

NodePM: Um Sistema de Monitoramento Remoto do Consumo de Energia Elétrica via Redes de Sensores sem Fio

Geraldo P. R. Filho¹, Jó Ueyama¹, Leandro A. Villas², A. R. Pinto³, Sibeliuss Seraphini¹

¹Instituto de Ciência Matemáticas e de Computação – ICMC
Universidade de São Paulo – USP
13566 – 590 – São Carlos – SP – Brasil

²Instituto de Computação – Universidade Estadual de Campinas – UNICAMP
Campinas – SP – Brasil

³Universidade Estadual Paulista – UNESP
São José do Rio Preto – SP – Brasil

{geralop, joueyama}@icmc.usp.br, leandro@ic.unicamp.br,
arpinto@ibilce.unesp.br, sibeliuss@grad.icmc.usp.br

Abstract. *This article proposes NodePM, a autonomous method for novelty detection in electronic equipments monitored by a smart grid. Considering the entropy of each equipment, which is calculated based on a markov chain model, the method uses a machine learning algorithm for the detection of novelties. NodePM is integrated to a energy consumption monitoring platform composed of a WSN and a cloud-based application. The results of the experiments showed the efficiency of the method in detecting novelties in electronic equipment which has made their use viable in our platform.*

Resumo. *Neste artigo, é proposto um método inteligente, nomeado de NodePM, para detectar novidades em equipamentos eletrônicos monitorados por uma smart grid. Considerando a entropia de cada equipamento monitorado, a qual é calculada com base em um modelo de cadeia de markov, o método proposto detecta as novidades por meio de um algoritmo de aprendizado de máquina. Além disso, o NodePM é integrado a uma plataforma de monitoramento remoto de consumo de energia, que consiste de uma RSSF associada a uma aplicação em nuvem. Os resultados obtidos evidenciaram a eficiência do método em detectar novidades nos equipamentos, demonstrando a viabilidade do uso deste na plataforma de monitoramento.*

1. Introdução

Nos últimos anos, percebeu-se uma crescente demanda de energia elétrica por parte das indústrias, comércios e residências. Essa situação está presente tanto no cenário brasileiro quanto mundial. No período entre 1999 e 2009, o consumo *per capita* de energia elétrica no Brasil e no mundo aumentou 20% e 22% respectivamente [DATA 2012]. Tal cenário, portanto, exige um sistema elétrico mais inteligente que permita reduzir o consumo de energia elétrica em cada equipamento eletrônico, encorajando os consumidores a implementar estratégias eficientes para a redução do consumo de energia.

A tecnologia da informação para o sistema elétrico de potência, integrada aos sistemas de comunicação e infraestrutura de rede elétrica, conhecida como *smart grid*

[ENERGY 2012], permite monitorar e gerenciar o sistema de energia elétrica, em qualquer lugar a qualquer momento. Nesse sentido, a utilização das *smart grids* tem se tornado cada vez mais importante no cenário urbano, pois oferecem integração em diversas fontes de energia, tais como hidrelétrica, eólica, solar e atômica.

Espera-se que a utilização da *smart grid* venha a ser uma realidade nos próximos anos, uma vez que as indústrias, universidades e governos do mundo inteiro têm dedicado recursos expressivos para o desenvolvimento dessa tecnologia. Essa tendência pode ser confirmada por meio de diferentes projetos e iniciativas nacionais e internacionais do governo, indústria e academia, dedicados às *smart grids* [AlertMe 2012], [ENERGY 2012], [MAGGI 2012], [Erol-Kantarci and Mouftah 2010], [Will et al. 2009].

Apesar de todos os avanços conquistados nessa área, as concessionárias de energia elétrica não oferecem suporte para as instalações da rede elétrica nas residências (para, por exemplo, gerenciar o consumo de energia remotamente). Isso significaria mais custos para a empresa. Além disso, essas concessionárias fornecem somente o consumo total de energia gasto em uma casa. Determinar quais são as fontes individuais (equipamentos eletrônicos) que possuem maior influência na conta de luz não é uma tarefa trivial.

Assim, um dos possíveis caminhos propostos [Power-Meter 2012], [AlertMe 2012], [Campus-Metabolism 2012], [Erol-Kantarci and Mouftah 2011], [Duarte et al. 2011], [Jota et al. 2006], [Botte et al. 2005], para resolver esse problema é integrar uma Rede de Sensores Sem Fio (RSSF) nos equipamentos eletrônicos da residência. A partir de uma RSSF é possível estabelecer um sistema que monitora em tempo real o uso de cada tomada da casa, permitindo ao usuário conhecer o seu perfil de consumo de energia. A partir disso, o usuário pode descobrir se existe algum tipo de desperdício de forma pontual e efetuar as devidas correções.

Estudos mostram que fornecer informações sobre como, quando, onde e o que os usuários estão usando os ajuda a tomar decisões corretas [Mcmakin et al. 2002], [Stern 1999], [Ester 1985]. É fundamental, portanto, usar nas *smart grids* métodos inteligentes para detectar e enviar novidades, de forma individual e autônoma, para os usuários quando algo anômalo surge nos seus equipamentos eletrônicos. Tais anomalias podem surgir, por exemplo, quando um equipamento consome mais energia do que o esperado e/ou quando o equipamento começa a ter um comportamento fora do padrão.

Nesse contexto, este artigo propõe um método inteligente, nomeado de *Novelty Detection Power Meter* (NodePM), para detectar novidades em equipamentos eletrônicos monitorados por uma *smart grid*. Considerando a entropia de cada equipamento monitorado, a qual é calculada com base em um modelo de cadeia de markov, o método proposto detecta as novidades por meio de um algoritmo de Aprendizado de Máquina (AM). O NodePM é integrado a uma plataforma de monitoramento remoto de consumo de energia elétrica, que consiste de uma RSSF associada a uma aplicação em nuvem. Dessa forma, é possível enviar alertas autônomos de forma inteligente aos usuários (por exemplo, em um *smartphone*) quando algo anômalo surge nos equipamentos eletrônicos. Para avaliar a qualidade do método proposto, foi realizada uma análise de desempenho tendo como parâmetro de comparação o *Self-Organizing Novelty Detection* (SONDE).

Os resultados obtidos mediante a análise estatística evidenciaram a viabilidade do método na plataforma desenvolvida, obtendo desempenho satisfatório na detecção de

novidades no ambiente monitorado (veja a Seção 6).

Nesse sentido, este trabalho diferencia-se das soluções existentes em pelo menos três aspectos: (i) faz uso de um método inteligente para monitorar o consumo de energia elétrica; (ii) usa técnicas de AM para analisar o comportamento dos equipamentos eletrônicos por meio das RSSFs; (iii) e envia alertas de forma inteligentes ao *smartphone* quando algo anômalo surge.

O restante deste artigo está organizado da seguinte forma. A Seção 2 apresenta os trabalhos relacionados. A Seção 3 descreve a estratégia utilizada para o desenvolvimento do trabalho e o método proposto. A Seção 4 apresenta o *baseline* para o NodePM. A Seção 5 apresenta a base de dados utilizada. A Seção 6 apresenta os resultados dos experimentos. Finalmente, a Seção 7 apresenta as conclusões e trabalhos futuros.

2. Trabalhos Relacionados

Diversos trabalhos têm sido publicados na área de *smart grid* ao longo dos anos e esta seção apresenta trabalhos científicos [AlertMe 2012, Campus-Metabolism 2012, Power-Meter 2012, Duarte et al. 2011, Erol-Kantarci and Mouftah 2010, Jota et al. 2006, Botte et al. 2005] cujo foco corresponde, principalmente, ao monitoramento do consumo de energia elétrica. No entanto, apesar de todos os recentes avanços conquistados, ainda existem vários desafios e problemas em aberto nessa área. Por exemplo, a inexistência de uma metodologia para detectar anomalias/novidades no ambiente monitorado.

Um dos modelos mais antigo de uma *smart grid* é o projeto Telegestore. O Telegestore [Botte et al. 2005] é um sistema que gerencia os medidores residenciais e comerciais remotamente, visando explorar a rede de distribuição de baixa tensão entre os transformadores e os medidores. As principais desvantagens do Telegestore são: (i) a desconsideração de um método para detectar anomalias (por exemplo, *black-out*) na rede de baixa tensão; (ii) e, o monitoramento por setores não é explorado.

[Erol-Kantarci and Mouftah 2010] propõem o uso da RSSF para gerenciar a energia elétrica de uma residência por meio da *smart grid*. Para tanto, os pesquisadores propuseram um *Appliance Coordination* (ACCORD) para reduzir o custo de energia na hora de pico. As desvantagens do ACCORD são: (i) não aproveitar os benefícios que as RSSFs oferecem (por exemplo, monitoramento por setores); (ii) a falta de um método para detectar novidades no ambiente monitorado; (iii) e, por fim, os experimentos são simulados.

No Brasil, também há iniciativas em relação às *smart grids*, por exemplo, o Centro de Monitoramento de Usos Finais (CMUF) [Jota et al. 2006]. O projeto CMUF objetiva ajudar as pessoas a gerenciar o consumo de energia elétrica remotamente, consistindo em um sistema de baixo custo. Porém, as principais desvantagens do CMUF são: (i) não integrar na plataforma um método para detectar novidades no ambiente monitorado; (ii) e, devido à sua natureza centralizada, o sistema pode ser vulnerável a sobrecarga de dados.

O trabalho que mais se assemelha a este artigo é do [Duarte et al. 2011]. Os pesquisadores propõem um *smart meter* para fazer a medição da energia elétrica em uma residência. O medidor consiste em monitorar o consumo de energia elétrica indicando os equipamentos de maior consumo dentro do domicílio. Apesar da semelhança, as desvantagens do *smart meter* são: (i) a plataforma não utiliza um método para detectar novidades nos equipamentos eletrônicos; (ii) e, não usa um servidor em nuvem para facilitar o geren-

ciamento das informações.

3. Sistema de Monitoramento Remoto do Consumo de Energia em uma Residência

Algumas etapas foram necessárias para a realização deste trabalho, sendo elas: (i) o desenvolvimento de uma plataforma na qual utiliza uma RSSF e computação em nuvem; (ii) o uso do NodePM para detectar novidades no ambiente a ser monitorado; (iii) e, por fim, o desenvolvimento de uma aplicação para o usuário visualizar as informações de consumo de energia em sua residência.

Deve-se ressaltar, portanto, que a plataforma desenvolvida e apresentada a seguir é de simples instalação e fácil utilização, assim não é necessário ter um especialista na área. O NodePM é integrado na plataforma a fim de detectar novidades no ambiente monitorado. Além disso, o sistema desenvolvido permite monitorar o consumo de energia elétrica em qualquer lugar a qualquer momento.

3.1. Plataforma construída usando RSSF e computação em nuvem

A plataforma proposta é composta por três etapas:

- A Etapa 1 é responsável pela aquisição dos dados de consumo dos equipamentos eletrônicos por meio de uma RSSF.
- A Etapa 2 está relacionada com o processamento dos dados coletados na Etapa 1. Para tanto, foi desenvolvida uma aplicação em um servidor em nuvem (*app engine*¹), no qual gerencia os dados recebidos da RSSF.
- Por fim, a Etapa 3 é responsável pela disponibilização das informações sobre o consumo de energia elétrica para os interessados. Por isso, uma aplicação para o *smartphone* foi desenvolvida.

Na Figura 1, é apresentado o funcionamento da plataforma de monitoramento remoto do consumo de energia elétrica. Para isso, foram construídos protótipos de RSSF (Rótulo 1 Figura 1) a partir da montagem de *wattímetros* (equipamentos que mede o consumo de energia elétrica) *Kill-a-Watt* da empresa P3². Os *wattímetros* são ligados diretamente em uma tomada e junto a ela é conectado o equipamento que deseja monitorar. Como os *wattímetros* não possuem acesso a nenhum meio de comunicação, foi adicionado a eles um módulo XBee³ a fim de transmitir as informações de consumo de energia para um servidor. Dessa forma, construiu-se uma infraestrutura, *gateway* Arduino⁴, (Rótulo 2 da Figura 1) para enviar as informações via RSSFs e transmitir os dados lidos para o servidor em nuvem (Rótulo 3 da Figura 1, onde é implementado o método proposto). O servidor recebe os dados da RSSF, processa tais dados (detectando novidades) e gerencia (dados de consumo) o envio das informações, sendo possível fazer o monitoramento remoto por meio de um *smartphone* (Rótulos 4 e 5 da Figura 1).

Nesse sentido, o usuário recebe os alertas e visualiza as informações mais importantes do consumo de energia no Rótulo 4 da Figura 1. Já no Rótulo 5 da Figura 1, é feito o monitoramento por setores, tais como: o equipamento, cômodo da casa e/ou toda região de atuação, somente possível por causa da RSSF.

¹App Engine - Developers, <https://developers.google.com/appengine/>

²P3 International innovative electronic solutions, <http://www.p3international.com>

³Xbee-pro module datasheet, <http://ftp1.digi.com/support/documentation/90000976G.pdf>

⁴Arduino, <http://www.arduino.cc>

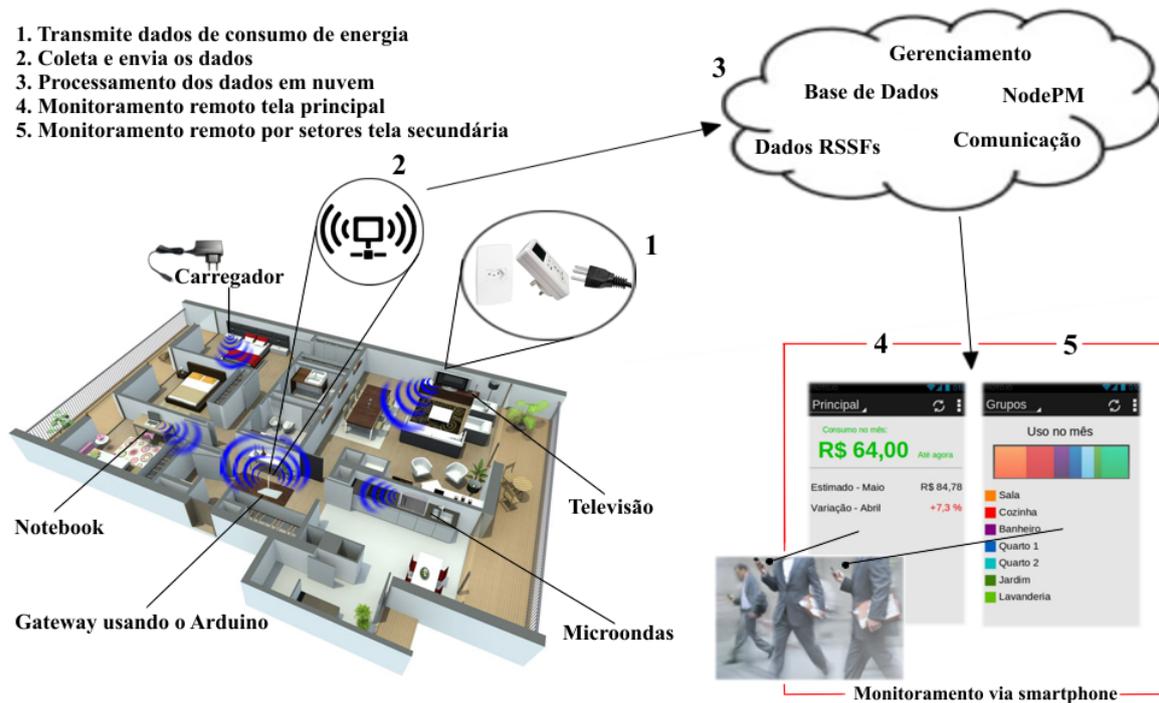


Figura 1. Cenário de funcionamento da plataforma.

3.2. Visão Geral do *Novelty Detection Power Meter* (NodePM)

O método proposto, Algoritmo 1, é nomeado como *Novelty Detection Power Meter* (NodePM) justamente por ser integrada na plataforma, servindo como um ponto de interconexão de envios de alertas entre o medidor e o usuário.

Algoritmo 1: *Novelty Detection Power Meter*

```

Entrada:  $X=(X_1 \dots X_n)$ ; /* conjunto de dados */
Saída: mensagens de alertas para o dispositivo móvel
Dados: Threshold, Markov;
1 inicio
2 Estado  $\leftarrow$  KNN( $X$ ); /* classifica a instância desconhecida */
3 Markov  $\leftarrow$  AtualizarCadeiaMarkov(Markov, Estado); /* incorporar a
  instância classificada em um estado da cadeia de Markov */
4 Probabilidade  $\leftarrow$  GetProbabilidade(Markov); /* vetor de probabilidade */
5 DeltaEntropia  $\leftarrow$  Shannon(Probabilidade); /* variação da entropia */
  /* detectando novidade no equipamento eletrônico */
6 se DeltaEntropia  $\geq$  Threshold então
7   | Enviar(Dispositivo); /* envia mensagem para o dispositivo móvel */
8   fim
9 fim

```

O NodePM é modelado usando os conceitos de AM, cadeias de Markov e entropia. Foram identificadas três etapas distintas para o funcionamento do Algoritmo 1.

- Primeira etapa, processar os dados por meio de um classificador, *K-Nearest Neighbors* (KNN) (Linha 2 do Algoritmo 1), para catalogar o estado comportamental

dos equipamentos eletrônicos.

- Segunda etapa, capturar o elemento catalogado na primeira etapa e adicionar no estado comportamental da cadeia de Markov (Linha 3 do Algoritmo 1).
- Terceira etapa, obter a matriz de probabilidade da cadeia de Markov e calcular o grau de incerteza dos equipamentos eletrônicos usando a variação da entropia (Linhas 4 e 5 do Algoritmo 1).

As etapas descritas anteriormente são reorganizadas, nas próximas subseções, para um melhor entendimento do método.

3.2.1. Classificação de Padrões Comportamentais

Nesta subseção é apresentada a técnica escolhida, KNN, para classificar os dados recebidos da RSSF, com o intuito de incorporá-los na cadeia de Markov (Subseção 3.2.2).

A escolha do KNN é justificada por dois motivos: (i) apesar de simples, tem mostrado ser uma das técnicas mais eficazes já proposta na literatura; (ii) e, é uma técnica, tradicional e efetiva aplicada em vários problemas de classificação [Yang and Liu 1999].

Para facilitar o entendimento da técnica e o seu funcionamento na plataforma, a Figura 2 ilustra como os dados de entrada são classificados. Os gráficos da Figura 2 representam a potência de um equipamento em função do tempo (minuto), sendo possível perceber quatro comportamentos distintos (*standby*, desligado, normal⁵ e não esperado⁶) de consumo de energia.

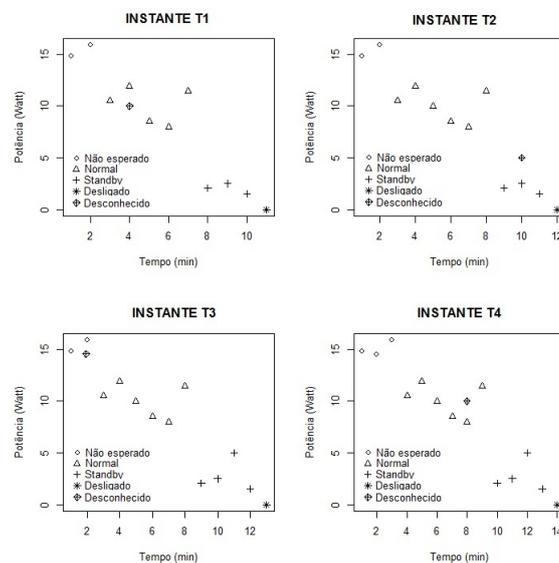


Figura 2. Exemplo de classificação do KNN.

Deve-se ressaltar, portanto, que o KNN classifica as instâncias desconhecidas analisando os K vizinhos mais próximos. Neste caso, a instância “desconhecida” no INSTANTE T1 (Figura 2) é classificada como uma instância “normal”. Após a classificação, é fundamental destacar, que o elemento classificado é adicionado incrementalmente na base de treinamento, a fim do método se adaptar de forma autônoma em ambientes dinâmicos.

⁵Representa o equipamento dentro de um padrão, a geladeira por exemplo fica ligada constantemente.

⁶Representa o equipamento fora do padrão e/ou consumindo energia acima do esperado.

A adaptação do ambiente, neste trabalho, ocorre a cada 5 segundos, tempo estipulado pela plataforma para enviar os dados. Dessa forma, o INSTANTE T2, T3 e T4 seguem o mesmo processo que o INSTANTE T1.

Assim, a próxima etapa do método consiste em adicionar a instância classificada, a cada instante de tempo, em um estado da cadeia de Markov.

3.2.2. Obtenção de Padrões Comportamentais

Para obtenção de padrões comportamentais a cadeia de Markov [Markov 1971] é utilizada. Assim, conforme o KNN classifica as instâncias desconhecidas, uma nova cadeia de Markov é gerada. Na Figura 3 é exibida a matriz de probabilidade e a cadeia de Markov em cada um dos instantes de tempo, como apresentado anteriormente na Figura 2, a fim de representar o comportamento do equipamento naquele instante de tempo.

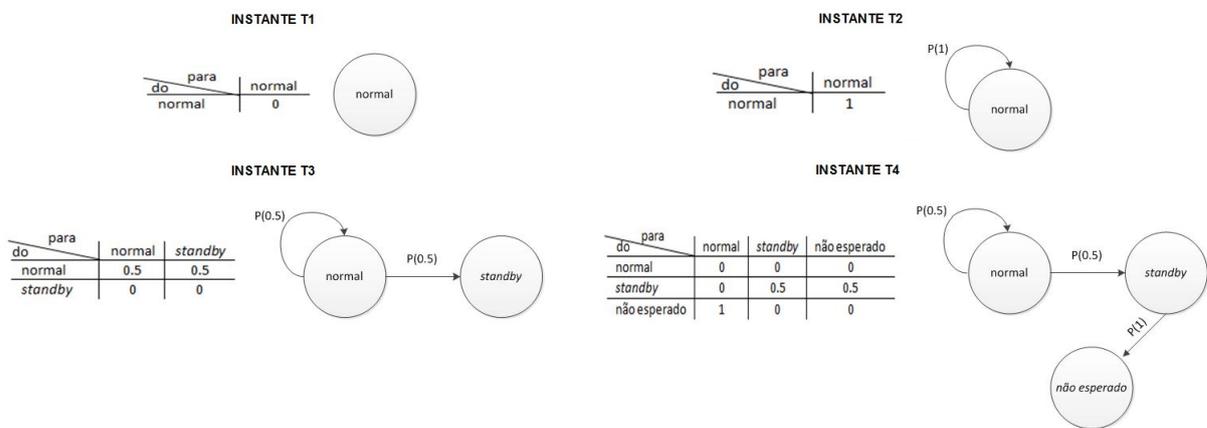


Figura 3. Matriz de probabilidade e cadeia de Markov a cada instante de tempo de acordo com a Figura 2.

Nesse sentido, a matriz de probabilidade e a cadeia de Markov são atualizadas de acordo com cada nova classificação do KNN (veja o paralelo da Figura 2 com a Figura 3). Neste caso, no INSTANTE T1 (Figura 3) classificou-se a primeira instância como “normal”, porém sem transição na cadeia de Markov. No INSTANTE T2 (Figura 3), a segunda instância também foi classificada como “normal”, no entanto há uma transição entre a primeira e segunda instância classificada, com isso, a matriz de probabilidade é atualizada, apresentando uma transição **do** estado “normal” **para** o estado “normal”. O INSTANTE T3 e T4 da Figura 3 seguem o mesmo processo. Assim, os estados da cadeia de Markov são criados dinamicamente a medida que o KNN classifica o comportamento do equipamento para aquele instante de tempo.

Após classificar, representar o comportamento dos equipamentos mediante a cadeia de Markov e obter as probabilidades a cada instante de tempo é necessário calcular o grau de incerteza dos equipamentos por meio da entropia da cadeia de Markov.

3.2.3. Detecção de Novidades com a Variação da Entropia da Cadeia de Markov

Com o conjunto das probabilidade das cadeias de Markov gerado torna-se possível calcular a variação da entropia a cada instante de tempo, por meio da seguinte fórmula:

$$H(X) = - \sum_{i=1} p(x_i) \log_b p(x_i)$$

Onde \log_b denota o logaritmo de x na base 2 e o $p(x_i)$ é a probabilidade de um evento ir para um outro estado (matriz de probabilidade da Figura 3). Dessa forma, cada interação do usuário com o equipamento gera uma curva de energia, a qual representa as alterações comportamentais dos equipamentos. Na Seção 6.1 são realizados experimentos dos tipos de novidades detectadas nos equipamentos.

4. O *Baseline* para o Método Proposto

O *Self-Organizing Novelty Detection* (SONDE) é uma rede neural que adapta incrementalmente sua estrutura de conhecimento (neurônio), a fim de detectar novidades em ambientes dinâmicos. Para isso, a SONDE classifica em um mesmo neurônio padrões de entradas similares. Quando nenhum neurônio é capaz de classificar um padrão de entrada, um novo neurônio é criado indicando uma novidade no ambiente [Albertini and Mello 2007].

A escolha da SONDE pode ser justificada por quatro motivos: primeiro, a SONDE pode ser utilizada em qualquer sistema e/ou aplicação independente da base de dados; segundo, o método detecta eventos inesperados em ambientes dinâmicos; terceiro, o treinamento e adaptação são realizados sem a intervenção de um especialista; e, por fim, quarto, a SONDE obteve melhor desempenho quando comparado com *Grow When Required* [Albertini and Mello 2007] e [Marsland et al. 2002].

5. Base de Dados

O método proposto utilizou um conjunto de dados (*dataset*), contendo informações reais sobre a interação do usuário com o equipamento eletrônico. Essa base de dados possui quatro atributos, sendo eles: (i) identificador do equipamento atribuído pelo sensor; (ii) potência em *Watt*; (iii) a data de utilização do equipamento; (iv) e o tempo de uso do equipamento naquele momento.

Deve-se ressaltar, no entanto, que o único atributo normalizado para o tipo inteiro é a data, a qual representa o número de dias decorridos a partir de uma data fixa. Esse atributo é normalizado pela própria plataforma. Esta utiliza o JSON⁷ para transmitir os dados em um formato padronizado. Assim, o método não necessita fazer qualquer normalização dos atributos recebidos.

Após o envio dos dados, foram realizadas oito partições das amostras do conjunto de dados coletados de três meses (veja Seção 6). Cada partição contém um conjunto de dados de 1 e 2 semanas. Cada elemento do conjunto (1 semana e 2 semanas) foi dividida em 25% de instâncias de treinamento e 75% em instâncias de testes. Essa técnica é conhecida como *hold-out*, no qual divide a base de dados em dois conjuntos (conjunto de treino e teste) [Mitchell 1997]. Para não existir grau de dependência nos experimentos, da Subseção 6.2, a divisão do conjunto de treino e teste foi replicada aleatoriamente 11 vezes para cada conjunto de parâmetros da Tabela 1. As replicações em um mesmo subconjunto têm o objetivo de contemplar os diferentes comportamentos que os usuários possuem (por exemplo, comportamento diferente no final de semana em relação aos demais dias).

⁷JSON, <http://www.json.org/>

6. Experimentos e Análise dos Dados Obtidos

A validação do método proposto, foi dividida em duas etapas: a primeira, Seção 6.1, apresenta e discute as novidades/anomalias encontradas nos equipamentos eletrônicos; a segunda, Seção 6.2, é realizado uma avaliação de desempenho do NodePM com a SONDE.

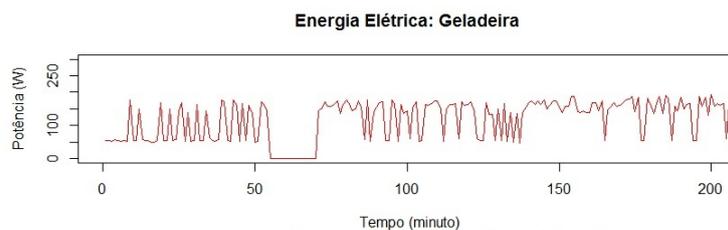
Para produzir resultados com boa precisão, foi montando um ambiente real para monitorar o consumo de energia dos equipamentos eletrônicos em uma residência. Nesse cenário, os equipamentos se localizavam em cômodos diferentes e, por conveniência, o *gateway* Arduino estava junto do roteador para comunicar com a Internet. A coleta de dados foi feita durante três meses e nesse período foi observado o comportamento do sistema por meio do perfil de consumo do usuário.

6.1. Experimentos de Detecção de Novidades com o NodePM

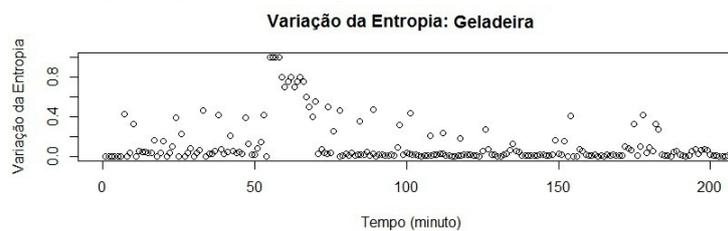
Com os dados coletados foi possível realizar experimentos com o método, o qual apresentou dois tipos de novidades (qualitativa e quantitativa):

- A 1ª ocorre quando o equipamento eletrônico tem uma mudança abrupta no seu comportamento padrão, ou seja, ocorreu uma troca não habitual no seu estado, passando do estado x para um estado y , no qual não era esperado. (qualitativa)
- A 2ª acontece quando o equipamento eletrônico começa a consumir energia mais que o esperado durante um determinado período. (quantitativa)

A Figura 4 apresenta o primeiro tipo de novidade, como descrito anteriormente, para um equipamento eletrônico (geladeira). O Gráfico 4b, no qual apresenta a variação da entropia em função do tempo (minuto) é possível notar tanto o comportamento padrão (sem novidade) entre o período de 0 a 54 e 71 a 200 minutos quanto o comportamento não esperado (com novidade) entre o período de 55 a 70 minutos.



(a) Consumo de energia elétrica da geladeira durante um período acontecendo algo inesperado entre o período de 55 a 70 minutos.

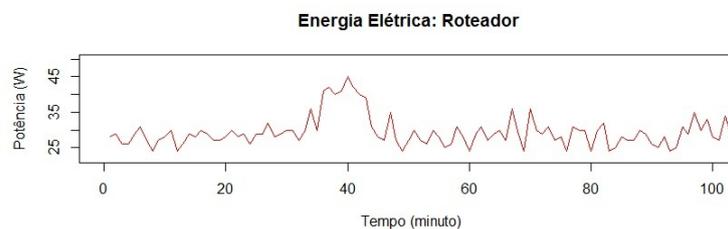


(b) Detectando novidade com uma mudança inesperada no comportamento da geladeira de acordo com o Gráfico 4a.

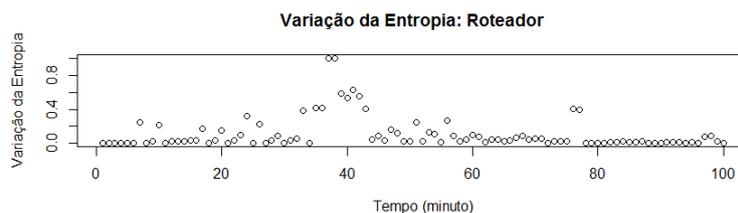
Figura 4. Detecção de novidades causada pela mudança no comportamento do equipamento.

Como o Gráfico 4a apresenta a potência consumida em função do tempo (minutos) é essencial observar o comportamento da geladeira a cada período. Nesse sentido, no intervalo de tempo entre 55 e 70 (Gráfico 4b) nota-se uma alteração repentina na entropia, causada pela mudança no padrão de comportamento da geladeira (Gráfico 4a). Esta mudança abrupta surgiu devido ao desligamento da geladeira durante 15 minutos (veja o Gráfico 4a no intervalo de 55 a 70), ou seja, o equipamento saiu do seu estado habitual o que não era esperado. Apesar disso, a variação da entropia começa a estabilizar depois dos 66 minutos. Essa estabilização pode ocorrer de duas maneiras: a primeira é devido à mudança de hábito do usuário (causada por *feedback* por intermediação do *smartphone*), e a segunda é que no decorrer do tempo o método considera que aquela novidade já não é algo inesperado.

O segundo tipo de novidade (Figura 5) está relacionado com o consumo elevado de energia elétrica, do roteador, durante um determinado período (quantitativo). O Gráfico 5a representa bem esse comportamento, expondo a potência consumida em função do tempo (minuto). Nesse gráfico é perceptivo um elevado pico de energia no período de 36 a 43 minutos. Sendo assim, o Gráfico 5b que representa a variação da entropia em função do tempo (minuto) constata a ocorrência de uma novidade no mesmo instante de tempo, que o Gráfico 5a, devido a uma mudança abrupta na variação da entropia. Vale destacar, que a estabilização da variação da entropia ocorre similar ao primeiro tipo de novidade, ou seja, a estabilização pode ocorrer por dois motivos, como descrita anteriormente.



(a) Consumo de energia elétrica do roteador durante um determinado período.



(b) Detectando novidade no mesmo instante de tempo que o Gráfico 5a.

Figura 5. Detecção de novidade com o aumento do consumo inesperado de energia elétrica do roteador.

Apesar de existir dois tipos de novidades, ambas podem estabilizar no erro como ocorrido no Gráfico 4b depois de 66 minutos. Este tipo de situação é considerado uma vantagem do ponto de vista do usuário. Visto que, caso o usuário não queira mudar o seu comportamento diante do equipamento defeituoso e/ou do consumo exagerado de energia elétrica, a metodologia utilizada não irá incomodar o usuário com os alertas enviados para o seu *smartphone*.

6.2. Avaliação de Desempenho com o NodePM

Esta subseção avalia o desempenho do NodePM, considerando as seguintes medidas de desempenho: sensibilidade (conhecida por alguns autores como taxa de detecção), precisão, especificidade e acurácia. Essas são calculadas a partir de uma matriz de confusão apresentada na Figura 6, no qual avalia os resultados com base nas perdas causadas.

		Classe Predita		
		Positivo	Negativo	
Classe Verdadeira	Positivo	Verdadeiro	Falso	Sensibilidade: $VP/(VP + FN)$
		Positivo (VP)	Negativo (FN)	Precisão: $VP/(VP + FP)$
	Negativo	Falso	Verdadeiro	Especificidade: $VN/(VN + FP)$
		Positivo (FP)	Negativo (VN)	Acurácia: $(VP + VN)/(VP + VN + FP + FN)$

Figura 6. Medidas de desempenho calculadas a partir da matriz de confusão.

Segundo [Fawcett 2006], tais medidas possuem características inerentes, sendo a sensibilidade o total de amostras cuja classe resultante é realmente positiva (verdadeiros positivos); a precisão é o total de exemplos classificados como positivo, mas que nem sempre são, ou seja, os verdadeiros negativos; já a especificidade é o oposto da sensibilidade, importando somente em classificar como negativo os exemplos que de fato são negativos (falsos positivos) e a acurácia permitirá analisar o quanto preciso o método classifica adequadamente o comportamento de um equipamento.

Os conjuntos de parâmetros estabelecidos para realizar a avaliação de desempenho são apresentados pela Tabela 1. O sistema é avaliado com a alteração do *dataset* (1 semana e 2 semanas), com a mudança do equipamento (geladeira e roteador), e, com a troca dos métodos (NodePM e SONDE). Tais métodos utilizam o *dataset* no qual contém informações reais da interação do usuário com o equipamento. Cada conjunto de parâmetros foi executado 11 vezes para cada partição do conjunto de dados (apresentado anteriormente, Seção 5), a fim de alcançar uma estabilidade dos dados. Os resultados são apresentados na Figura 7, com um intervalo de confiança de 95%.

Tabela 1. Conjunto de parâmetros escolhidos para serem avaliados.

Experimento	Método	<i>Dataset</i>	Equipamento
E1	NodePM	1 Semana	Geladeira
E2	NodePM	1 Semana	Roteador
E3	NodePM	2 Semanas	Geladeira
E4	NodePM	2 Semanas	Roteador
E5	SONDE	1 Semana	Geladeira
E6	SONDE	1 Semana	Roteador
E7	SONDE	2 Semanas	Geladeira
E8	SONDE	2 Semanas	Roteador

O Gráfico 7 apresenta a porcentagem dos resultados obtidos dos experimentos de E1 até E8 (Tabela 1) em função das medidas sensibilidade, precisão, especificidade e acurácia para o *dataset* de 1 e 2 semanas. O resultado destaca que o método NodePM possui um desempenho superior em relação a SONDE quando considerado o *dataset* de 1 semana, independente do equipamento utilizado (experimentos $\{(E1,E5),(E2,E6)\}$). Isso ocorre devido ao NodePM usar uma abordagem supervisionada, ou seja, a base de dados possuem exemplos que estão rotulados com uma classe predita. Logo, o método não

necessita de muitas instâncias de treinamento para atingir bons resultados, pois com o *dataset* de 2 semanas o seu desempenho foi significativo (aproximadamente 3% de aumento, experimentos de E1 a E4). Porém, ao levar em consideração o *dataset* de 2 semanas os papéis se invertem, ou seja, a SONDE tem melhor desempenho que o NodePM, independente do equipamento utilizado (experimentos $\{(E3,E7),(E4,E8)\}$). Isto faz sentido, pois a SONDE é uma rede neural artificial não-supervisionada e auto-organizável. Assim, é necessário ter um *dataset* maior para que os neurônios da SONDE se adapte, de forma incremental, de acordo com novos padrões de entrada. Por isso, o seu elevado desempenho (aproximadamente 13% de aumento, experimentos de E5 a E8) está relacionado com o aumento das instâncias de treinamento.

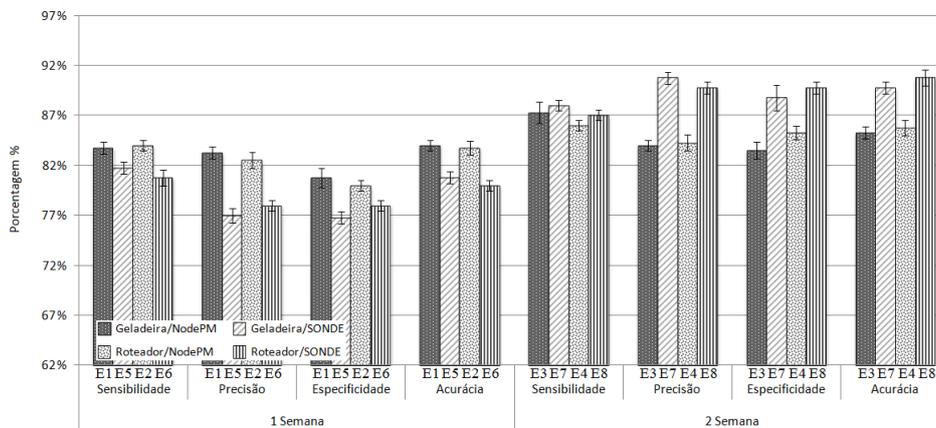


Figura 7. Análise de desempenho do NodePM de acordo com a Tabela 1.

6.2.1. Análise das Influências dos parâmetros

Esta subseção apresenta as análises das influências do conjunto de parâmetro consideradas na Tabela 1 em função das medidas de desempenho. As análises foram feitas utilizando um modelo de regressão multivariado, no qual inclui uma relação causal com mais de duas variáveis [Jain 1991]. Neste caso, o comportamento de uma medida de desempenho é explicado por mais de um parâmetro. Dessa forma, é possível analisar quais parâmetros e/ou suas combinações que mais influenciou nos resultados.

Tabela 2. Influência dos parâmetros em função das medidas de desempenho.

Medidas de Desempenho	A (%)	B (%)	C (%)	AB (%)	BC (%)	AC (%)	ABC (%)
Sensibilidade	3,06	81,05	2,25	12,25	0,57	0,25	0,57
Precisão	0,17	59,14	0,02	39,9	0,07	0,02	0,68
Especificidade	1,09	77,85	0,82	19,16	0,38	0,12	0,58
Acurácia	0,64	64,97	0,07	33,16	0,95	0,01	0,02

Na Tabela 2, são apresentados os parâmetros métodos, *dataset* e equipamento com as respectivas letras A, B e C. As combinações de duas ou mais letras são os percentuais das interações entre os parâmetros. Nota-se que o parâmetro com mais influência para todas as medidas de desempenho é justamente o *dataset* (B), no qual também pode ser observado nos resultados pelos pares dos experimentos (E1,E3), (E2,E4), (E5,E7), (E6, E8). O segundo parâmetro com mais influência é a combinação do *dataset* com o método

(AB). Tal influência faz sentido, pois é relacionada com a forte interação que o parâmetro A possui com o B. Deve-se ressaltar, que o parâmetro com menos influência é o equipamento independente do conjunto de parâmetro considerado. Essas afirmações estatísticas evidenciam a eficácia do NodePM na plataforma desenvolvida.

7. Conclusões e Trabalhos Futuros

Com base nos resultados, constatou-se que o NodePM foi viável na plataforma. Porém, cada método teve suas peculiaridades. Assim sendo, a escolha do método ficou intrínseco ao *dataset*, pois os dados coletados em tempo real da interação do usuário com o equipamento possuem características distintas. Dessa forma, se usuário utilizar equipamentos que são ligados constantemente, tais como geladeira, roteador e/ou *freezer*, no qual possuem mais dados com interação, o mais adequado é utilizar a SONDE (veja Figura 7, experimentos E7 e E8). Porém, caso o usuário interaja com equipamentos, tais como microondas, máquina de lavar e/ou cafeteira, no qual os aparelhos não são ligados constantemente tendo poucos dados de interação, o mais adequado é o NodePM (veja Figura 7, experimentos E1 e E2).

Este artigo apresentou as seguintes contribuições: (i) o uso de uma RSSF para a construção de um protótipo, no qual monitora o consumo de energia dos equipamentos eletrônicos individualmente; (ii) um método inteligente, NodePM, baseado nos conceitos de AM para detectar novidades em um ambiente monitorado; (iii) uma aplicação para *smartphone* no qual recebe as informações do consumo de energia pelo servidor em nuvem, apresentando as informações para o usuário; (iv) a utilização de um servidor em nuvem como solução para gerenciar tanto as informações recebidas das RSSFs quanto da utilização do método; (v) e finalmente, uma avaliação de desempenho do NodePM, apresentando o seu desempenho na detecção de novidades nos equipamentos eletrônicos monitorados por uma *smart grid*.

Como trabalho futuro, pretende-se desenvolver um *Middleware* adaptativo com base no OpenCom [Coulson et al. 2008], para mudar em tempo de interação qualquer método de acordo com o ambiente monitorado.

8. Agradecimentos

Os autores agradecem o apoio financeiro da CAPES, processo DS-7901025/M; FAPESP, processos 2012/12061-1, 2008/05346-4 e 2008/57870-9; e CNPq, 573963/2008-8.

Referências

- Albertini, M. K. and Mello, R. F. (2007). A self-organizing neural network for detecting novelties. In *Proceedings of the 2007 ACM symposium on Applied computing*.
- AlertMe (2012). Energy monitor save energy & lower electricity costs alerme. <http://www.alertme.com>. visitado em 13 de Março de 2012.
- Botte, B., Cannatelli, V., and Rogai, S. (2005). The telegestore project in enel's metering system. *International Conference on Electricity Distribution*, 18th.
- Campus-Metabolism (2012). Arizona state university, (asu). <http://cm.asu.edu/#app=4c23&6741-selectedIndex=1>. visitado em 05 de julho de 2012.
- Coulson, G., Blair, G., Grace, P., Taiani, F., Joolia, A., Lee, K., Ueyama, J., and Sivaharan, T. (2008). A generic component model for building systems software.

- DATA, P. (2012). Indicadores do desenvolvimento mundial. http://www.google.com.br/publicdata/explore?ds=d5bncppjof8f9_. visitado em 02 de julho de 2012.
- Duarte, L. F. C., Zambiacco, J. D., Airoidi, D., Ferreira, E. C., and Dias, J. A. S. (2011). Characterization and breakdown of the electricity bill using custom smart meters: a tool for energy-efficiency programs. *International journal of circuits, system and signal processing*.
- ENERGY (2012). Smart grid — department of energy. <http://energy.gov/oe/technology-development/smart-grid>. visitado em 05 de março de 2012.
- Erol-Kantarci, M. and Mouftah, H. (2011). Wireless sensor networks for smart grid applications. In *Electronics, Communications and Photonics Conference (SIECPC)*.
- Erol-Kantarci, M. and Mouftah, H. T. (2010). Wireless sensor networks for domestic energy management in smart grids. *Biennial Symposium on Communications*, 25th.
- Ester, P. (1985). Consumer behavior and energy conservation : a policy-oriented experimental field study on the effectiveness of behavioral interventions promoting residential energy conservation / by peter ester.
- Fawcett, T. (2006). An introduction to roc analysis. *Elsevier*, 27:861 – 874.
- Jain, R. (1991). The art of computer systems performance analysis: Techniques for experimental design, measurement, simulation, and modeling. *SIGMETRICS Perform.*
- Jota, F. G., Jota, P. R. S., and Nobre, E. C. (2006). Gerenciamento efetivo de energia por uso final: Um sistema de monitoramento de baixo custo via internet. *Seminário Nacional de Distribuição de Energia Elétrica*, XVII.
- MAGGI, B. (2012). Projeto de lei nº 84/2012. <http://www.senado.gov.br/atividade/materia/-getPDF.asp?t=105231&tp=1>.
- Markov, A. A. (1971). *Extension of the Limit Theorems of Probability Theory to a Sum of Variables Connected in a Chain*. In Dynamic Probabilistic System.
- Marsland, S., Shapiro, J., and Nehmzow, U. (2002). A self-organising network that grows when required. *Neural Netw.*, 15:1041–1058.
- Mcmakin, A. H., Malone, E. L., and Lundgren, R. E. (2002). Motivating residents to conserve energy without financial incentives abstract:.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill Science.
- Power-Meter (2012). Google powermeter: A google.org project. <http://www.google.com/powermeter>. visitado em 14 de Março de 2012.
- Stern, P. C. (1999). Information, incentives, and proenvironmental consumer behavior. *Journal of Consumer Policy*, 22:461–478.
- Will, J., O’Connell, T., and Lange, C. (2009). Smart grid is a global priority. *Smart Grid*. <https://courses.cit.cornell.edu/crp384/2009reports/LangeOt>
- Yang, Y. and Liu, X. (1999). A re-examination of text categorization methods. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*.