



Grupo de Estudo de Sistemas de Informação e Telecomunicação para Sistemas Elétricos-GTL

ASSISTENTE COMPUTACIONAL PARA SUPERVISÃO DE GRANDEZAS OPERATIVAS

**FILIFE VENTURA MUGGIATI(1); FELIPE TREVISAN(1); JEFFERSON FERNANDO STUMPF(1);
PAULO ZANELLI JUNIOR(1); MILTON BROWN DO COUTTO FILHO(2); JULIO CESAR STACCHINI DE
SOUZA(2); BERNARDO PAULO TROST MORCELLI(1); ARSENIO ESTECHE RAMOS(1);
ITAIPU(1); UFF(2)**

RESUMO

Este informe técnico apresenta uma ferramenta de apoio à Operação que aplica técnicas de mineração de dados e aprendizado de máquina para a detecção de anomalias em séries temporais de grandezas físicas, registradas pelo sistema de supervisão e controle da usina de Itaipu. Inicialmente, os dados do SCADA são filtrados, transformados e agregados em um formato compatível com algoritmos de mineração de dados. Em seguida, uma rede neural não-supervisionada é aplicada para realizar uma pré-rotulagem e, posteriormente, novos dados são classificados em tempo real por um algoritmo supervisionado para indicar aos operadores grandezas que que apresentam comportamento anômalo.

PALAVRAS-CHAVE

Mineração de Dados, Aprendizado de Máquina, Séries Temporais, Detecção de Anomalias, Geração de Energia Elétrica

1.0 - INTRODUÇÃO

Os complexos processos envolvidos na produção de energia elétrica congregam uma grande quantidade de equipamentos e dispositivos operando ininterruptamente de modo a satisfazer padrões de qualidade e confiabilidade. A supervisão deste processo constitui tarefa difícil e importante, a ser realizada de forma a evidenciar possíveis situações de operação adversas. Para auxiliar na tarefa de supervisão, sistemas computacionais conhecidos como SCADA (*Supervisory Control and Data Acquisition*) são usualmente utilizados (1).

Tradicionalmente, a unidade mestre do sistema SCADA, conhecida como *Master Terminal Unit* (MTU), está integrada a um sistema de armazenamento de dados históricos, chamado de *historiador*, que funciona como um grande banco de dados de séries temporais. Isto permite que todas as grandezas físicas, aquisitadas através das diversas unidades remotas ou *Remote Terminal Units* (RTU), sejam registradas, tais como potência elétrica, tensão e comandos enviados aos atuadores ao longo do tempo para controle do processo. A título de exemplo, em Itaipu são historiadas aproximadamente 30.000 grandezas operativas entre medidas analógicas e indicações de estado provenientes do sistema SCADA. Adequadamente processados e analisados, esses dados podem revelar características importantes sobre equipamentos e sistemas da usina, como comportamentos anômalos que podem indicar falhas iminentes.

Neste contexto, a enorme quantidade de dados históricos registrados pelos sistemas digitais de supervisão e controle permite que técnicas de mineração de dados e aprendizado de máquina sejam utilizadas para criar modelos de identificação e correlação de anomalias, em espaços de busca multidimensionais. Este informe técnico apresenta uma ferramenta concebida para auxiliar as equipes de operação em tempo real da usina de Itaipu, capaz de construir modelos computacionais a partir da aplicação de algoritmos de mineração de dados em dados históricos operativos, destinados à identificação de anomalias e suporte à supervisão.

2.0 - DESCOBERTA DE CONHECIMENTO EM DADOS

A mineração de dados busca extrair informações úteis de bases de dados através da aplicação de algoritmos especializados. Trata-se de um processo que pode gerar um grande valor agregado devido aos benefícios que a informação minerada pode trazer a uma instituição, reduzindo custos e aumentando a produtividade (2). Porém, apesar da execução do algoritmo de mineração propriamente dito se tratar de um processo automático, esta

atividade está inserida em um contexto mais amplo, conhecido como *Knowledge Discovery from Data – KDD*, em que uma série de outros processos, nem sempre automáticos, são necessários para “preparar” os dados para serem minerados. Neste contexto, o termo “mineração” refere-se a uma atividade trabalhosa de extração de informações realmente valiosas, não constituindo tarefa trivial.

Adotando o processo de *KDD* como uma metodologia para a análise de dados, inicialmente se dá a seleção, pré-processamento e transformação dos dados. Uma das maneiras de executar esta primeira etapa denomina-se ETL (*extraction, transformation and load*), em que os dados de entrada são extraídos de suas diversas fontes originais, limpos, integrados, transformados e, por fim, carregados em um único banco de dados, conhecido por *Data Warehouse*. Este banco reúne todos os dados que serão futuramente processados pelos algoritmos de mineração de dados. Posteriormente, algoritmos de mineração de dados podem ser aplicados para extrair informações úteis à tomada de decisão.

2.1 Séries Temporais em Plantas Industriais

Conjuntos de valores de grandezas de interesse ordenados em função do tempo são denominados séries temporais. Assim, cada valor pertencente à série está associado a uma marca temporal (conhecida como estampa de tempo) referente ao comportamento dinâmico do processo em estudo.

Do ponto de vista computacional, um conjunto de séries temporais pode ser visto como uma tabela de um banco de dados com campos de identificação da série, estampa de tempo e valor. Tais campos podem ser usados para a construção de bancos de dados derivados, como um *Data Warehouse*, descrito no processo de descoberta de conhecimento - *KDD*. Para isso, deve-se definir quais os dados a importar e escolher as funções de limpeza e transformação que serão aplicadas às séries selecionadas. Estes dados são transferidos para um banco de dados (*DW*) e a estampa de tempo de cada amostra transforma-se em um atributo, relacionado a todos os eventos que aconteceram em um mesmo intervalo.

As séries temporais associadas ao problema abordado neste informe técnico compreendem em geral dois tipos básicos (e distintos) de valores: séries analógicas, que contêm valores reais contínuos no tempo e proporcionais às grandezas físicas; e séries de estado, que representam diferentes estados de um equipamento ou subsistema no tempo. A diferenciação entre tais séries é importante para o pré-processamento dos atributos e aplicação dos algoritmos, pois as séries analógicas devem ser tratadas considerando a continuidade de sua magnitude e sinal ao longo do tempo, enquanto os valores que representam as séries de estados não possuem qualquer associação do estado representado com a magnitude ou com o sinal do número que o representa, o qual poderia ser considerado como um símbolo, por exemplo.

Importante destacar que bancos de dados de séries temporais em plantas industriais utilizam diferentes frequências de amostragem das grandezas de interesse. Neste trabalho se emprega uma etapa de pré-processamento em que é realizada uma interpolação de dados para permitir a sincronização, ainda que em intervalos irregulares, de todas as grandezas envolvidas em uma determinada análise. Na Figura 1, pode-se observar que as três séries representadas (referentes às grandezas nível de óleo, temperatura e potência) têm frequências de amostragem diferentes e irregulares.

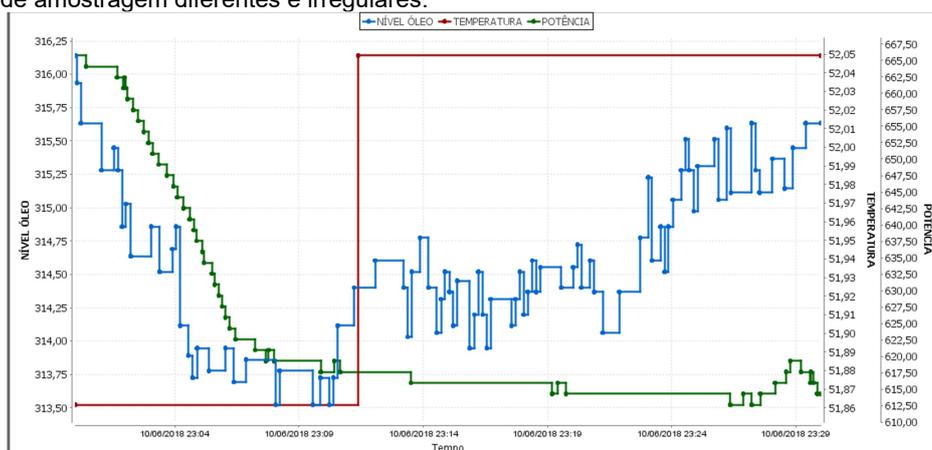


Figura 1 - Séries temporais com amostragens irregulares.

2.2 Técnicas de Mineração de Dados para a Detecção de Anomalias

Uma anomalia (*outlier*) representa um conjunto de amostras de uma série temporal que tem comportamento diferente do esperado (normal) (2, 3, 4). Sua origem difere daquela de ruídos no processo de aquisição dos dados e representa alterações físicas no processo que está sendo observado. Enquanto ruídos são inerentes ao processo de observação e indesejados, as anomalias representam informações importantes sobre o processo, uma vez que podem identificar falhas (inclusive em estágios iniciais) em um equipamento (5).

As anomalias podem ser caracterizadas como globais ou contextuais. As globais são caracterizadas por se manifestarem em relação a todo o conjunto de amostras, enquanto que as contextuais dizem respeito a um conjunto de amostras específico (contexto) e podem não ser consideradas anomalias em outro conjunto de amostras (contexto diverso) (6).

Existe ainda a definição de anomalias pontuais, representadas por uma única amostra no conjunto de dados, e coletivas, aquelas formadas por um conjunto de amostras que juntas representam um comportamento anômalo. Nestes dois casos, as anomalias podem se manifestar num contexto global ou contextual (2,7). Neste trabalho, foi utilizada uma estratégia de redução do conjunto de amostras em atributos específicos, transformando assim o problema em um caso de detecção de anomalia pontual, com diversos atributos (8). Porém, vale observar que esta agregação pode resultar em perda substancial de informação, como ocorre por exemplo quando os dados são agregados por média, deixando-se assim de representar comportamentos importantes para a detecção de anomalias.

Os algoritmos de mineração de dados podem ser também classificados de acordo com a existência ou não de rotulagem que indique se cada registro corresponde ou não a uma anomalia. Os métodos supervisionados utilizam todos os dados rotulados por especialistas para treinar modelos com base na associação de regras previamente apresentadas. Métodos semi-supervisionados usam apenas um subconjunto dos dados rotulados, e.g. somente os dados normais. Por fim, métodos não-supervisionados utilizam dados sem nenhuma rotulagem, situação muito comum no mundo real. No caso da detecção de anomalias, algoritmos não-supervisionados se dedicam ao aprendizado do que pode ser considerado como comportamento normal (7) e assumem que os dados correspondentes a essas situações são muito mais frequentes do que aqueles correspondentes à anomalias.

3.0 - METODOLOGIA PROPOSTA

A ferramenta apresentada neste trabalho foi construída a partir de uma metodologia que busca promover uma interação equilibrada entre o conhecimento de especialistas do setor elétrico de potência e as técnicas de descoberta do conhecimento (9). Para isso, foi definido um processo contendo três etapas principais (apresentadas nas próximas seções): seleção e transformação de dados, construção do conjunto de treinamento supervisionado e execução periódica para fins de diagnóstico.

3.1 Seleção e Transformação dos Dados

O objetivo desta etapa é transformar um conjunto de tamanho variável de séries temporais de diversas grandezas físicas, irregularmente intervaladas no tempo, em um conjunto de registros estruturados, com intervalos regulares e um número fixo de atributos, para que seus elementos possam ser comparados entre si e empregados em algoritmos de mineração de dados. Isto implica em agregar de alguma maneira este conjunto de dados no intervalo que foi definido para realizar a redução descrita na Seção 2.2. Além disso, como a análise pode estar relacionada a uma ou mais séries temporais, uma equação parametrizável, que relaciona estes atributos entre si, deve ser definida para transformar este número variável de entradas em um número finito de atributos.

Após a aplicação da equação para todos os instantes (amostras) de todas as séries de um certo intervalo, as funções de agregação mostradas na Tabela 1 serão aplicadas ao vetor resultante. Cada um dos elementos desta tabela será disponibilizado como um atributo para o algoritmo de mineração de dados. A quantidade de dados a ser analisada, ou seja, o intervalo ou janela de tempo que deve ser considerada adequada para a agregação é um parâmetro muito importante para a identificação de anomalias e representa um desafio, pois existem grandezas, como a potência de uma unidade geradora, que não apresentam um comportamento regular como a frequência elétrica (valores em torno de 60Hz), e podem apresentar grandes variações ao longo do tempo e, mesmo assim, estão dentro de uma condição de normalidade. Desta forma, é importante definir o intervalo temporal que representa todo um ciclo que pode conter a anomalia coletiva, do início ao fim, para determinar se uma ou mais amostras realmente representam uma anomalia. Uma segunda equação pode ainda ser definida como filtro para seleção dos dados, permitindo o descarte de dados não relevantes à análise.

Tabela 1 - Funções de Agregação Aplicadas

Atributo	Funções de Agregação Aplicadas
Número de Amostras	Contagem de amostras do intervalo
Média	$\frac{1}{n} \sum_{i=1}^n a_i$
Moda	Valor mais frequente do intervalo
Desvio Padrão	$\sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}}$
Mínimo	Valor mínimo do intervalo
Máximo	Valor máximo do intervalo
Frequência	Contagem do número de transições entre a moda e os outros valores
Duração	Tempo decorrido com valores diferentes da moda
Inclinação	$\frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2}$

Após a definição de todos os parâmetros feita pelo especialista, os dados podem ser coletados das fontes selecionadas e processados considerando-se o período de análise de interesse. Os resultados são armazenados em um banco de dados, chamado de *Data Warehouse*, permitindo o seu uso para o treinamento de algoritmos de classificação, conforme as etapas descritas a seguir.

3.2 Construção do Conjunto de Treinamento Supervisionado

Nesta etapa é realizada a criação do conjunto de treinamento supervisionado de classificação que será utilizado posteriormente na detecção de anomalias, rotulando os dados como "normais" ou "anormais". Como a proposta é permitir a análise de grandes períodos de tempo e de um conjunto amplo de problemas, a rotulagem manual de todos os dados extraídos para cada caso seria extremamente trabalhosa e custosa (avaliação de cada intervalo pelo especialista). Por isso, foi proposta a aplicação de um método não-supervisionado de classificação ao conjunto não rotulado para realizar uma pré-rotulagem. O resultado desta classificação inicial deve ser avaliado pelo especialista, que fará pequenos ajustes apenas em casos que considerar inadequados (9).

Para isso, foi adotado um método de classificação não supervisionado chamado de Rede Neural Replicante RNN. Inspirado em modelos utilizados para compressão de dados de imagem, uma rede neural replicante utiliza uma rede neural *Multi-Layer Perceptron* com o objetivo de reproduzir os dados de entrada na saída (10). A principal característica da Rede Neural Replicante é ser formada pelo mesmo número de neurônios na camada de entrada e de saída para ser treinada para reproduzir a entrada na saída. Além disso, a camada oculta contém um número menor de neurônios do que as camadas externas, provocando o efeito de compressão dos dados, já que para ter os mesmos dados da entrada na saída, será preciso representá-los, de alguma forma, com menos neurônios nas camadas ocultas. Desta maneira, ao treinar esta rede (considerando que a grande maioria dos dados sejam normais) espera-se que ela "aprenda" a reproduzir melhor estes tipos de dados. Assim, a partir do módulo das diferenças entre os valores de entrada e saída pode-se inferir sobre a raridade dos eventos e, por conseguinte, suspeitar da ocorrência de uma anomalia.

3.3 Execução Periódica

Nesta etapa, novos dados são classificados periodicamente, de acordo com o intervalo de agregação definido inicialmente, utilizando classificadores supervisionados, treinados utilizando o conjunto construído na etapa descrita na Seção 3.2. Anomalias detectadas pelos modelos nesta etapa são reportadas na forma de notificações às equipes de operação em tempo real.

Para a classificação de novos dados foi utilizado o algoritmo *k-Nearest Neighbors* (k-NN), um algoritmo supervisionado que determina a classe de uma nova instância com base na classe das instâncias mais próximas (vizinhas). Permite a parametrização de quantas instâncias (k) serão utilizadas para determinar a classe resultante e qual a forma de cálculo da distância que será avaliada. É definido como um classificador *lazy* pois faz a avaliação no momento da classificação, ou seja, calcula a distância dos elementos e faz a busca pelos mais próximos em tempo de execução (11). Vale destacar que os novos dados classificados também podem ter sua classificação revisada pelo especialista caso falsos positivos ou falsos negativos sejam identificados, possibilitando um aperfeiçoamento contínuo do modelo de classificação. A Figura 2 mostra um diagrama com

todas as etapas da metodologia proposta juntamente com a inter-relação entre as configurações e dados registrados em uma etapa com as outras etapas.

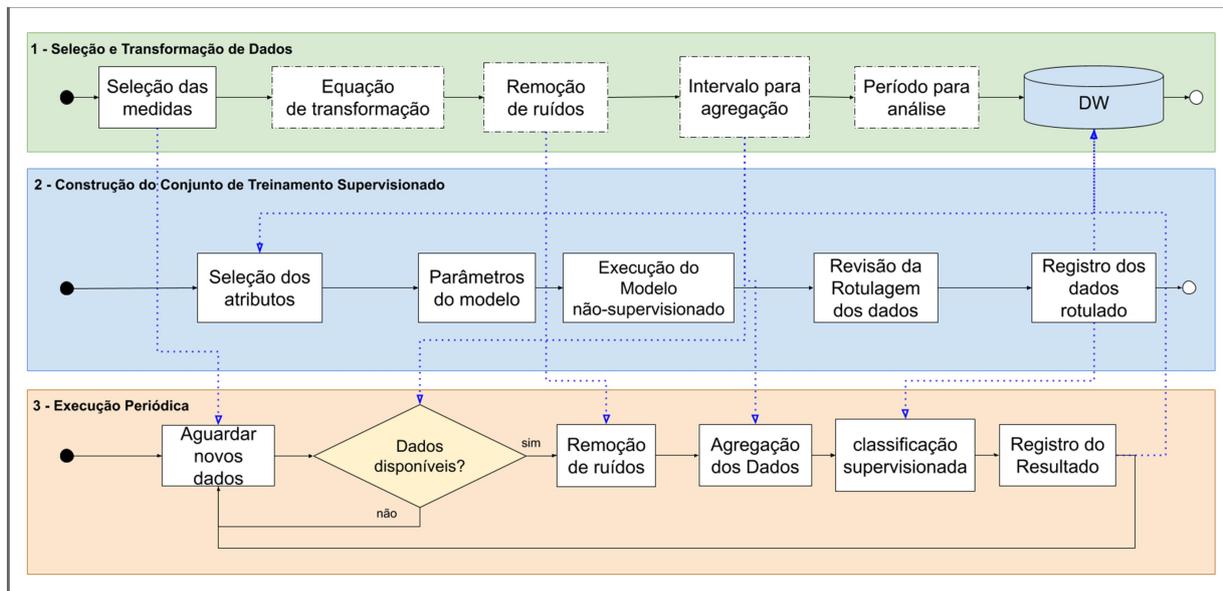


Figura 2 - Diagrama de atividade da metodologia proposta (9).

4.0 - RESULTADOS

Foi desenvolvida uma ferramenta computacional para permitir que o usuário, especialista do sistema elétrico de potência, configure uma extração, realize a criação do conjunto de treinamento supervisionado e, por fim, estabeleça a execução periódica para a identificação de anomalias a partir de novos dados. Esta ferramenta, chamada de Assistente para Supervisão de Grandezas Operativas - ASGO, foi desenvolvida em Java, utilizando um framework de redes neurais, para implementar a RNN, chamado *Neuroph* (12). Os dados pré-processados e os rótulos aplicados pelos algoritmos não supervisionados ou pelos especialistas foram armazenados em um SGBD *PostgreSQL* (*Data Warehouse*). Criou-se também uma biblioteca padronizada de acesso a séries temporais, para acessar não só dados históricos do sistema SCADA como também dados de outros sistemas de apoio, como o sistema de solicitação de serviços e de intervenções da manutenção. A Figura 3 mostra os diferentes componentes da ferramenta ASGO, bem como ilustra as diferentes etapas de interação do especialista com a ferramenta.



Figura 3 - Arquitetura da ferramenta ASGO.

4.1 Anomalias na Pressão do Acumulador Ar/Óleo do Sistema Regulador de Velocidade

Para avaliação da ferramenta, realizou-se um estudo de caso que trata de uma anomalia real no sistema regulador de velocidade das unidades geradoras de Itaipu, a qual pode ser observada através da análise do comportamento de uma grandeza analógica (dados de pressão do acumulador ar/óleo do RV). O óleo pressurizado é utilizado para movimentar as palhetas do distribuidor, controlando o torque hidráulico na turbina através de sua abertura ou fechamento. Assim, é possível manter a velocidade constante do conjunto turbina-gerador durante variações de carga. Normalmente, variações no comportamento de pressão indicam defeito na válvula detectora de pressão. Na Figura 4 é possível observar que a variação da pressão (em verde) sofreu significativa alteração após as 13 horas do dia 22 de novembro de 2016, distorcendo visivelmente seu formato “dente de serra”. Um dos principais desafios da aplicação de técnicas de mineração de dados para solução deste tipo de problema é tentar traduzir essas variações de comportamento das grandezas em atributos numéricos obtidos a partir da agregação de dados (médias, máximos, mínimos, desvio padrão, etc) de modo que não se percam informações importantes.

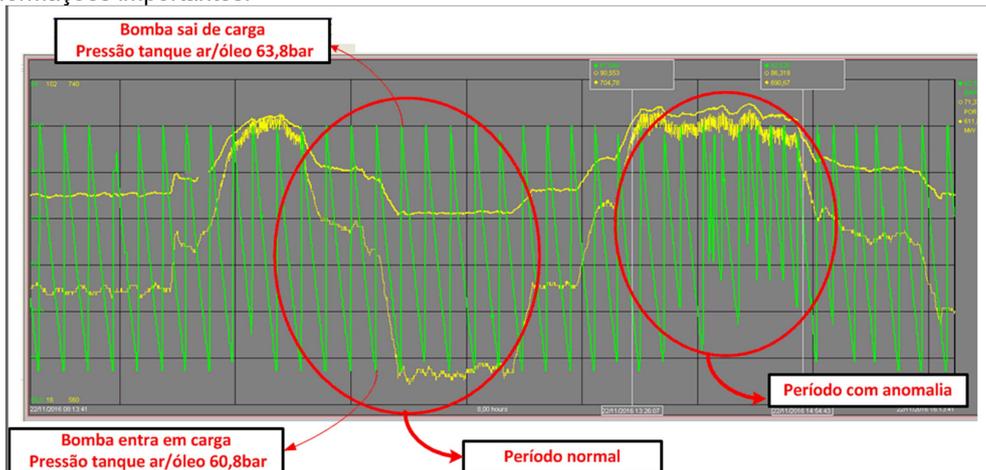


Figura 4 - Anomalia na pressão no acumulador ar/óleo do sistema regulador de velocidade em 22/11/2016. (9)

Para a configuração da ferramenta, foi utilizada uma janela de cálculo de uma hora para a agregação dos dados e feita uma análise usando dados de aproximadamente 1 ano, totalizando 7.582 instâncias (intervalos de agregação), com 28 instâncias rotuladas como anômalas. A Figura 5 apresenta a configuração feita na ferramenta ASGO para extração dos dados e construção do conjunto de treinamento rotulado.

A interface de configuração do ASGO apresenta os seguintes campos e opções:

- Identificação:** Local: GSS, Sistema: CASO 5, Nome: PRESSÃO TANQUE DE ÓLEO DO REGULADOR DE VELOCIDADE - U14, Tipo da Grandeza: Analógica.
- Transformação:** Intervalo de Agregação: A cada 01 hora, Fórmula: =PRESSAO, Filtro: =(ESTADO=1).
- Variáveis:** Tabela com colunas Nome e Tag, contendo ESTADO e PRESSAO.
- Período para Coleta de Dados:** Data Inicial: 01/07/2016 00:00:00, Data Final: 01/07/2017 00:00:00.
- Botões:** Coletar Dados, Exportar, Deletar Dados Coletados, Gráfico das Variáveis, Salvar.

Figura 5 - Tela de configuração do ASGO para a extração descrita no estudo de caso.

Na etapa de construção do conjunto de treinamento supervisionado, foram selecionados os atributos de desvio padrão e frequência após análise gráfica de seus valores pelo especialista. Posteriormente, foi aplicado o

algoritmo não supervisionado (RNN), para realizar a pré-rotulagem dos dados e, em seguida, realizou-se uma revisão para identificar corretamente todos os períodos anômalos.

A Figura 6 ilustra os valores agregados a cada 01 hora dos atributos frequência e desvio padrão da grandeza pressão no acumulador ar/óleo do sistema regulador de velocidade entre o período sob análise. As marcas destacadas em vermelho nos gráficos indicam as instâncias classificadas pela ferramenta como anomalias durante o processo de rotulagem dos dados levando-se em conta simultaneamente os valores dos dois atributos selecionados.

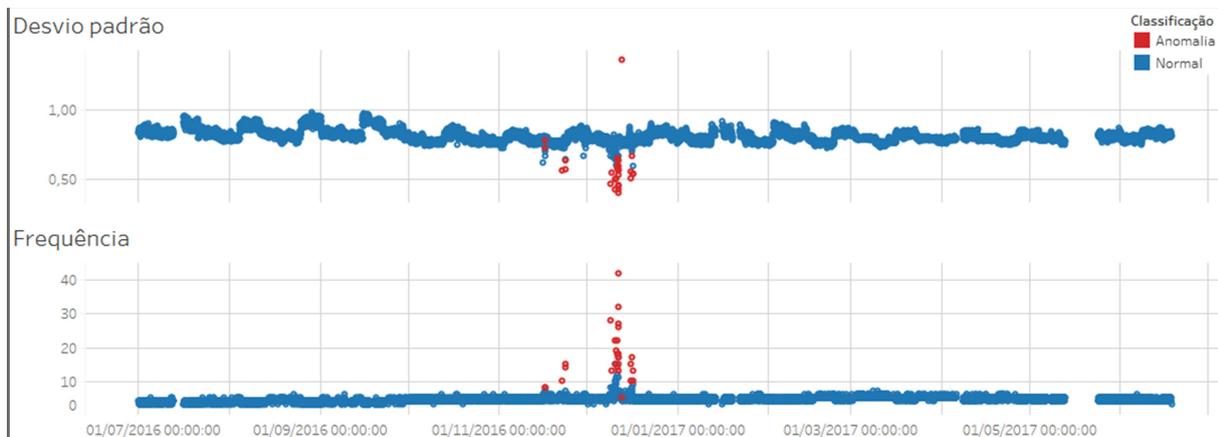
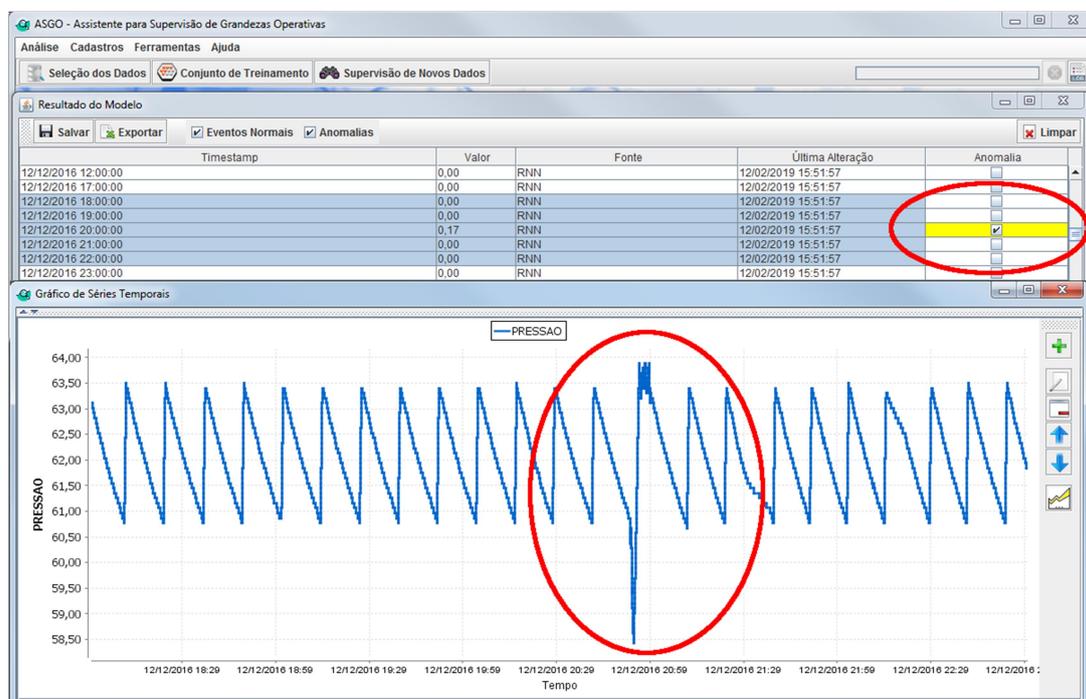


Figura 6 - Valores agregados a cada 01 hora dos atributos Desvio Padrão e Frequência da grandeza pressão.

A avaliação do resultado dos algoritmos foi feita através do método *10-fold cross validation* que apresentou uma precisão de 92,6%¹ e um *recall* de 85,7% na identificação de intervalos anômalos no período de 01/07/2016 a 01/07/2017.

Por fim, a Figura 7, mostra a interface da ferramenta ASGO que permite avaliar em detalhe o comportamento da grandeza original (pressão no acumulador ar/óleo do RV) para instâncias classificadas como anômalas, destacadas em amarelo na tabela de valores (20h dia 12/12/2016). Nesta interface o especialista tem ainda a possibilidade de fazer a correção da classificação realizada automaticamente pela ferramenta.



1

A análise considera apenas o conjunto de anomalias, pois a grande maioria dos dados são normais e uma métrica de todo o conjunto não seria representativa para avaliar o resultado.

Figura 7 - Detalhe da grandeza pressão para instâncias classificadas como anômalas.

5.0 - CONCLUSÕES

Este informe técnico apresentou uma ferramenta que permite a análise de grandezas armazenadas no banco de dados histórico do Scada através da aplicação de técnicas de mineração de dados. Ao longo do desenvolvimento e realização dos testes, constatou-se que a conjunção da experiência de especialistas e da capacidade de processamento computacional de dados promove uma boa relação de custo-benefício. Por um lado, concentram-se no especialista atividades rápidas e complexas, realizadas com base no conhecimento técnico e operacional do equipamento, tais como a seleção de dados, definição da janela de agregação e parametrização básica dos algoritmos de mineração de dados. E por outro, confere-se à ferramenta computacional a tarefa de processar grande quantidade de dados, por meio de estruturas e algoritmos bem estabelecidos. Dificuldades na preparação e transformação dos dados a processar foram contornadas, mormente no que diz respeito a ruídos, dados faltantes e séries com amostras não sincronizadas. Verificou-se que a utilização do algoritmo *KNN* representa uma forma simples de realizar a melhoria contínua do modelo, disponibilizando a opção ao especialista de corrigir uma classificação incorreta e usar esta informação na próxima classificação, resultando em um efeito de aprendizado contínuo do sistema. Além disso, o algoritmo não supervisionado apresentou um desempenho satisfatório e se mostrou adequado para a atividade de pré-rotulagem dos dados. Outro ponto a destacar refere-se à compreensão do especialista acerca do que ocorre em cada etapa do problema em estudo. Ademais, constatou-se que a interface da aplicação, os termos utilizados e a sequência do processo de análise devem ser claras e objetivas no sentido de aproximar a ferramenta à realidade dos especialistas em operação de usinas. Atualmente o ASGO encontra-se em fase final de testes e preparação para implantação em produção em Itaipu.

6.0 - REFERÊNCIAS BIBLIOGRÁFICAS

- (1) Stuart A Boyer. 2018. SCADA supervisory control and data acquisition. The Instrumentation, Systems and Automation Society.
- (2) Jiawei Han, Jian Pei, and Micheline Kamber. 2011. Data mining: concepts and techniques Elsevier.
- (3) Animesh Pacha and Jung-Min Park. 2007. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer networks* 51, 12 (2007), 3448–3470.
- (4) Manish Gupta, Jing Gao, Charu C Aggarwal, and Jiawei Han. 2014. Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and Data Engineering* 26, 9 (2014), 2250–2267.
- (5) SP King, DM King, K Astley, L Tarassenko, P Hayton, and S Utete. 2002. The use of novelty detection techniques for monitoring high-integrity plant. In *Control Applications, 2002. Proceedings of the 2002 International Conference on*, Vol. 1. IEEE, 221–226.
- (6) Stan Salvador, Philip Chan, and John Brodie. 2004. Learning States and Rules for Time Series Anomaly Detection. In *FLAIRS Conference*. 306–311.
- (7) Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2009. Anomaly detection: A survey. *ACM computing surveys (CSUR)* 41, 3 (2009), 15
- (8) Markus Goldstein and Seichi Uchida. 2014. Behavior analysis using unsupervised anomaly detection. In *The 10th Joint Workshop on Machine Perception and Robotics (MPR 2014)*. Online.
- (9) Filipe Ventura Muggiati, Milton Brown Do Coutto Filho e Julio Cesar Stacchini de Souza. 2019. Mineração de Dados Operativos de Usinas Hidrelétricas – Dissertação de Mestrado - Universidade Federal Fluminense.
- (10) Simon Hawkins, Hongxing He, Graham Williams, and Ra Baxter. 2002. Outlier detection using replicator neural networks. *Proceedings of the 4th International Conference on Data Warehousing and Knowledge Discovery (2002)*, 170–180. <https://doi.org/10.1.1.167.4245>
- (11) Thomas Cover and Peter Hart. 1967. Nearest neighbor pattern classification. *IEEE transactions on information theory* 13, 1 (1967), 21–27.
- (12) Sevarac, Z.; Goloskokovic, I.; Tait, J.; Carter-Greaves, L.; Morgan, A.; Steinhauer, V. Neuroph-java neural network framework. *Retrieved in January (2012)*.

7.0 - DADOS BIOGRÁFICOS



FILIFE VENTURA MUGGIATI possui mestrado em Computação pela Universidade Federal Fluminense e graduação em Engenharia Elétrica pela UNIOESTE. Atualmente trabalha em Itaipu no desenvolvimento de sistemas de suporte à operação em tempo real e ferramentas de apoio às atividades de pré e pós-operação.