714	48,1150
5/3/08 18:48	35,5086	35,1755	35,9808	35,4924	34,3295	35,5446	35,4754	35,6589	48,1116
5/3/08 18:49	35,5308	35,1856	35,5904	35,5146	34,3110	35,6374	35,5086	35,6932	48,2174
5/3/08 18:50	35,5350	35,6757	35,6205	35,5547	34,3611	35,7385	35,5816	35,5509	48,1430
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
27/11/09 23:54	36,0477	35,5068	36,1096	36,0807	34,5542	36,0748	36,0446	36,0664	49,0794
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
16/11/09 8:23	37,4818	37,0849	37,2376	37,3458	35,8129	37,3097	37,7448	37,5017	49,9974

Tabela 2. **Limites de operação.**

Tag	Descrição	Alto	Alarme	Trip
GAF1-8	Temperatura de ar frio do radiador nº 1,2,...,8	>44°C	>45°C	
GAQ1	Temperatura de ar quente	>70°C	>75°C	>85°C

A arquitetura proposta neste artigo é mostrada na Fig. (1). Os dados coletados do banco de dados tem a estrutura assim como mostrado na Tab. (1) com gráfico de tendências mostrado na Fig. (2). Essas variáveis passam pela etapa de extração de características no qual a traves do método PCA, extrai-se os componentes mais relevantes. O agrupamento é realizado usando como dados os principais componentes, gerando padrões de operação normal e falhas. Com os indicadores de desempenho o algoritmo de agrupamento pode ser realimentado para melhorar seu desempenho. Os padrões classificados pelo algoritmo será uma entrada para o processamento em tempo real do sistema de detecção e isolamento de falhas.

No processamento em tempo real, o sistema adquire dados dos sensores formando um vetor de variáveis. A detecção de uma falha é realizada verificando se o vetor de variáveis associadas esta for da região de operação normal, delimitado por um limiar de operação normal. Uma vez detectada uma falha, na etapa de isolamento determina-se a tendência da falha a uma das regiões de falhas estabelecidas no processo de agrupamento. A função da etapa de diagnóstico é calcular o grau de pertinência a um dado aos grupos de falha. O resultado final do módulo de diagnóstico será o índice de saúde (Eq. (9)) da máquina monitorada que é uma função que depende do valor de pertinência e da distância euclidiana do vetor de variáveis aos grupos de falhas.

$$h_i = f(\mu_{ij}, d_{ij}) \quad (9)$$

## 4. ESTUDO DE CASO: SISTEMA DE RESFRIAMENTO

Este sistema é introduzido ao sistema de resfriamento do gerador (SRG) da unidade geradora hidráulica (UGH) tipo Kaplan da usina hidrelétrica de Balbina (UHB). A UHB é localizada no Município de Presidente Figueiredo no Estado de Amazonas. A UHB é constituído de cinco UGH cada uma com uma potência nominal de 50MW, somando um total de 250 MW. A velocidade de rotação é 105,88 rpm e tem uma queda líquida de 21.85m.

O SRG consiste de oito trocadores de calor instalados equidistantes um dos outros ao redor do gerador elétrico, assim como é mostrado na Fig. (3). Os trocadores de calor são de tubo, usando como fluido de resfriamento água a temperatura ambiente. As medições realizadas no SRG são as oito medições de temperatura de ar frio, com transmissores GAF1, GAF2, GAF3, GAF4, GAF5, GAF6, GAF7 e GAF8. Também é monitorada a temperatura de ar quente com o transmissor GAQ1. Alguns dos valores destas variáveis coletadas durante o monitoramento são mostradas na Tabela (1).

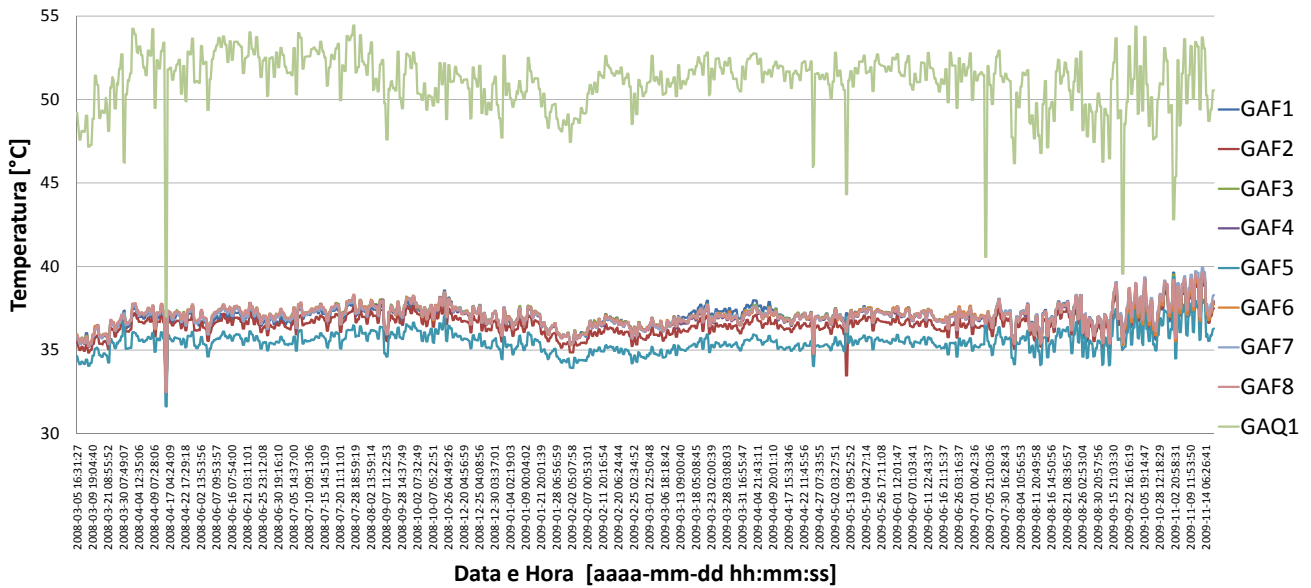


Figura 2. Tendência dos sinais de temperatura de ar frio e ar quente.

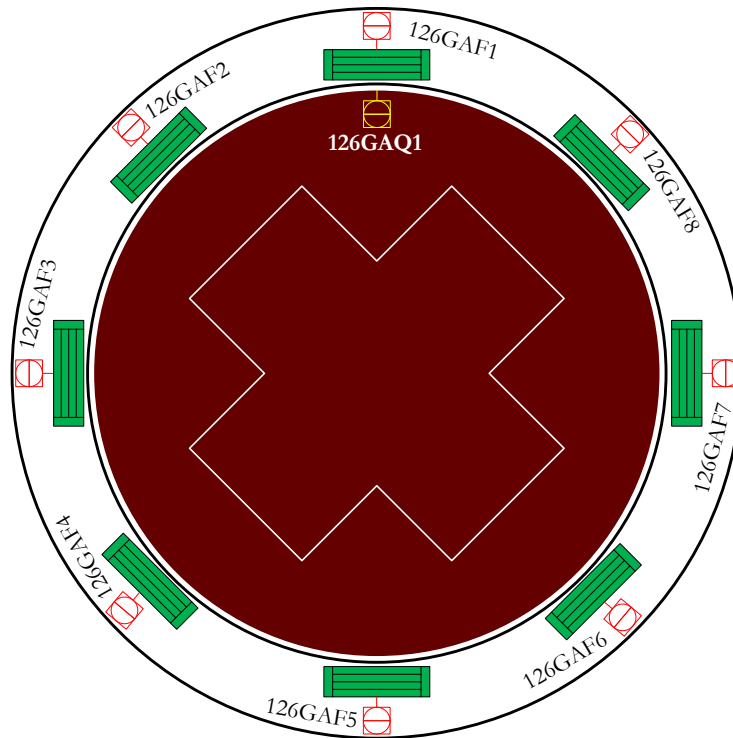


Figura 3. Localização dos transmissores de temperatura de ar frio e ar quente no gerador elétrico.

Na Figura (3) mostra-se a gráfica de tendência dos dados coletados, observa-se a presença de transientes e regiões de evolução de falhas, os limites de operação destas variáveis (alto, alarme e trip) são mostrados na Tabela (2). Os dados coletados passam pelo de processo de análise dos componentes principais no qual reduzimos a dimensão dos dados. Na Figura (4) mostra-se a porcentagem acumulada dos componentes principais, com três componentes consegue reproduzir mais de 97% dos valores verdadeiros.

A aplicação do algoritmo de agrupamento possibilístico permitiu descobrir a presença de grupos de falhas e grupo de operação normal do sistema que pode ser observado na Fig. (5). O número de grupo de falhas escolheu-se baseado na análise dos modos e efeitos das falhas (FMEA – *Failure Mode and Effect Analysis*), baseado na árvore de falhas (FTA – *Fault Tree Analysis*) e as falhas mais comuns dos últimos anos.

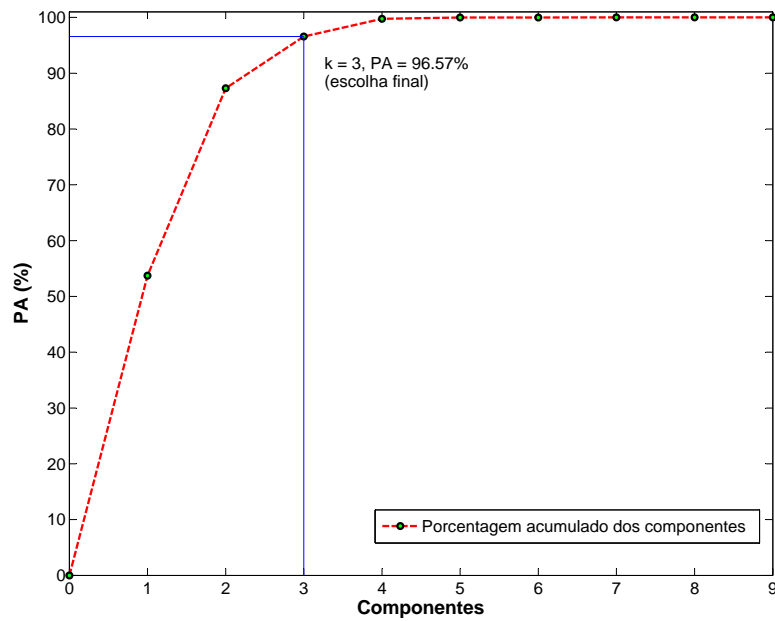


Figura 4. Escolha do número componentes principais.

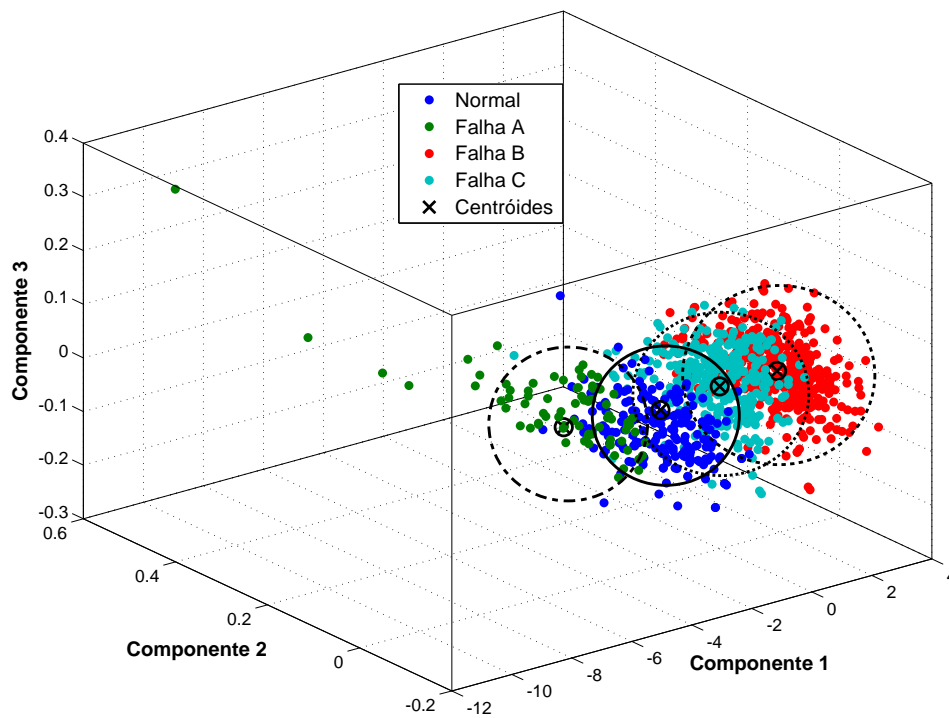


Figura 5. Regiões de operação normal e falhas.

## 5. CONCLUSÕES

Neste artigo apresenta-se os resultados do algoritmo para detecção e isolamento de falhas usando o agrupamento possibilístico. O algoritmo é aplicado ao agrupamento dos dados em regiões de operação normal e de falhas. Os modos de operação de novos dados poderam ser diagnosticados calculando a distância euclidiana de cada conjunto de dados aos centróides dos grupos. O algoritmo processa os dados do sistema de resfriamento de gerador elétrico da UHB e mostrou-se capaz de isolar e detectar três tipos de falhas ocultas aos limites de operação normal, alto, alarme e trip que são verificados pelo pessoal de operação e manutenção. Este algoritmo será introduzido no módulo de diagnóstico do sistema de manutenção preditiva da UHB descrito em outros artigos por Amaya *et al.* 2008, Amaya *et al.* 2009a e Amaya *et al.* 2009b. Os trabalhos futuros incluem determinar a função do índice de saúde, a estimação dos valores futuros dos

índices de saúde serão usados pelo módulo de prognóstico com o objetivo de calcular o RUL (*Remaining Useful Life*) do estado de saúde das máquinas.

## 6. AGRADECIMENTOS

Agradecemos o suporte da Eletronorte e da Manaus Energia no projeto de pesquisa e desenvolvimento com contrato 4500052325, projeto número 128 "Modernização das usinas hidrelétricas de Balbina e Samuel" que tem como responsável ao professor Alberto Álvares da Universidade de Brasília. Agradecemos também ao Eng. Antonio Araújo da Eletronorte pelas informações cedidas que foram de extrema importância para o desenvolvimento do presente trabalho.

## 7. REFERÊNCIAS

- Amaya, E. J., Alvares, A., Tonaco, R., Souza, R. e R. Gudwin, R., 2008, "An Intelligent Kernel for the Maintenance System of a Hydroelectric Power Plant", in *Proc. ABCM Symposium Series in Mechatronics*, Vol. 3, pp. 821-830.
- Amaya, E. J., Alvares, A. e Gudwin, R., 2009a, "Open System Architecture for Condition Based Maintenance Applied to a Hydroelectric Power Plant", The 8th Latin-American Congress on Electricity Generation and Transmission (CLAG-TEE), ISBN - 978-85-61065-01-0, October 18-22, Ubatuba, SP, Brazil.
- Amaya, E. J., Alvares, A. e R. Gudwin, R., 2009b, "An Expert System for Fault Diagnostics in Condition Based Maintenance", 20th International Congress of Mechanical Engineering (COBEM), November 15-20, Gramado, RS, Brazil, 2009.
- Bezdek, J.C., 1980, "A convergence theorem for the fuzzy ISODATA clustering algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 2, No. 1, pp. 1-8.
- Choi, S.W., Yoo, C.K. e Lee, I., 2003, "Overall statistical monitoring of static and dynamic patterns", *Industrial and Engineering Chemistry Research*, Vol. 42, pp. 108-117.
- Oliveira, E. C., Garcia, G., Matsuura, J. P. e Galvão, R. K. H., 2009, "Fault detection in a 3DOF helicopter system", 20th International Congress of Mechanical Engineering (COBEM), Gramado, RS, Brasil.
- Oliveira, E. C., 2010, "Detecção de Falhas no Helicóptero 3DOF", Tese de Mestrado, 79f, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, Brasil.
- Detroja, K.P., Gudi, R.D. e Patwardhan, S.C., 2006, "A possibilistic clustering approach to novel fault detection and isolation", *Journal of Process Control*, Vol. 16, pp. 1055-1073.
- Duda, R.O., Hart, P.E. e Stork, D.G., 2003, "Pattern Classification", Wiley-Interscience publication, New York,.
- Gustafson, D.E. e Kessel, W.C., 1979, "Fuzzy clustering with a fuzzy covariance matrix", in: Proceedings of IEEE CDC, San Diego, CA, USA, pp. 761-766.
- Krishnapuram, R. e Keller, J.M., 1993, "A possibilistic approach to clustering", *IEEE Transactions on Fuzzy Systems*, Vol. 1, No. 2, pp. 98-110.
- Peter He, Q., Wang, J. e Joe Qin, S., 2005, "A new fault diagnosis method using fault directions in Fisher discriminant analysis", *AICHE*, Vol. 51, No. 2, pp. 555-571.
- Tay, E.H., 1999, "Contingency management in flexible manufacturing systems using modal state logic", *Journal of Manufacturing Systems*, Vol. 18, No. 5, pp. 345-357.
- Tay, E.H. e Shen, L., 2003, "Fault diagnosis based on Rough Set Theory", *Engineering Applications of Artificial Intelligence*, Vol. 16, pp. 39-43.
- Teppola, P. e Minkkinen, P., 1999, "Possibilistic and fuzzy c-means clustering for process monitoring in an activated sludge waste-water treatment plant", *Journal of Chemometrics*, Vol. 13, pp. 445-459.

## 8. DIREITOS AUTORAIS

Os autores são os únicos responsáveis pelo conteúdo do material impresso incluído neste artigo.

## FAULT DETECTION AND ISOLATION BASED ON POSSIBILISTIC CLUSTERING TECHNIQUES

Edgar Jhonny Amaya, eamaya@unb.br

Alberto José Álvares, alvares@AlvaresTech.com

University of Brasília (UnB), Departamento de Engenharia Mecânica e Mecatrônica, Grupo de Inovação em Automação Industrial (GIAI), Grupo de Automação e Controle (GRACO), CEP 70910-900, Brasília, DF, Brasil.

**Abstract:** *In this paper is presented an alternative approach for fault detection and isolation based in the possibilistic clustering technique. The proposed algorithm is applied do fault machine clustering and pattern classification. The clustering is realized from historic fault data and its associates variables that was collected from monitoring system*



*database. The possibilistic clustering technique address some of the shortcomings of the fuzzy c-means(FCM) algorithm. The probabilistic constraint imposed by the FCM is relaxed in the possibilistic clustering. Because of this relaxation, the possibilistic approach presented in this paper will be give better results in the for fault detection and isolation tasks. This clustering technique is used to calculate the normal operation cluster and fault cluster of the machine, and determine the centroid of the clusters. The defective behavior of a machine is quantified determining the Euclidian distance between one data and the fault centroid clusters. The case study of the proposed algorithm is to perform in cooling system of the electric generator of the hydroelectric power plant of Balbina. An advantage of this method is that not require the mathematical model of the system reducing the complexity and cost implementation. The results are the clustering of the acquire variables in normal operation cluster and fault clusters that permit the fault detection and isolation. This algorithm will be implemented in the diagnostic module of a predictive maintenance system.*

**Keywords:** *Fault detection, fault isolation, c-means, possibilistic clustering, pattern classification.*