

Detecção de Novidades em Aparelhos Eletrônicos através do Monitoramento do Consumo de Energia

Thamires C. Luz, Tiago A. Almeida, Fábio L. Verdi

Departamento de Computação (DComp)

Universidade Federal de São Carlos (UFSCar)

Rod. João Leme dos Santos, Km 110, Itinga

18052-780, Sorocaba, São Paulo

thamiluz@gmail.com, talmeida@ufscar.br, verdi@ufscar.br

Resumo—A energia elétrica no Brasil é gerada em sua maioria pelas usinas hidrelétricas. Essas hidrelétricas dependem do volume de água dos reservatórios para a geração de energia e, devido à diminuição dos índices pluviométricos, métodos alternativos de geração com custos mais altos são necessários. O consumo exagerado e o desperdício da energia são fatores que influenciam para este cenário. Um alerta poderia ser enviado ao usuário caso o comportamento padrão fosse alterado. Sendo assim, neste trabalho propomos um sistema que realiza a detecção de novidade no comportamento do consumo de energia elétrica dos aparelhos eletrônicos, de forma que alertas sejam enviados para o usuário. Para detectar a alteração de comportamento do consumo de energia elétrica dos aparelhos eletrônicos foram implementados os métodos Janela Deslizante, *Exponentially Weighted Moving Averages* (EWMA), Agrupamento, Média por Ciclo, Média por Estágio, Distribuição Gaussiana e *Self-Organizing Novelty Detection* (SONDE). Os experimentos demonstraram que os métodos foram eficientes na detecção de novidade em tempo real apresentando taxa de acerto acima de 90% e falso-positivo abaixo de 10%, além de um tempo médio de resposta baixo.

I. INTRODUÇÃO

A energia elétrica no Brasil é gerada predominantemente por usinas hidrelétricas através da água acumulada nos reservatórios. Consequentemente, a capacidade de geração de energia está diretamente relacionada com o volume de chuva necessário para abastecer as represas. Em virtude do aquecimento global e da recente diminuição dos índices pluviométricos, a capacidade de geração das usinas hidrelétricas vem reduzindo significativamente. Para evitar a escassez de energia, o governo precisa recorrer a métodos alternativos e mais caros de geração de energia elétrica, como as usinas termoelétricas. Dessa forma, para suprir o crescente aumento nos custos, foi implantado em 2013, o sistema tarifário diferenciado para ajustar a cobrança do consumo de energia elétrica de acordo com a capacidade dos reservatórios. Com isso, a tarifa padrão é acrescida nos períodos em que há escassez de chuva [1].

Recentemente, a conta de energia elétrica é uma grande preocupação para a maioria da população, pois além do consumo de energia ter aumentado significativamente nos últimos anos, o valor da energia também vem sofrendo grande reajuste, impactando severamente o orçamento de muitas famílias.

Relatórios recentes demonstram que o consumo de energia elétrica *per capita* no Brasil aumentou cerca de 25% somente

nos últimos dez anos [2], principalmente devido à aquisição e dependência de novos aparelhos eletrônicos. Contudo, uma parcela desse aumento pode ser creditada a desperdícios resultantes do mau funcionamento dos aparelhos e de falhas humanas, como o simples ato de esquecer a porta da geladeira aberta, por exemplo. Geralmente, esses eventos demoram ou não são notados pelos usuários, contribuindo para o aumento desnecessário e inconsciente da conta de energia elétrica.

Partindo da hipótese de que cada aparelho possui um padrão de consumo, qualquer alteração nesse padrão pode ser entendido como uma *novidade* no comportamento esperado para tal aparelho. Assim sendo, através do monitoramento e análise contínua do consumo de energia de cada aparelho, é possível detectar alterações comportamentais e alertar os usuários para que alguma providência possa ser tomada rapidamente.

De forma geral, a tarefa de detectar alterações em padrões de comportamento pode ser dividida em três categorias: detecção de novidades, anomalias e *outliers* [3]. A diferença entre elas é que a anomalia e os *outliers* são sempre considerados alterações indesejadas, sendo que o *outlier* pode, ainda, ser um ruído ou falha de leitura que deve ser descartado. Por outro lado, a novidade é considerada como um comportamento que pode ser ou não incorporado ao padrão em determinada situação, ou ainda, ser ou não considerado como algo indesejado dependendo da aplicação. Os termos são relacionados e dependendo do contexto podem ser empregados mais de um dos termos para a definição no cenário pretendido [3].

De acordo com Spinosa [4], a detecção de novidade pode ser interpretada como a tarefa de descobrir situações que diferem do padrão esperado, ou como a alteração abrupta de um comportamento determinado, em que essa diferenciação na maioria dos casos é algo indesejado. Recentemente, esse tipo de problema vem ganhando destaque em pesquisas científicas, principalmente para evitar defeitos de funcionamento que podem ser gerados por pequenas alterações de comportamento. Neste trabalho, a detecção de alteração no padrão de consumo elétrico foi tratada como um problema de detecção de novidades (DN), pois apenas em algumas situações a novidade estará associada a um evento indesejado.

Detectar novidades com base no consumo de energia elétrica de cada aparelho envolve realizar a análise individualizada e contínua do consumo no decorrer do tempo. Em outras

palavras, isso significa que as técnicas empregadas para tal propósito necessitam manipular um fluxo contínuo de dados (FCD) de consumo elétrico para cada aparelho envolvido. Segundo Aggarwal [5], um FCD é uma sequência contínua e infinita de dados que apresenta ou não dependência em relação ao tempo.

Os FCDs podem ser considerados como séries temporais quando apresentam dependência direta com o tempo [6]. Neste trabalho, utilizaremos a nomenclatura fluxo contínuo de dados, entretanto, somente serão analisados os FCD que sejam *Série Temporal*. De acordo com Babcock *et al.* [7], os FCDs possuem as seguintes características:

- Os dados chegam de forma contínua;
- Possuem tamanho ilimitado;
- Uma vez que um exemplo (sinal) é processado, ele não é mais utilizado; e
- Os dados não ficam armazenados em memória.

De maneira geral, a tarefa de detectar alterações de padrões é estudada desde o século XIV. Inicialmente, soluções genéricas foram propostas com base no uso de métodos estatísticos e, no decorrer dos anos, outras técnicas foram propostas para resolver problemas cada vez mais específicos.

Recentemente, grande parte dos métodos especialistas empregados para detectar novidades e anomalias vêm sendo obtidos através de técnicas consolidadas de aprendizado de máquina (AM). De forma ampla, técnicas de AM desempenham atividades de acordo com a experiência adquirida em uma determinada tarefa [8]. Através desses métodos, atualmente, é possível realizar a detecção de novidades em FCD provenientes da coleta de consumo de energia de dispositivos eletrônicos [9].

Nesse cenário, este trabalho propõe e avalia um conjunto consolidado de técnicas estatísticas e de aprendizado de máquina empregados na detecção de novidades no comportamento de aparelhos eletrônicos com base no consumo energético. As detecções são realizadas através do monitoramento e da análise temporal do fluxo contínuo de consumo de energia elétrica de cada aparelho.

A partir das técnicas desenvolvidas, foi projetado um sistema que realiza duas atividades principais: (a) detecção da novidade, que consiste no reconhecimento da alteração do padrão de consumo de energia elétrica de cada aparelho; e (b) envio de um alarme em tempo real, que consiste no disparo de um alerta para o usuário, como por exemplo um e-mail ou SMS, assim que a novidade é detectada.

O restante deste trabalho está organizado da seguinte forma: na Seção II, são apresentados os principais trabalhos relacionados com o tema abordado. Na Seção III, são detalhadas as técnicas implementadas e avaliadas. A metodologia experimental é descrita na Seção IV. Na Seção V, são apresentados os resultados obtidos. Finalmente, as conclusões e direcionamentos para trabalhos futuros são discutidos na Seção VI.

II. TRABALHOS RELACIONADOS

Os estudos para realizar a detecção de novidades vêm crescendo cada vez mais devido à importância do assunto e dos

desafios encontrados na área. Diversos trabalhos apresentam estudos relacionados à detecção de novidade e anomalia [3], [10], [11], entretanto essa seção apresenta de forma resumida apenas os principais trabalhos que se relacionam à detecção de novidade em FCD.

Um dos principais métodos para realizar previsões em séries temporais é a Janela Deslizante (JD), ou Média Deslizante. Mesmo com a diversidade de métodos disponíveis para realizar tal tarefa, a Janela Deslizante continua sendo uma das mais utilizadas [12]. Neste contexto, Hansun [12] apresenta comparações dos métodos derivados da JD, demonstrando que a JD e suas variações apresentam alta taxa de acerto para a previsão do valor esperado. Wang *et al.* [13] também demonstra resultados obtidos com a JD.

Vários trabalhos destacam os métodos baseados em Rede Neural Artificial (RNA) como sendo os principais métodos utilizados para a detecção de anomalias [14]. Para realizar a DN em FCD, Albertini e de Mello [15] criaram o método *Self-Organizing Novelty Detection* (SONDE). A detecção de novidade é realizada quando um novo comportamento é apresentado para o método e nenhum neurônio pode representá-lo. Aguayo e Barreto [16] também utilizaram a RNA para analisar as detecções de novidades e anomalias em séries temporais. Os autores apresentam o projeto *Detection of Anomalies and Novelties in Time sEries with self-organizing networks* (DANTE) em que se realiza a comparação de diferentes métodos baseados em RNA e *Self-Organizing Map* para realizar a detecção de novidade.

Outro método bastante utilizado na literatura para detecção de novidades em FCD é o agrupamento. Estudos demonstram que os métodos de agrupamento apresentam bons resultados para solucionar problemas em série temporal e comparações de vários métodos são realizadas [17], [18], [19]. Além disso, outros métodos baseados em agrupamento são propostos para a detecção de novidade como o método *OnLine Novelty and Drift Detection Algorithm* (OLINDDA) [4] e *DiscrETE Cosine Transform based NOvelty and Drift detection* (DETECTNOD) [20].

De modo geral, os métodos citados acima foram propostos para cenários genéricos. Entretanto, trabalhos específicos para realizar a detecção de novidades para o consumo elétrico são encontrados também. Para realizar a detecção de novidade nesse cenário, Chou e Telaga [21] partem da ideia que a detecção de novidade pode ser realizada caso o valor previsto para o consumo elétrico apresente uma grande diferença do valor real. Para realizar essa tarefa, os autores realizam comparações do método *Auto-Regressive Integrated Moving Average* (ARIMA) clássico com uma modificação do método para ser rede neural (NNARIMA).

Outro trabalho que utiliza o cenário de consumo de energia elétrica, é o apresentado por Filho *et al.* [22]. Nele, é proposto o método *Novelty Detection Power Meter* (NodePM) para detectar as novidades presentes no consumo de energia coletado dos aparelhos eletrônicos pelos sensores de rede sem fio em *smart grid*. Os autores compararam os resultados do método proposto com o método SONDE citado anteriormente [15].

E com isso, constataram que embora o NodePM apresente melhores resultados com uma base de dados menor (de uma semana), o SONDE apresenta melhores resultados em bases de dados maiores (com dados de duas ou mais semanas) [22].

Como é possível observar, há uma grande variedade de métodos que são propostos para a DN. Porém, ainda são encontrados desafios de acordo com o cenário utilizado, principalmente quando relacionados aos FCDs. Dessa maneira, para a detecção de novidade ser realizada em tempo real para o consumo de energia elétrica foram implementados os principais métodos mais citados pela literatura. Com isso, pode-se sumarizar a contribuição deste trabalho em três principais aspectos: (a) realização da detecção de novidades no cenário de consumo de energia elétrica; (b) aplicação e comparação de sete métodos de AM para este cenário; e (c) envio de alerta para o usuário quando detectada uma novidade em tempo real.

III. MÉTODOS

De acordo com a revisão da literatura foram implementados os principais métodos utilizados para a detecção de novidade: Janela Deslizante, EWMA, Agrupamento, SONDE, Distribuição Gaussiana, Média por Ciclo e Média por Estágio, sendo os dois últimos métodos estatísticos.

Os aparelhos eletrônicos foram classificados de acordo com o tipo de consumo de energia elétrica em duas categorias: (a) cíclicos, aparelhos que possuem comportamentos diferentes que se repetem ao longo do tempo, como por exemplo a geladeira e máquina de lavar roupa; e (b) não cíclicos, aparelhos que seguem sempre um mesmo consumo ao longo do tempo, como por exemplo o computador e televisão. A forma de detecção de novidade para cada categoria é feita de maneira diferente e, portanto, os métodos implementados seguem essas categorias.

Nas sub-seções a seguir é detalhado cada método implementado.

1) Cíclicos:

a) Média por Ciclo:

O método nomeado de Média por Ciclo (MC) foi criado com base na frequência de tempo em que ocorrem os ciclos de consumo elétrico dos aparelhos eletrônicos. Na fase de treinamento o método identifica os ciclos de consumo de energia elétrica de acordo com a mudança de estágio de consumo e um ciclo é definido quando esse consumo volta a ter novamente o mesmo comportamento inicial. Após a identificação do ciclo, o método calcula qual foi a duração do ciclo completo, além da média dos valores de consumo. A DN ocorre quando um ciclo completo possui um tempo de permanência inferior ao tempo de duração padrão acrescido da margem de aceite, ou quando a média de valores do consumo do ciclo seja divergente do padrão de média encontrada para o ciclo na fase de treinamento. A Figura 1 demonstra que o ciclo identificado como novidade apresentou a duração t menor que n (tempo padrão acrescido da margem de aceite).

b) Média por Estágio:

O método Média por Estágio (ME) é similar ao Média por Ciclo, porém verifica a duração de tempo (t) e a média de

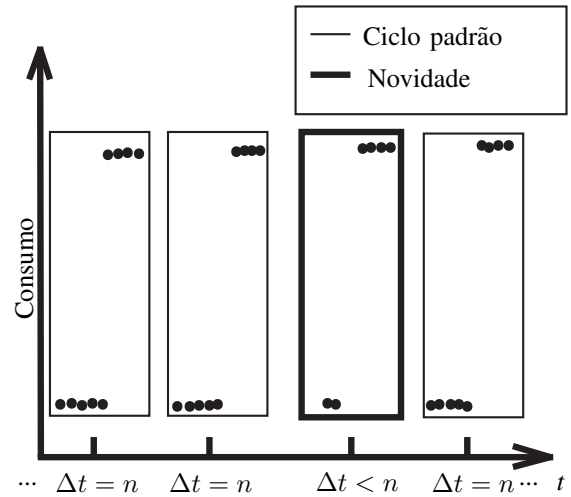


Figura 1. Exemplificação do método Média por Ciclo em que o ciclo padrão apresenta duração (t) igual ao tempo padrão acrescido da margem de aceite (n). O ciclo é identificado como novidade quando a duração é inferior a n .

consumo de energia para cada estágio de consumo ao invés de um ciclo completo. Na fase de treinamento é calculada a duração de tempo de cada estágio. Para identificar as novidades, a cada vez que o consumo de energia é alterado para um estágio diferente, o tempo de permanência do estágio é comparado ao tempo padrão de permanência calculado no treinamento acrescido da margem de aceite para estágios que tenham a mesma média de consumo. A Figura 2 demonstra que o estágio identificado como novidade apresentou duração (t) inferior a duração de tempo padrão acrescido da margem de aceite (n).

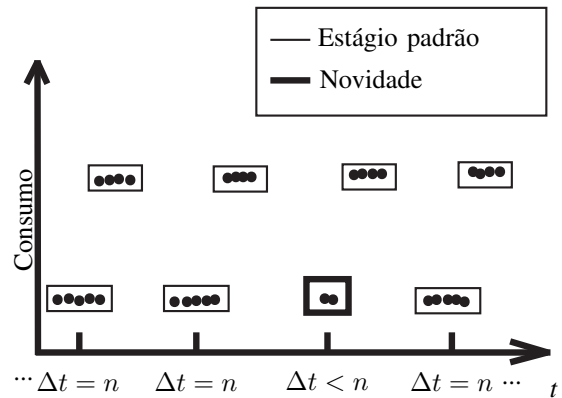


Figura 2. Exemplificação do método Média por Estágio em que o estágio padrão de consumo baixo apresenta duração (t) igual ao tempo padrão acrescido da margem de aceite (n). O estágio de consumo baixo é identificado como novidade quando a duração é inferior a n .

2) Não cíclicos:

a) Distribuição Gaussiana:

A Distribuição Gaussiana (GAUSS), também conhecida como Distribuição Normal, é uma das mais utilizadas em estatística e segue a ideia de que os dados se distribuem levando em conta a média e o desvio padrão, de forma que

a média se refere ao centro da distribuição e o desvio padrão ao espalhamento dos dados [23]. Este método é bastante utilizado para detecção de anomalias [24]. Através do cálculo de probabilidade de distribuição do dado é possível determinar uma novidade caso essa probabilidade seja inferior a um limiar estipulado. A Figura 3 apresenta a margem de distribuição calculada. Os pontos que ficam fora da distribuição calculada são considerados novidades.

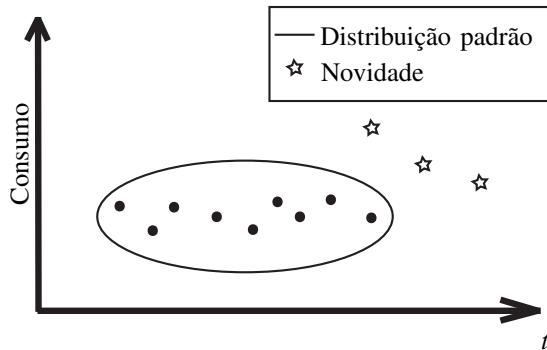


Figura 3. Exemplo do método GAUSS em que a distribuição dos dados devem estar dentro da margem estipulada. Os pontos que estão com a distribuição fora da margem calculada são considerados novidades.

b) *Exponentially Weighted Moving Averages:*

O método *Exponentially Weighted Moving Averages* (EWMA) realiza a estimativa do próximo ponto com base em n amostras passadas (pesos), sendo que os dados mais recentes têm mais influência na previsão [12]. A detecção de novidade acontece quando o previsto é divergente do real. A Figura 4 apresenta a variação aceita da previsão realizada pelo método. Os pontos que ultrapassam os limites estipulados são considerados novidades.

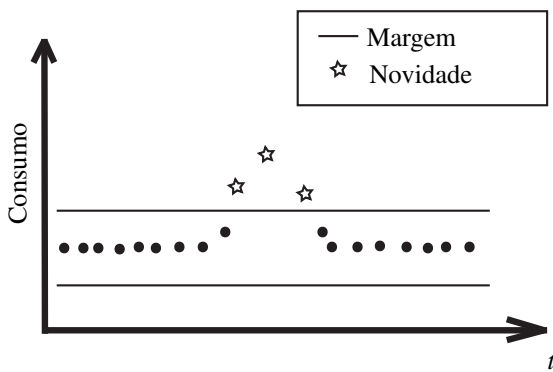


Figura 4. Exemplo do método EWMA em que a variação da previsão realizada para o consumo no próximo período deve estar dentro da margem de aceite estipulada. Os pontos que ultrapassam a margem estipulada são considerados como novidade.

c) *Janela Deslizante:*

O método Janela Deslizante (JD) é popular na tarefa de prever dados que mudam ao longo do tempo [12]. Esse método, também chamado de *Simple Moving Average* (SMA), é baseado em *Moving Average* [25]. O método segue a ideia

que dada uma série de tempo T sejam selecionados alguns dados dentro de uma janela de tamanho n , que se desloca em k tempos dos dados apresentados. De acordo com essa janela selecionada, a média é calculada, e os pontos são considerados novidades quando ultrapassam uma margem definida. A Figura 5 demonstra que a novidade é identificada quando a média dos consumos da janela selecionada ultrapassa a margem estipulada.

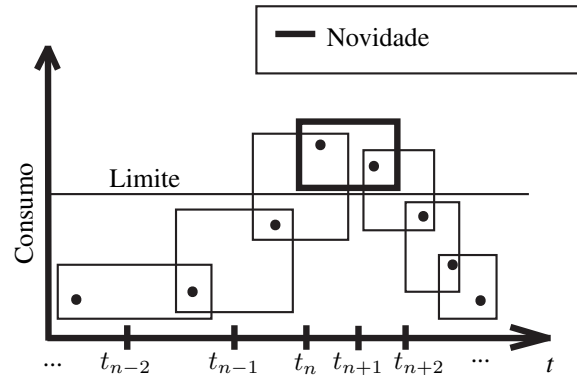


Figura 5. Exemplo do método de Janela Deslizante em que a novidade é identificada quando a média de consumo da janela ultrapassa o limite da margem de aceite estipulada.

d) *Self-Organizing Novelty Detection:*

Conforme já mencionado, o *Self-Organizing Novelty Detection* (SONDE) é um método baseado em RNA. O método cria neurônios que representam padrões apresentados. Para isto, leva-se em conta a similaridade e a proximidade dos dados. Caso estas características estejam dentro de um raio calculado, esse padrão é representado por um mesmo neurônio. Dessa forma, os dados com mesmo comportamento são representados por um mesmo neurônio, e quando um novo padrão é apresentado e o método não identifica nenhum neurônio para representar este novo padrão, um novo neurônio é criado, caracterizando a novidade [15]. A Figura 6 demonstra que os pontos de novidades não são representados por neurônios existentes.

3) *Genéricos:*

a) *Agrupamento:*

Dentre as principais categorias existentes para o agrupamento (AG) [17], neste trabalho implementou-se um método de agrupamento baseado em densidade. Este método agrupa os dados que sejam semelhantes entre si como os demais métodos de agrupamento [17], entretanto se diferencia dos outros métodos de agrupamentos por criar vários grupos que teriam o mesmo centro se não levasse em conta o tempo. Na fase de treinamento, os grupos são definidos de acordo com a similaridade de consumo e proximidade de tempo, isto é, quando uma nova amostra apresenta a mesma média de consumo das amostras passadas e seu tempo se distancia por alguns segundos do tempo das amostras passadas, esta nova amostra é agrupada no mesmo grupo das amostras passadas. Um novo grupo é formado caso o novo consumo seja muito maior ou muito menor que os consumos anteriores ou quando

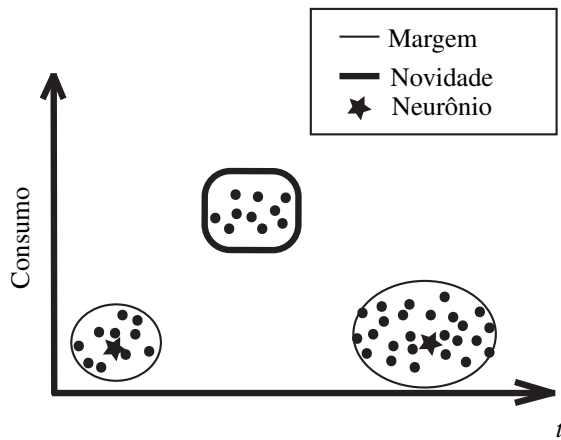


Figura 6. Exemplo de neurônios de comportamento padrão do método SONDE. Os pontos destacados como novidades não são representados pelos neurônios existentes devido a estarem fora do raio de aceite dos neurônios.

a amostra apresenta a mesma média de consumo de um grupo já existente, porém sua proximidade em relação ao tempo é maior que alguns minutos. Após definidos os grupos, é calculada a densidade e o centro de cada grupo. Grupos com o mesmo centro de consumo devem ter uma densidade similar. A DN ocorrerá quando a densidade calculada para um novo grupo apresentar uma grande diferença com relação as densidades dos grupos com centro de consumo similares dos grupos criados na fase de treinamento ou não há grupos com a mesma média de centróide. A Figura 7 demonstra a novidade detectada para um grupo que apresenta densidade inferior a densidade padrão calculada para a mesma média de consumo.

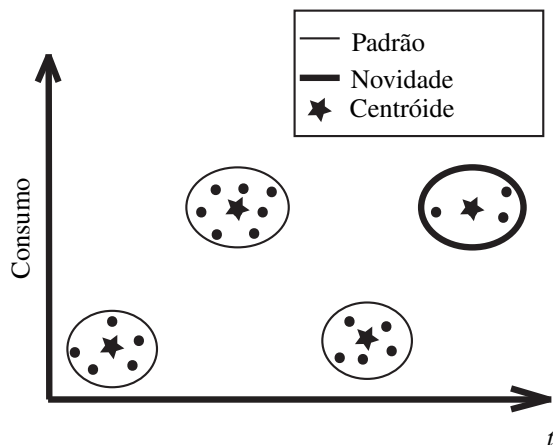


Figura 7. No método de AG os grupos com mesma média de consumo devem ter densidades próximas. A novidade é detectada quando o grupo apresenta uma densidade muito inferior ao padrão calculado para a mesma média de consumo.

IV. METODOLOGIA EXPERIMENTAL

O objetivo geral desse trabalho é identificar e enviar alertas das novidades encontradas no consumo de energia elétrica dos aparelhos eletrônicos para o usuário em tempo real. Para isso,

Tabela I
DETALHAMENTO DAS COLETAS REALIZADAS POR EQUIPAMENTO ELETRÔNICO.

Equipamento	Categoria	Tempo	Tipo de novidade
Geladeira Electrolux R280 110V	Cíclico	3 meses	Porta aberta por cerca de 1 hora
Computador Macbook Pro 13 Late 2011	Não Cíclico	2 dias	Sobrecarga de processamento
Eletroforese Capilar Agilent Technologies 7100 de 110v [26]	Não Cíclico	2 dias	Criada sinteticamente

o desenvolvimento dividiu-se em duas tarefas principais: (a) coletar o consumo de energia elétrica de aparelhos eletrônicos, e (b) implementar algoritmos capazes de processar dados contínuos em larga escala para detecção de novidade em aparelhos eletrônicos. A seguir são apresentados os detalhes da coleta de dados, os parâmetros e medidas de desempenho utilizadas.

A. Coleta de dados

Para realizar a coleta de dados montou-se um ambiente com um wattímetro Kill-a-Watt da empresa P3 ¹, em que fundiu-se ao wattímetro uma placa transmissora *ZigBee* (dispositivo de rede sem fio), responsável pelo envio do consumo de watts coletados. Para receber os dados, configurou-se um mini-servidor local com um mini-computador *Raspberry Pi*. A este mini-servidor conectou-se uma placa *ZigBee* em sua porta USB para agir como receptora. Dessa forma, os dados são enviados para o mini-servidor e armazenados em uma planilha CSV. Como casos de uso para este trabalho, foram coletados o consumo energético de uma geladeira, um computador e um eletroforese capilar. A Tabela I apresenta os detalhes de cada equipamento, o tempo de coleta, a categoria e o tipo de novidade presente nas simulações.

B. Parâmetros

Cada método implementado possui um conjunto de parâmetros que podem ser ajustados pelo usuário. Esses parâmetros influenciam em quais situações a DN deverá ocorrer. Para selecionar os melhores conjuntos de parâmetros para cada aparelho utilizado foram realizados vários experimentos de forma que os parâmetros foram selecionados empiricamente, pois cada equipamento eletrônico apresenta um padrão próprio de consumo de energia elétrica sendo necessário a parametrização para cada situação. Os melhores parâmetros selecionados para cada método foram:

- Cíclico
 - 1) Média por Ciclo
 - Margem de aceite: 15%
 - 2) Média por Estágio
 - Margem de aceite: 20%
- Não Cíclico

¹P3 - <http://www.p3international.com/products/p4400.html> - Acessado em: jan/2015.

- 1) Distribuição Gaussiana
 - Margem de aceite: 1 Desvio Padrão
 - Taxa de aceite: 0,00001
 - 2) EWMA
 - Margem de aceite: 1 Desvio Padrão
 - Intervalo: 5 minutos
 - Peso: 5
 - 3) Janela Deslizante
 - Tamanho da Janela: 10 minutos
 - Deslocamento: 1 minuto
 - Margem de aceite: 1 Desvio Padrão
 - 4) SONDE
 - Similaridade mínima: 0,001
- Genérico
- 1) Agrupamento
 - Margem de aceite: 30%

C. Medidas de Desempenho

Para determinar quais métodos obtiveram melhores resultados foram utilizadas as métricas de (a) Novidade Detectada (ND), que contabiliza quantas vezes o método identificou corretamente o que realmente é novidade; e (b) Falso Alarme (FA), identificado como quantas vezes o método apontou como novidade, mas na verdade é um ponto padrão. Além disso, foi levado em conta o Tempo Médio de Alarme (TMA), que corresponde ao tempo médio que o método precisa para realizar a detecção de novidade e enviar o alarme para o usuário.

D. Procedimentos

Todos os métodos foram implementados em Python. As experimentações foram realizadas com duas bases de dados: (a) treinamento e (b) teste. A base de treinamento contém somente os dados de consumo de energia elétrica padrão que foram coletados em um ambiente controlado para cada dispositivo eletrônico, e é apresentada para que o método aprenda qual é o comportamento padrão esperado. Para verificar a detecção de novidades dos métodos, utilizaram-se as bases de testes. Estas bases de testes apresentam, além do comportamento padrão de energia elétrica do equipamento eletrônico, trechos com as novidades geradas pelas simulações com as alterações do comportamento padrão do usuário.

Para se obter melhores resultados, os dados para os métodos não cíclicos foram normalizados por padronização, isto é, os dados foram escalados com média zero e desvio padrão um. Entretanto, como a normalização deixa a dispersão dos dados menor, a normalização não devem ser empregada para os métodos cíclicos que precisam da dispersão para determinar os ciclos.

Após realizar os experimentos com as bases de dados, os métodos foram expostos ao FCD de uma coleta em tempo real para determinar se os alarmes seriam enviados corretamente. Para analisar esse comportamento, os métodos foram implantados em um protótipo inicial, porém, funcional. Nesse protótipo, o mini-computador Raspberry Pi recebe os dados

coletados via interface sem-fio *Zigbee* e os armazena em um banco de dados SQLite. Os métodos de detecção de novidade executam continuamente neste computador analisando os FCDs. Ao detectarem uma novidade, um e-mail é enviado para o usuário alertando-o sobre a mudança de comportamento. A Figura 8 demonstra a arquitetura utilizada desde a coleta até o envio do alerta.

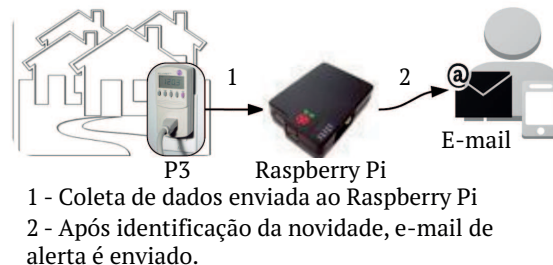


Figura 8. Arquitetura da coleta de consumo elétrico. O walltímetro P3 envia os dados ao Raspberry Pi. Os métodos implementados e executando continuamente identificam as novidades existentes e realizam o envio do e-mail de alerta para o usuário.

V. RESULTADOS

A seguir são apresentados os resultados obtidos com as melhores parametrizações dos métodos para cada equipamento eletrônico.

A. Aparelhos cíclicos

As coletas de consumo de energia na geladeira realizadas demonstraram que o padrão de ciclo de consumo de energia da geladeira é de em média 20 minutos no estágio de consumo alto (média de 160 watts), e 40 minutos no estágio de consumo baixo (média de 90 watts). Considerando esse padrão, as novidades presentes no arquivo de teste tem uma permanência no estágio de consumo baixo próximo de 20 minutos. O parâmetro “Margem de aceite” para os métodos Agrupamento, Média por Ciclo e Média por Estágio é relacionado ao percentual de aceite que o método deve considerar (acima ou abaixo) do valor padrão calculado na fase de treinamento como margem de tolerância.

A Tabela II apresenta os resultados dos experimentos para o cenário cíclico. São apresentados os percentuais da ND, FA e TMA. O método MC apresentou melhores resultados na detecção correta das novidades, pois não apresentou nenhum falso-positivo. Entretanto, o MC é o método que mais demora para enviar o alerta para o usuário nesse cenário, devido a detecção da novidade somente ser realizada após a conclusão do ciclo completo. Dessa forma, o método de AG e ME apresentam o tempo de envio mais eficiente, e apesar de apresentarem falsos-positivos, a taxa de FA é considerada baixa, sendo recomendado a utilização dos métodos AG e ME para o cenário cíclico.

B. Aparelhos não cíclicos

As coletas de consumo de energia no computador demonstram que, em média, o uso do computador segue um mesmo

Tabela II
RESULTADOS SUMARIZADOS DOS MÉTODOS ANALISADOS PARA O CENÁRIO CÍCLICO.

Aparelho	Método	ND (%)	FA (%)	TMA (min)
Geladeira	AG	100	3	20
	MC	100	0	40
	ME	100	3	20

padrão de consumo de energia, e quando mais processamento é necessário o consumo de energia elétrica aumenta. A média do consumo padrão calculada apresentou um resultado próximo de 95 watts. Portanto, para ser considerado uma novidade, a média a ser verificada deve ser superior a 95 watts mais uma porcentagem definida como um parâmetro.

O consumo de energia elétrica do Eletroforese Capilar oscila entre 100 watts e 350 watts, apresentando uma média de consumo próximo de 200 watts. Não foi possível realizar a alteração do consumo padrão para esse aparelho, devido a seu uso constante pelo laboratório em que se encontra. Portanto, definiu-se que as novidades para este aparelho seriam picos de consumo de energia, como o caso do computador. Com isto, estipulou-se que a novidade seria de picos próximos de 500 watts. Este é o único equipamento com novidades geradas sinteticamente.

Na Tabela III são apresentados os resultados sumarizados para o cenário não cíclico. Todos os métodos apresentaram altas taxas de ND (acima de 80%), e baixas taxas de FA (abaixo de 5%). Porém, os métodos GAUSS, EWMA e SONDE se destacam pelo baixo índice de TMA para o envio de alerta, pois esses métodos realizam a detecção de novidade a cada ponto e, portanto, não apresentam nenhum tipo de atraso para a emissão do alerta. O método JD depende do tamanho da janela para a emissão de alerta, uma vez que a detecção somente será realizada quando a média da janela no tempo atual for superior a margem estipulada. Nos exemplos do computador e do eletroforese capilar, foram utilizadas janelas de 2 minutos, sendo o alerta enviado após esse período. Já o método AG somente realizará o envio do alerta quando o evento novidade terminar, pois sua detecção somente é realizada a cada mudança de grupos identificados e, portanto, seu TMA varia conforme o tempo de duração do evento novidade. No exemplo do consumo do computador e do eletroforese capilar, o TMA foi de 120 segundos devido a novidade durar esse período, sendo o alerta apenas enviado ao fim do evento novidade.

VI. CONCLUSÃO

A economia de energia elétrica e atitudes que auxiliem para evitar os desperdícios são de suma importância nos dias atuais. A detecção de novidade para o consumo energético dos aparelhos eletrônicos pode contribuir para a economia de energia, evitando desperdícios desnecessários pela intervenção do usuário após o alerta de novidade recebido. A tarefa de detectar novidades vem sendo cada vez mais alvo de estudos e ainda são encontrados desafios dependendo do cenário

Tabela III
RESULTADOS SUMARIZADOS DOS MÉTODOS ANALISADOS PARA O CENÁRIO NÃO CÍCLICO.

Aparelho	Método	ND (%)	FA (%)	TMA (seg)
Comp.	AG	100	0	120
	GAUSS	100	0	1
	EWMA	100	4	1
	JD	80	0	60
	SONDE	100	0	1
Eletrof. Cap.	AG	100	3	120
	GAUSS	100	3	1
	EWMA	100	4	1
	JD	100	5	60
	SONDE	100	0	1

aplicado. Diversos métodos são propostos para solucionar problemas cada vez mais complexos.

Para realizar a detecção de novidades no fluxo contínuo de dados provenientes da coleta de consumo energético dos aparelhos eletrônicos foram implementados sete métodos conceituados pela literatura. Tais métodos são citados na literatura como sendo os principais para a detecção de novidade e, para esse cenário, a eficiência dos métodos foi comprovada com os altos índices de acertos. Os resultados demonstraram que o Média por Ciclo, Média por Estágio e Agrupamento para os aparelhos eletrônicos que tenham cenários cíclicos apresentam altos índices de acertos (ND 100%), e poucos falso-positivos (menos de 5% de FA), sendo promissores para a DN nessa categoria. Entretanto, a utilização do método de Agrupamento deve ser realizada com cautela, pois, embora o Agrupamento apresente alta taxa de acerto, o TMA é maior devido a sua detecção somente ocorrer ao fim do evento novidade, não contribuindo de imediato para o alerta ao usuário.

Para a DN em aparelhos eletrônicos não cíclicos os métodos também apresentaram alta taxa de acerto (taxa de ND de 100%), e apresentaram índices de falso-positivos baixos (menos de 5% de FA). Embora todos os métodos para o cenário não cíclico apresentem boas taxas de detecção, os métodos Distribuição Gaussiana, SONDE e EWMA apresentam os menores TMAs, sendo os mais eficientes para a emissão de alerta ao usuário.

Como trabalhos futuros pretende-se disponibilizar os dados em uma nuvem para proporcionar a consulta remota do histórico de consumo de energia elétrica do usuário. Além disso, é importante que o usuário tenha acesso ao sistema através de uma interface gráfica amigável de forma a realizar configurações necessárias assim como proporcionar a visualização do histórico de consumo de energia. Finalmente, os algoritmos implementados neste trabalho podem ser usados em outros cenários cujas novidades precisam ser detectadas tais como cenários que envolvem monitoramento humano, ambiental, trânsito e riscos de saúde.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro da CAPES.

REFERÊNCIAS

- [1] “Bandeiras Tarifárias - Agência Nacional de Energia Elétrica,” <http://www.aneel.gov.br/area.cfm?idArea=758>, acessado em: jan/2015.
- [2] “The World Bank,” <http://data.worldbank.org/>, acessado em: jun/2015.
- [3] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Comput. Surv.*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009. [Online]. Available: <http://doi.acm.org/10.1145/1541880.1541882>
- [4] E. J. Spinosa, “Detecção de novidade com aplicação a fluxos contínuos de dados,” *Tese de Defesa de Doutorado. Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (ICMC-USP)*, 2008.
- [5] C. C. Aggarwal, *Data streams: Models and algorithms (advances in database systems)*. Springer, 2006.
- [6] E. R. d. F. Paiva, “Detecção de novidade em fluxos contínuos de dados multiclasse,” Ph.D. dissertation, Tese de Defesa de Doutorado. Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (ICMC-USP), 2014.
- [7] B. Babcock, S. Babu, M. Datar, R. Motwani, and J. Widom, “Models and issues in data stream systems,” in *Proceedings of the Twenty-first ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, ser. PODS '02. New York, NY, USA: ACM, 2002, pp. 1–16. [Online]. Available: <http://doi.acm.org/10.1145/543613.543615>
- [8] T. M. Mitchell, *Machine learning*, 1st ed. McGraw-Hill, 1997.
- [9] J. Gama and M. M. Gaber, *Learning from Data Streams: Processing Techniques in Sensor Networks*. Springer-Verlag New York Inc, 2007.
- [10] X. Ding, Y. Li, A. Belatreche, and L. P. Maguire, “An experimental evaluation of novelty detection methods,” *Neurocomputing*, vol. 135, no. 0, pp. 313 – 327, 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0925231213011314>
- [11] M. A. F. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, “A review of novelty detection,” *Signal Processing*, vol. 99, no. 0, pp. 215 – 249, 2014.
- [12] S. Hansun, “A new approach of moving average method in time series analysis,” in *New Media Studies (CoNMedia), 2013 Conference on*, Nov 2013, pp. 1–4.
- [13] Y. Wang, C. Zhao, and C. Chang, “Anomaly detection using sliding causal windows,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, July 2014, pp. 4600–4603.
- [14] S. Markou, M.; Singh, “Novelty detection: a review—part 2: neural network based approaches,” *Signal Processing*, vol. 88, no. 2499–2521, 2003.
- [15] M. K. Albertini and R. F. de Mello, “A self-organizing neural network for detecting novelties,” in *Proceedings of the 2007 ACM Symposium on Applied Computing*, ser. SAC '07. New York, NY, USA: ACM, 2007, pp. 462–466. [Online]. Available: <http://doi.acm.org/10.1145/1244002.1244110>
- [16] L. Aguayo and G. A. Barreto, “Novelty detection in time series through self-organizing networks: An empirical evaluation of two different paradigms,” *Neural Networks, Brazilian Symposium on*, vol. 0, pp. 129–134, 2008.
- [17] T. W. Liao, “Clustering of time series data—a survey,” *Pattern Recognition*, vol. 38, no. 11, pp. 1857 – 1874, 2005.
- [18] H. Nguyen, Y. Woon, and W. Ng, “A survey on data stream clustering and classification,” *Knowledge and Information Systems*, pp. 1–35, 2014.
- [19] J. A. Silva, E. R. Faria, R. C. Barros, E. R. Hruschka, A. C. P. L. F. d. Carvalho, and J. Gama, “Data stream clustering: A survey,” *ACM Comput. Surv.*, vol. 46, no. 1, pp. 13:1–13:31, Jul. 2013.
- [20] M. Hayat and M. Hashemi, “A dct based approach for detecting novelty and concept drift in data streams,” in *Soft Computing and Pattern Recognition (SoCPar), 2010 International Conference of*, 2010, pp. 373–378.
- [21] J. Chou and A. Telaga, “Real-time detection of anomalous power consumption,” *Renewable and Sustainable Energy Reviews*, vol. 33, no. 0, pp. 400 – 411, 2014.
- [22] G. P. R. Filho, J. Ueyama, L. A. Villas, A. R. Pinto, and S. Seraphini, “Nodepm: Um sistema de monitoramento remoto do consumo de energia elétrica via redes de sensores sem fio,” *Proceedings of: 31º Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos – SBRC*, 2013.
- [23] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed. John Wiley & Sons, 1968.
- [24] K. Limthong, K. Fukuda, Y. Ji, and S. Yamada, “Unsupervised learning model for real-time anomaly detection in computer networks,” vol. E97-D, Agosto 2014, pp. 2084–2094.
- [25] W. B. Jun and Z. Ying, “A survey and performance evaluation on sliding window for data stream,” in *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*, Maio 2011, pp. 654–657.
- [26] “Capillary Electrophoresis - Agilent Technologies,” <http://www.chem.agilent.com/en-US/products-services/Instruments-Systems/Automated-Electrophoresis/7100-CE-System/Pages/default.aspx>, acessado em: jan/2015.