

## Análise de Compromisso entre Granularidade e Interpretabilidade em Sistemas Granulares Evolutivos<sup>\*</sup>

Luiz A. Q. Cordovil<sup>\*</sup> Pedro H. S. Coutinho<sup>\*</sup>  
Iury V. de Bessa<sup>\*,\*\*</sup> Murilo O. Camargos<sup>\*</sup>  
Reinaldo Martinez Palhares<sup>\*\*\*</sup>

<sup>\*</sup> Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Minas Gerais, MG, (e-mails: luiz.cordovil05@gmail.com, phc Coutinho@ufmg.br, murilo.camargosf@gmail.com).

<sup>\*\*</sup> Departamento de Eletricidade, Universidade Federal do Amazonas, AM (e-mail: iurybessa@ufam.edu.br)

<sup>\*\*\*</sup> Departamento de Engenharia Eletrônica, Universidade Federal de Minas Gerais, MG, (e-mail: rpalhares@ufmg.br)

**Abstract:** In this work, a procedure for data abstraction and knowledge formation based on evolving granular systems is introduced. With the use of ellipsoidal information granules, an optimal granularity allocation approach for real-time data processing based on fuzzy sets is developed in order to obtain interpretable representations. From the real-time data analysis, the goal is to maximize the fuzzy rules support that characterizes the information granules preventing the representation becoming abstract, compromising the interpretability of the rules. In this sense, experiments regarding time-series granulation are performed to analyze the trade-off between granularity and interpretability of evolving systems.

**Resumo:** Neste trabalho, é apresentado um procedimento de abstração de dados e formação de conhecimento baseado em sistemas granulares evolutivos. Com a utilização de grânulos de informação elipsoidais, uma abordagem de alocação ótima de granularidade para processamento em tempo real de fluxo de dados baseada em conjuntos *fuzzy* é desenvolvida para obtenção de representações interpretáveis. A partir da análise em tempo real dos dados, busca-se a maximização do suporte das regras *fuzzy* que caracterizam os grânulos de informação sem que a representação se torne abstrata comprometendo a interpretabilidade das regras. Neste sentido, experimentos a respeito de granulação de séries temporais são realizados para analisar a relação de compromisso entre granularidade e interpretabilidade de sistemas evolutivos.

**Keywords:** Granular Computing; Evolving Information Granules; Interpretability.

**Palavras-chave:** Computação granular; Grânulos de Informação Evolutivos; Interpretabilidade.

### 1. INTRODUÇÃO

A análise de dados para a obtenção de informações é etapa fundamental para o desenvolvimento de uma base de conhecimento a respeito de um objeto ou processo. De forma geral, esta é feita pela abstração dos dados em batelada ou em fluxo, de forma a extrair características e apresentá-las de maneira objetiva (Yao, 2018). A partir da base de conhecimento formada, objetiva-se evidenciar as características, identificar tendências e similaridades, e alcançar interpretabilidade. Não obstante, não se trata apenas de formar conhecimento, mas formar conhecimento baseado em fatos, com significado e de maneira contínua.

No que se refere à inteligência computacional, há diferentes técnicas utilizadas para formação de conhecimento a partir de processamento em tempo real de fluxo de dados tais

<sup>\*</sup> Este trabalho foi apoiado em parte pelas agências brasileiras CNPq, FAPEMIG, CAPES - Código de Financiamento 001 e pelo Programa de Bolsas PROPG-CAPES/FAPEAM.

como as baseadas em lógica *fuzzy* (Leite et al., 2015; Lemos et al., 2011), redes neurais artificiais (Silva et al., 2014) e estatística (Costa et al., 2016). Em particular, no que diz respeito às técnicas baseadas em lógica *fuzzy*, um paradigma importante para o processamento de informações é o Princípio da Granularidade Justificável (PGJ) introduzido por Pedrycz e Homenda (2013). A partir do PGJ, os dados são traduzidos em grânulos de informação (GIs) que fornecem evidências numéricas com relevância semântica. GIs são representações interpretáveis de informação que possuem capacidade de abstração, isto é, eles podem representar dados como conceitos significativos de maneira simplificada (Wang et al., 2016). Com a utilização de modelos granulares, grandes quantidades de dados podem ser separadas em partes menores distinguíveis e com semântica bem definida (Yin et al., 2017).

A granularidade de um modelo está relacionada à quantidade de GIs obtidos, que por sua vez afeta a interpretabilidade do mesmo. As regras obtidas com os GIs

são propriedades que traduzem a incerteza dos dados em conjuntos *fuzzy* com suporte definido. A partir destas, pode-se então inferir o significado e chegar a conclusões com técnicas de raciocínio aproximado (Liu e Liu, 2002).

Neste artigo, é proposto um sistema granular evolutivo baseado em grânulos elipsoidais. Com uma adaptação do PGJ para processamento em tempo real, propõe-se um procedimento de alocação ótima de granularidade baseado em dados com o objetivo de maximizar evidência numérica e semântica. Ademais, uma discussão a respeito do compromisso entre granularidade e interpretabilidade é apresentada considerando o aspecto de granularidade de modelo como definido por Mencar e Fanelli (2008), que discutem restrições sobre a interpretabilidade de GIs.

O restante deste artigo é organizado da seguinte forma: na Seção 2, são apresentados os fundamentos de Computação Granular considerando o Princípio da Granularidade Justificável e a formulação do problema para obtenção de representações interpretáveis para fluxo de dados em tempo real; na Seção 3, discute-se a relação de conflito entre granularidade e interpretabilidade para sistemas granulares evolutivos, permitindo discutir o compromisso entre tais propriedades; a metodologia proposta para desenvolvimento de modelos granulares elipsoidais evolutivos é apresentada na Seção 4, compreendendo os procedimentos de alocação ótima de granularidade e parametrização de funções de pertinência hiper-elipsoidais; na Seção 5, o compromisso entre granularidade e interpretabilidade é analisado a partir de um estudo de caso sobre granulação de séries temporais provenientes de um *benchmark* de turbina eólica; e, por fim, conclusões a respeito da metodologia proposta são apresentadas na Seção 6. Em relação à notação, o uso de negrito irá denotar a representação vetorial como, por exemplo,  $\mathbf{m}$ . A ausência de negrito irá denotar um escalar como, por exemplo,  $m$ .

## 2. FUNDAMENTOS E FORMULAÇÃO DO PROBLEMA

Considere um fluxo de dados cuja  $k$ -ésima amostra  $\mathbf{x}_k = [\mathbf{x}_{k1}, \mathbf{x}_{k2}, \dots, \mathbf{x}_{kp}]^T \in \mathbb{R}^p$  é abstraída de tal forma a contribuir no desenvolvimento de representações interpretáveis com base em evidência numérica. Com o objetivo de extrair características ou informações, tais amostras são organizadas de acordo com certos critérios, os quais podem incluir medidas de proximidade e similaridade.

Neste sentido, algoritmos de clusterização são usualmente empregados para a obtenção de representações que proveem medidas quantitativas quanto ao particionamento dos dados no espaço como, por exemplo, distância à média e grau de pertinência referentes aos métodos  $K$ -means e *Fuzzy c-means* (FCM), respectivamente. De fato, estes algoritmos são base para o estabelecimento de protótipos de grânulos de informação ( $\mathcal{G}$ ) que, por sua vez, precisam ser aprimorados no sentido de garantir máxima compreensão dos dados sem serem abstratos a ponto de não representar conhecimento algum (Gacek e Pedrycz, 2015).

Em abordagens de formação de conhecimento baseada em dados, a evidência numérica pode ser definida como uma representação intervalar *fuzzy* de uma coleção de dados  $\mathbf{X} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_n^T]^T$ , que é parametrizada em termos

da média ( $\mathbf{m}$ ), e limites inferior ( $\mathbf{a}$ ) e superior ( $\mathbf{b}$ ) (Ren et al., 2018).

Por simplicidade, admita um fluxo de dados com  $x_k \in \mathbb{R}$ . A representação intervalar de  $\mathbf{X}$  é iniciada com a separação dos dados de acordo com a média, tal que:

$$\begin{aligned} \mathbf{X}_k^- &= \{x_k \mid x_k < m, k = 1, \dots, n\}, \\ \mathbf{X}_k^+ &= \{x_k \mid x_k \geq m, k = 1, \dots, n\}. \end{aligned} \quad (1)$$

A determinação dos limites  $a$  e  $b$  é obtida de  $\mathbf{X}_k^-$  e  $\mathbf{X}_k^+$ , respectivamente. O significado prático destes limites é garantir parâmetros que levem à alocação ótima de granularidade. Neste contexto, uma importante ferramenta para a criação de grânulos é o Princípio da Granularidade Justificável (PGJ) e a partir de sua aplicação para análise dos dados em batelada, busca-se a maximização da evidência numérica e do significado semântico da representação, com base em duas variáveis denominadas: cobertura (*coverage*) e especificidade (*specificity*) (Pedrycz e Wang, 2016).

A cobertura (*cov*) é uma medida relacionada à cardinalidade *fuzzy* do grânulo, ao passo que, a especificidade (*sp*) está relacionada ao nível de abstração do conjunto *fuzzy* a partir de suas funções de pertinência  $e$ , conseqüentemente, de suas regras. De maneira geral, cobertura e especificidade são definidas como (Liu et al., 2017):

$$\begin{aligned} \text{cov}(\mathbf{X}^-, g) &= \sum_{x \in \mathbf{X}^-} g(x; a, b), \\ \text{sp}(\mathbf{X}^-, g) &= 1 - \int_0^1 \frac{m - g^{-1}(\alpha; a, b)}{m - \min(\mathbf{X}^-)} d\alpha, \\ \text{cov}(\mathbf{X}^+, g) &= \sum_{x \in \mathbf{X}^+} g(x; a, b), \\ \text{sp}(\mathbf{X}^+, g) &= 1 - \int_0^1 \frac{g^{-1}(\alpha; a, b) - m}{\max(\mathbf{X}^+) - m} d\alpha, \end{aligned}$$

em que  $g(\cdot)$  é a função de pertinência do grânulo, definida como uma função unimodal de suporte compacto, ou seja, possui valor de máximo único e é diferente de zero apenas num domínio  $[a, b] \subset \mathbb{R}$ , com  $a < b$ . A partir destas definições, formulam-se dois problemas de otimização que, baseados na relação de conflito entre cobertura e especificidade, determinam os limites ótimos do grânulo:

$$\begin{aligned} a^{\text{opt}} &= \arg \max_a \tilde{Q}(\mathbf{X}^-, g), \\ b^{\text{opt}} &= \arg \max_b \tilde{Q}(\mathbf{X}^+, g), \end{aligned}$$

em que  $\tilde{Q}(\mathbf{C}, g) = \text{cov}(\mathbf{C}, g) \cdot \text{sp}(\mathbf{C}, g)$ . Quanto maior a cobertura, mais dados estão sendo considerados na representação granular, ao passo que, quanto maior a especificidade, menos abstrato é o grânulo. A função de rendimento do grânulo é definida como:

$$Q = \tilde{Q}(\mathbf{X}^-, g^{\text{opt}}) + \tilde{Q}(\mathbf{X}^+, g^{\text{opt}}),$$

em que  $g^{\text{opt}}(x) = g(x; a^{\text{opt}}, b^{\text{opt}})$ . Todas estas medidas estão diretamente relacionadas com a forma da função de pertinência e a alocação ótima de granularidade depende de  $a^{\text{opt}}$  e  $b^{\text{opt}}$  (Liu et al., 2018).

## 3. GRANULARIDADE E INTERPRETABILIDADE

Como retratado na Seção 1, a extração de características a partir de sistemas granulares evolutivos considera a análise

de similaridade entre os dados e, com o estabelecimento de regras de distinção, outras informações podem ser identificadas. Além disso, a representação das características a partir do uso de GIs remonta o conceito de cobertura e especificidade como discutido na Seção 2. Mencar e Fanelli (2008) discutem de que maneira sistemas granulares evolutivos são construídos e sua influência na interpretabilidade a partir da análise das regras *fuzzy* obtidas durante o processamento de dados. O importante é que, em função dos grânulos, se estabeleça uma distinção entre o que é mais detalhado e específico e o que é mais abstrato e geral (Pedrycz e Gomide, 2007).

Acerca do conceito de granularidade, diversas interpretações podem ser encontradas em artigos relacionados: no que se refere a um grânulo, à granularidade refere-se à cardinalidade deste, ou seja, a quantidade de dados que são cobertos por uma regra (Leite et al., 2012); considerando o modelo como um todo, a granularidade é medida como função da quantidade de grânulos obtidos (Ahmed e Isa, 2015). O conceito de interpretabilidade está relacionado com o significado que o grânulo pode representar (Antonelli et al., 2011). Isto está diretamente relacionado com as medidas que são utilizadas para organizar dados de acordo com algum critério, como discutido na Seção 2. Com isso, espera-se extrair o máximo possível de informações interpretáveis a partir de grânulos justificados em termos de granularidade, cobertura e especificidade.

Considerando estes aspectos, alguns questionamentos podem ser levantados, como: até que ponto sistemas evolutivos com alta granularidade podem implicar em representações interpretáveis? Grânulos compactos e específicos proveem de fato informação relevante, ou são apenas agrupamentos limitados em termos de semântica? A incerteza das informações precisa realmente de representações específicas e compactas, ou podem ser interpretáveis ao considerar representações mais gerais e abstratas?

Uma análise sobre o compromisso entre granularidade e interpretabilidade será discutido na Seção 5 a partir da utilização de grânulos de informação evolutivos, considerando formas elipsoidais.

#### 4. GRÂNULOS ELIPSOIDAIIS EVOLUTIVOS

Quando se trata de interpretabilidade no contexto de conjuntos *fuzzy*, é importante analisar a influência que a forma destes tem sobre a definição da função de pertinência (Mencar e Fanelli, 2008). A depender da forma como os dados estão distribuídos, a escolha da função de pertinência tem efeito na granularidade e na capacidade do sistema granular de representar os dados de forma interpretável.

De acordo com Zhu et al. (2017), quando os dados estão dispersos em relação ao centro de um protótipo, utilizar formas elipsoidais para a construção de grânulos de informação é uma alternativa adequada àquelas que apresentam certas limitações como por exemplo formas cúbicas. Neste sentido, outro aspecto importante diz respeito à dimensão dos dados, pois, em aplicações que não consideram formas elipsoidais, os grânulos de informação multidimensionais são construídos considerando o produto Cartesiano de componentes obtidas para cada dimensão.

Neste método, as formas elipsoidais são consideradas para a granulação em tempo real de fluxo de dados multidimensionais baseado no PGJ. Cada forma elipsoidal  $E$  é definida em termos de um centro e de parâmetros relacionados à covariância, tal que:

$$E = \{\mathbf{x} \in \mathbf{X} | M \leq 1\}, \quad (4)$$

em que  $M = (\mathbf{x} - \mathbf{m})^T C^{-1} (\mathbf{x} - \mathbf{m})$  é o quadrado da distância de Mahalanobis de uma amostra  $\mathbf{x}$  ao centro da elipse ( $\mathbf{m}$ ) e  $C$  é a matriz de covariância, a qual está relacionada à inclinação e região de cobertura da elipse.

Além disso, é necessário que sejam estabelecidos critérios para a detecção de pontos de mudança ou anomalias durante o processamento dos dados, já que o objetivo principal é a obtenção de representações discerníveis e interpretáveis. Moshtaghi et al. (2016) estabelecem uma medida de distinção entre dados comuns e anomalias de acordo com um limiar ( $\ell$ ), que, por sua vez, é função da dimensão dos dados ( $p$ ). Este limiar é definido como a inversa da distribuição qui-quadrado:

$$\ell = (\chi^2)^{-1}(\gamma, p). \quad (5)$$

Em termos práticos, (5) auxilia na distinção dos dados, pois fornece um valor que quando comparado com  $M$  indica que a amostra é diferente de  $\gamma\%$  das amostras de uma distribuição qui-quadrado com  $p$  graus de liberdade. Se  $M$  exceder o limiar  $\ell$ , uma anomalia é detectada podendo implicar ou não na criação de uma nova elipse e, conseqüentemente, de um novo grânulo. Desta forma, (4) é reescrita considerando este limiar da seguinte forma:

$$E = \{\mathbf{x} \in \mathbf{X} | M \leq \ell\}.$$

A criação de um novo grânulo é realizada após a comparação da distância com o limiar, considerando a persistência de anomalias durante um determinado intervalo discreto de tempo ( $T_s$ ) e a verificação da condição de  $c$ -separação de elipses introduzida por Dasgupta (1999). A detecção de pontos de mudança é feita utilizando uma estrutura (*Tracker*) para monitorar o estado do sistema. Essa estrutura é definida por um grânulo elipsoidal com fator de esquecimento, respondendo rapidamente à mudanças nos dados. Quando o *Tracker* se torna suficientemente diferente dos outros grânulos, ou seja, está  $c$ -separado de todos os grânulos existentes, um novo grânulo é criado.

À medida em que os dados são disponibilizados e as condições de criação de grânulos são atendidas, a granulação em tempo real se dá em três etapas:

- Obtenção recursiva dos parâmetros que definem a forma elipsoidal;
- Parametrização recursiva da função de pertinência ( $g$ ) de acordo com os dados;
- Atualização recursiva dos parâmetros do grânulo de informação de acordo com um índice de rendimento relacionado com a cardinalidade fuzzy e a distância da amostra ao centro de um dado GI.

Essas três etapas são descritas nas subseções a seguir.

##### 4.1 Obtenção recursiva dos parâmetros da forma elipsoidal

Considerando uma amostra  $p$ -dimensional  $\mathbf{x}_k$ , a evidência numérica é definida como função de  $\mathbf{m}_k^i$ ,  $L_{kj}^i$  e  $R_{kj}^i$

$$\mathbf{m}_k^i = \mathbf{m}_{k-1}^i + \frac{\omega^i(\mathbf{x}_k)}{\sum_{i=1}^{N_k} \omega^i(\mathbf{x}_k)} (\mathbf{x}_k - \mathbf{m}_{k-1}^i), \quad (6)$$

$$L_{kj}^i = \frac{1}{l_{kj}} \left( L_{(k-1)j}^i l_{(k-1)j} + m_{kj}^i - x_{kj}^i \right), \quad (7)$$

$$R_{kj}^i = \frac{1}{r_{kj}} \left( R_{(k-1)j}^i r_{(k-1)j} + x_{kj}^i - m_{kj}^i \right), \quad (8)$$

em que  $i = 1, \dots, N_k$ ,  $j = 1, \dots, p$ ,  $N_k$  é a quantidade de grânulos e  $\omega^i(\cdot)$  é a função de pertinência do  $i$ -ésimo grânulo. A atualização da matriz de covariância  $C$  em termos da distância de Mahalanobis é realizada de acordo com Moshtaghi et al. (2015). Os parâmetros  $l_{kj}$  e  $r_{kj}$  correspondem, respectivamente, às cardinalidades dos conjuntos  $\mathbf{X}_k^-$  e  $\mathbf{X}_k^+$  definidos em (1).

#### 4.2 Parametrização recursiva da função de pertinência

A função de pertinência elipsoidal também é definida como função da evidência numérica obtida até a  $k$ -ésima amostra considerando todas as dimensões. A partir do conceito de número fuzzy  $p$ -elipsoidal (Wang et al., 2014), a função de pertinência é definida da seguinte forma:

$$\omega(\mathbf{x}_k) = \exp \left( -2 \left[ (\mathbf{x}_k - \mathbf{m}_k) \Delta^{-1} (\mathbf{x}_k - \mathbf{m}_k)^\top \right]^{\frac{1}{2}} \right),$$

em que  $\Delta = \text{diag} \{ (b_{k1} - a_{k1})^2, \dots, (b_{kp} - a_{kp})^2 \}$ , sendo  $\left( \frac{b_j - a_j}{2} \right)$  os semieixos das elipses. Desta forma, a partir de (7) e (8), definem-se:

$$a_{kj} = \max \{ m_{kj} - \lambda L_{kj}, \alpha \}, \quad (9)$$

$$b_{kj} = \min \{ m_{kj} + \lambda R_{kj}, \beta \}, \quad (10)$$

tal que  $[\alpha, \beta] = [0, +\infty)$  e  $\lambda \in [2, 4]$ .

Para a quantidade atual de grânulos obtidos ( $N_k$ ), ao longo do processo de granulação, o grau de pertinência normalizado de uma amostra para o  $i$ -ésimo grânulo é dado por

$$g^i(\mathbf{x}_k) = \frac{\omega^i(\mathbf{x}_k)}{\sum_{i=1}^{N_k} \omega^i(\mathbf{x}_k)}.$$

#### 4.3 Atualização Recursiva dos parâmetros do grânulo

Baseado no PGJ, para a alocação ótima de granularidade a partir da definição dos semieixos das elipses em (9) e (10), é desenvolvida uma abordagem baseada nos dados para a avaliação do rendimento do grânulo, tal que

$$Q_k(\mathcal{G}^i) = \sum_{l=1}^{|\mathcal{G}^i|} \tilde{Q}^i(\mathbf{x}_l^i), \quad (11)$$

em que  $|\mathcal{G}^i|$  é o número de amostras alocadas em  $\mathcal{G}^i$  e o termo  $\tilde{Q}^i(\mathbf{x}_k)$  é uma medida de contribuição do dado  $\mathbf{x}_k$  em relação a um grânulo  $\mathcal{G}^i$  e é definida por

$$\tilde{Q}^i(\mathbf{x}_k) = M^i(\mathbf{x}_k) \cdot \sum_{l=1}^{|\mathcal{G}^i|+1} g_{k+1}^i(\tilde{\mathbf{x}}_l^i), \quad (12)$$

em que  $\tilde{\mathbf{x}}_l^i$  denota a amostra alocada no grânulo  $\mathcal{G}^i$  e  $M^i(\mathbf{x}_k)$  é o quadrado da distância de Mahalanobis entre a amostra  $\mathbf{x}_k$  e o  $i$ -ésimo grânulo. Esta análise é fundamental para verificar a atualização dos parâmetros dos grânulos sem considerar dados passados. A relação estabelecida

entre a contribuição do dado e o rendimento do grânulo caracterizam o compromisso entre medidas conflitantes assim como no PGJ, pois, quanto maior a distância, menor o grau de pertinência e vice-versa.

Com a disponibilização de uma nova amostra  $\mathbf{x}_k$ , avalia-se sua influência aos rendimentos dos grânulos criados. Caso o rendimento obtido após a nova amostra ser admitida seja maior do que o rendimento do grânulo no instante anterior, os parâmetros do grânulo são atualizados. Para avaliar esta condição, considere que a amostra  $\mathbf{x}_k$  é admitida ao  $i$ -ésimo grânulo, denotando-a por  $\mathbf{x}_k^*$ . O seu grau de pertinência ao  $i$ -ésimo grânulo criado é dado por  $g_k^i(\mathbf{x}_k^*)$ . Assim, a soma cumulativa dos graus de pertinência das amostras já admitidas no  $i$ -ésimo grânulo é computada de acordo com

$$\sum_{l=1}^{|\mathcal{G}^i|+1} g_{k+1}^i(\tilde{\mathbf{x}}_l^i) = \sum_{l=1}^{|\mathcal{G}^i|} g_k^i(\tilde{\mathbf{x}}_l^i) + g_k^i(\mathbf{x}_k^*) - \frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial \mathbf{a}_k^i} \Delta \mathbf{a}_k^{i\top} - \frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial \mathbf{b}_k^i} \Delta \mathbf{b}_k^{i\top}, \quad (13)$$

em que

$$\frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial \mathbf{a}_k^i} \Delta \mathbf{a}_k^{i\top} = \left[ \frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial a_1^i} \dots \frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial a_p^i} \right] \begin{bmatrix} a_{k,1}^i - a_{k-1,1}^i \\ \vdots \\ a_{k,p}^i - a_{k-1,p}^i \end{bmatrix}$$

$$\frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial \mathbf{b}_k^i} \Delta \mathbf{b}_k^{i\top} = \left[ \frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial b_1^i} \dots \frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial b_p^i} \right] \begin{bmatrix} b_{k,1}^i - b_{k-1,1}^i \\ \vdots \\ b_{k,p}^i - b_{k-1,p}^i \end{bmatrix}$$

Logo, para a  $j$ -ésima dimensão dos dados, os termos derivativos em relação a  $\mathbf{a}_k$  e  $\mathbf{b}_k$  são computados da seguinte forma:

$$\frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial a_j^i} = g_k^i(\tilde{\mathbf{x}}_l^i) \left[ \frac{\sum_{i=1}^{N_k} \omega^i(\tilde{\mathbf{x}}_l^i) \Lambda^i(\mathcal{F}^i)^{-1}}{\sum_{i=1}^{N_k} \omega^i(\tilde{\mathbf{x}}_l^i)} - \Lambda^i(\mathcal{F}^i)^{-1} \right]$$

$$\frac{\partial g_k^i(\tilde{\mathbf{x}}_l^i)}{\partial b_j^i} = -g_k^i(\tilde{\mathbf{x}}_l^i) \left[ \frac{\sum_{i=1}^{N_k} \omega^i(\tilde{\mathbf{x}}_l^i) \Lambda^i(\mathcal{F}^i)^{-1}}{\sum_{i=1}^{N_k} \omega^i(\tilde{\mathbf{x}}_l^i)} - \Lambda^i(\mathcal{F}^i)^{-1} \right]$$

sendo

$$\Lambda^i = \frac{2(\tilde{x}_{lj}^i - m_{kj}^i)^2}{(b_{kj}^i - a_{kj}^i)^3} \quad \text{e} \quad \mathcal{F}^i = \left[ \sum_{j=1}^p \frac{(\tilde{x}_{lj}^i - m_{kj}^i)^2}{(b_{kj}^i - a_{kj}^i)^2} \right]^{\frac{1}{2}}.$$

Além disso, o termo  $\sum_{l=1}^{|\mathcal{G}^i|+1} g_{k+1}^i(\tilde{\mathbf{x}}_l^i)$  denota o valor que a soma cumulativa dos graus de pertinência das amostras de  $\mathcal{G}^i$  terá no próximo instante de tempo caso  $\mathbf{x}_k^*$  seja efetivamente admitida. De (13), pode-se verificar que o somatório dos graus de pertinência dos dados é função da variação dos parâmetros da função de pertinência dada a admissão de  $\mathbf{x}_k^*$ .

Com isso, caso o valor de  $Q_k^*(\mathcal{G}^i)$ , que corresponde ao valor do rendimento  $Q_k(\mathcal{G}^i)$  calculado após a admissão de  $\mathbf{x}_k^*$ , seja maior do que o valor antes da admissão, os parâmetros  $\mathbf{m}_k^i$ ,  $\mathbf{a}_k^i$  e  $\mathbf{b}_k^i$  são atualizados. Caso contrário, estes parâmetros permanecem inalterados.

#### 4.4 Realização Algorítmica

O algoritmo de granulação em tempo real apresentado ao longo desta seção é apresentado resumidamente no Algoritmo 1.

---

**Algoritmo 1** Procedimento de granulação em tempo real

---

```

Defina  $N_1 = 1$ ;  $\leftarrow$  criação do primeiro grânulo
para  $k = 1, \dots, p$  faça
  leia  $\mathbf{x}_k$ ;
  calcule  $\mathbf{m}_k^1$ ,  $\mathbf{a}_k^1$  e  $\mathbf{b}_k^1$  segundo (6), (9) e (10), respectivamente;
fim
faça anom = matriz vazia;
faça  $N_a = 0$ ;  $\leftarrow$  Número de anomalias contínuas
para  $k \geq p + 1$  faça
  leia  $\mathbf{x}_k$ ;
  faça anomalia = 1;
  faça atualizar = 0;
  para  $1 \leq i \leq N_k$  faça
    se  $M^i(\mathbf{x}_k) \leq \ell_1$  então
      faça atualizar = 1;
      faça anomalia = 0;
    fim
  senão
    se  $|\mathcal{G}^i| > T_s$  e  $M^i(\mathbf{x}_k) \leq \ell_2^*$  então
      faça atualizar = 1;
    fim
  fim
  compute  $Q_k^*(\mathcal{G}^i)$  usando (13);
  se atualizar = 1 e  $Q_k^*(\mathcal{G}^i) > Q_k(\mathcal{G}^i)$  então
    atualize o grânulo  $\mathcal{G}^i$  em termos de  $\mathbf{m}_k^i$ ,  $\mathbf{a}_k^i$  e  $\mathbf{b}_k^i$  segundo (6), (9) e (10), respectivamente;
  fim
fim
se anomalia = 1; então
  adicione  $\mathbf{x}_k$  à anom;
   $N_a = N_a + 1$ ;
fim
senão
   $N_a = 0$ ;
fim
se Tracker é c-separado de todo  $\mathcal{G}^i$  e  $N_a > 10$  então
  crie novo grânulo com dados em anom;
  faça anom=matriz vazia;
   $N_a = 0$ ;
fim
fim

```

---

\*  $\ell_1$  e  $\ell_2$ , são limiares como descrito em (5), mas que podem apresentar intervalos de confiança iguais ou distintos. Para este artigo, foi admitido o valor de 0.95 para ambos.

## 5. GRANULAÇÃO DE SÉRIES TEMPORAIS E ANÁLISE DE COMPROMISSO

Nesta seção, serão apresentados exemplos de granulação de séries temporais considerando a formulação de grânulos elipsoidais evolutivos. Em relação à maximização do suporte das regras *fuzzy*, isto é, dos semieixos das elipses ao longo da dimensão dos dados, dois procedimentos serão realizados:

- (1) Verificação das condições de atualização para todos os grânulos (alocação compartilhada);
- (2) Verificação das condições de atualização apenas para o grânulo em que a amostra  $\mathbf{x}_k$  possui maior grau de pertinência.

O principal objetivo destas verificações é analisar o efeito de cada procedimento na granularidade do sistema evolutivo e, conseqüentemente, na semântica da representação granular.

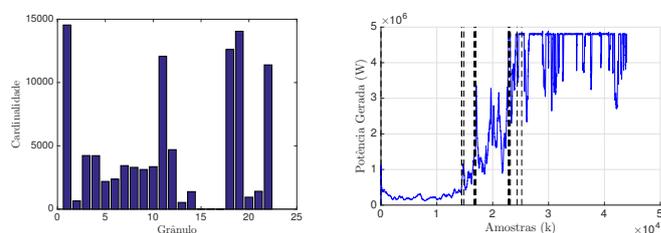
Para a verificação de cada um dos procedimentos de granulação descritos anteriormente, foi utilizado o *benchmark* de turbina eólica proposto por Odgaard et al. (2013)

usualmente empregado para avaliação de estratégias de diagnóstico e isolamento de falhas (Bessa et al., 2016) e controle tolerante a falhas (Odgaard e Stoustrup, 2015). Nos experimentos apresentados a seguir, os algoritmos de granulação propostos serão aplicados às séries temporais provenientes de sensores do *benchmark* considerando a ocorrência e a ausência de falhas. O *benchmark* em questão dispõe de dois sensores redundantes *m1 em2* para medição do ângulo de arfagem para cada pá ( $\beta_{1,m1}$ ,  $\beta_{1,m2}$ ,  $\beta_{2,m1}$ ,  $\beta_{2,m2}$ ,  $\beta_{3,m1}$  e  $\beta_{3,m2}$ ), de velocidades angulares no rotor ( $\omega_{r,m1}$  e  $\omega_{r,m2}$ ) e gerador ( $\omega_{g,m1}$  e  $\omega_{g,m2}$ ), medições de velocidade do vento ( $v_{hub}$ ), torque gerado ( $\tau_{g,m}$ ) e potência gerada ( $P_{g,m}$ ). Todas as medidas ainda que obtidas considerando o modelo físico e computacional da turbina eólica, os dados de entrada relativos à velocidade do vento são provenientes de medições reais (Odgaard et al., 2013).

Para o cenário de ocorrência de falhas, são consideradas as seguintes falhas em sensores

- Falha 1 ( $\mathcal{F}_1$ ): do tipo incipiente, trata-se de uma falha que começa a ocorrer entre as amostras 4000 até 4999, é total entre as amostras 5000 e 8000, e retorna gradativamente à normalidade entre as amostras 8001 até 9000. Essa falha afeta simultaneamente o sensor 1 da velocidade angular do gerador ( $\omega_{g,m1}$ ) e o sensor 2 da velocidade angular do rotor ( $\omega_{r,m2}$ ), tal que o primeiro apresenta velocidade medida 10% acima da velocidade real e o último apresenta atenuação de 10% da velocidade real.
- Falha 2 ( $\mathcal{F}_2$ ): do tipo abrupta, trata-se de falha no sensor 1 do ângulo de arfagem da pá 1 ( $\beta_{1,m1}$ ) da turbina eólica. O valor medido se torna constante e igual a 5° entre as amostras 34000 e 39000.

Na Figura 1(a), é apresentada a cardinalidade de cada grânulo para o caso de ausência de falhas. As amostras são distribuídas por entre todos os grânulos à medida em que o rendimento como indicado em (11) tem sua condição de atualização atendida. Desta forma, uma amostra, de acordo com o valor de sua contribuição (12), pode alterar o rendimento de mais de um grânulo. Na prática, o que ocorre é a maximização do suporte das funções de pertinência que compõem a base de regra *fuzzy* granular. As representações obtidas durante o processo de granulação, tem seu instante de criação representados pelas linhas tracejadas verticais na Figura 1(b).

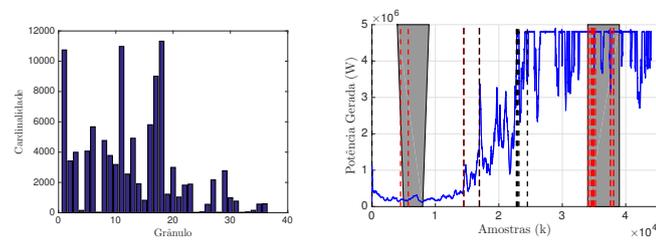


(a) Cardinalidade dos grânulos. (b) Pontos de criação dos grânulos.

Figura 1. Algoritmo de granulação em tempo real aplicado ao conjunto de dados sem falha considerando o Procedimento 1.

Com a ocorrência de falhas, pode-se verificar um aumento de aproximadamente 61% na quantidade de grânulos criados em relação ao caso sem falhas. Na Figura 2(a), pode-se

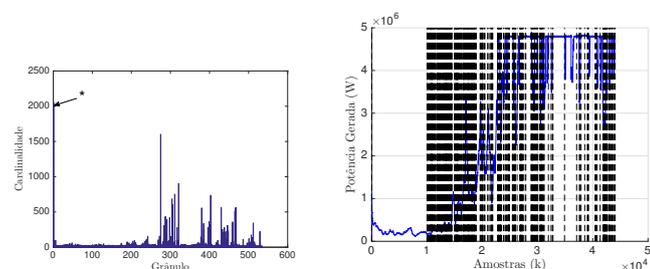
observar que a cardinalidade dos grânulos se altera devido a criação de novos grânulos nas regiões de ocorrência das falhas ( $\mathcal{F}_1$ ) e ( $\mathcal{F}_2$ ), representadas em cinza. A criação de grânulos nesta região está possivelmente relacionada à alterações nas séries temporais das variáveis causadas pela ocorrência de novidades. Esta constatação indica a possibilidade de detectar pontos de mudança em séries temporais.



(a) Cardinalidade dos grânulos. (b) Pontos de criação dos grânulos.

Figura 2. Algoritmo de granulação em tempo real aplicado ao conjunto de dados com falha considerando o Procedimento 1.

Considerando o procedimento 2, há um aumento na quantidade de grânulos em ambos os casos de ausência e ocorrência de falhas, como mostrado na Tabela 1. As amostras são distribuídas ao longo de grânulos com baixa cardinalidade e picos de cardinalidade alta como pode ser observado nas Figuras 3(a) e 3(b). No que se refere ao instante de criação dos grânulos, há alterações que podem ser percebidas nas regiões em que as falhas ocorrem. No caso da falha  $\mathcal{F}_1$ , verificou-se a criação de um número muito maior de grânulos quando comparado com o mesmo caso utilizando o procedimento 1. O mesmo é notado para o caso da falha  $\mathcal{F}_2$ .

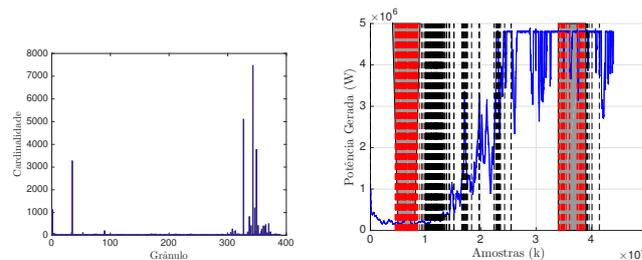


(a) Cardinalidade dos grânulos. (b) Pontos de criação dos grânulos.

Figura 3. Algoritmo de granulação em tempo real aplicado ao conjunto de dados sem falha considerando o Procedimento 2. Note que em (\*) a cardinalidade do grânulo 1 foi multiplicada por um fator de 0.2 para fins de escalabilidade para visualização.

Os resultados da granulação utilizando os dois procedimentos propostos são sumarizados na Tabela 1. É evidente que tanto para o caso sem falha, quanto com falha, o procedimento 1 levou a um número muito menor de grânulos do que o procedimento 2.

Em termos de interpretabilidade, a análise será realizada considerando dois aspectos: interpretabilidade baseada em complexidade (número de grânulos) e interpretabilidade baseada em semântica (cobertura dos grânulos) (Gacto



(a) Cardinalidade dos grânulos. (b) Pontos de criação dos grânulos.

Figura 4. Algoritmo de granulação em tempo real aplicado ao conjunto de dados com falha considerando o Procedimento 2.

Tabela 1. Número de grânulos obtido com os diferentes procedimentos.

Procedimento	Sem Falha	Com Falha
1	22	36
2	534	390
Aumento relativo	24.27	10.83

et al., 2011). A complexidade do modelo gerado pelo procedimento 2 é muito maior quando comparada com o procedimento 1, já que uma quantidade maior de grânulos foi criada para cobrir os dados. Por consequência, grânulos com baixa cardinalidade e alta especificidade foram obtidos. Em situações como esta, a interpretabilidade pode ser comprometida pelo fato do modelo apresentar partições não discerníveis e com baixa cobertura. Por outro lado, no procedimento 1, o fato da complexidade do modelo ser menor implica na criação de grânulos com maior cobertura e menor especificidade, sendo discerníveis entre si. Neste caso, a interpretabilidade do modelo é maior pois os grânulos formados apresentam semântica melhor definida, facilitando conclusões a respeito do significado das representações.

## 6. CONCLUSÃO

Neste trabalho, foi proposta uma metodologia de granulação em tempo real de séries temporais. A granulação foi baseada no PGJ, considerando-se uma medida quantitativa para alocação ótima de granularidade. Com isso, garantiu-se que os limites dos grânulos com formato elipsoidal fossem atualizados de modo que o suporte das funções de pertinência fossem maximizados em termos de cobertura e semântica. Além disso, realizou-se uma análise do compromisso entre granularidade e interpretabilidade de sistemas granulares evolutivos.

Para avaliar experimentalmente a relação de compromisso entre granularidade e interpretabilidade, foram considerados dois procedimentos de criação de grânulos. O primeiro levou em conta a influência da amostra disponibilizada pelo fluxo de dados na atualização dos índices de desempenho de todos os grânulos já criados, enquanto que o segundo foi baseado na atualização do índice de desempenho do grânulo o qual a amostra disponibilizada possui maior grau de pertinência.

O procedimento de alocação compartilhada de granularidade favorece a interpretabilidade dos grânulos, pois produziu um modelo com menor granularidade e maior

interpretabilidade que o segundo, já que grânulos com maior cobertura e menor especificidade foram obtidos, apresentando semântica melhor definida.

Além disso, com os resultados observados a partir da granulação do conjunto de dados da turbina eólica, verificou-se a capacidade do método proposto de indicar a ocorrência de novidades durante o processamento da série temporal que apresentava a ocorrência de falhas. Como trabalho futuro, essa capacidade deverá ser melhor explorada a fim de desenvolver estratégias de detecção e diagnóstico de falhas baseadas em granulação evolutiva.

## REFERÊNCIAS

- Ahmed, M.M. Isa, N.A.M. (2015). Information granularity model for evolving context-based fuzzy system. *Applied Soft Computing*, 33, 183–196.
- Antonelli, M., Ducange, P., Lazzerini, B., Marcelloni, F. (2011). Learning concurrently data and rule bases of mamdani fuzzy rule-based systems by exploiting a novel interpretability index. *Soft Computing*, 15(10), 1981–1998.
- Bessa, I.V., Palhares, R.M., D'Angelo, M.F.S.V., Chaves Filho, J.E. (2016). Data-driven fault detection and isolation scheme for a wind turbine benchmark. *Renewable Energy*, 87, 634 – 645.
- Costa, B.S.J., Bezerra, C.G., Guedes, L.A., Angelov, P.P. (2016). Unsupervised classification of data streams based on typicality and eccentricity data analytics. In *Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on*, 58–63. IEEE.
- Dasgupta, S. (1999). Learning mixtures of Gaussians. In *40th Annual Symposium on Foundations of Computer Science*, 634–644. IEEE.
- Gacek, A. Pedrycz, W. (2015). Clustering granular data and their characterization with information granules of higher type. *IEEE Transactions on Fuzzy Systems*, 23(4), 850–860.
- Gacto, M., Alcalá, R., Herrera, F. (2011). Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. *Information Sciences*, 181(20), 4340 – 4360. Special Issue on Interpretable Fuzzy Systems.
- Leite, D., Ballini, R., Costa, P., Gomide, F. (2012). Evolving fuzzy granular modeling from nonstationary fuzzy data streams. *Evolving Systems*, 3(2), 65–79.
- Leite, D., Palhares, R.M., Campos, V.C., Gomide, F. (2015). Evolving granular fuzzy model-based control of nonlinear dynamic systems. *IEEE Transactions on Fuzzy Systems*, 23(4), 923–938.
- Lemos, A., Caminhas, W., Gomide, F. (2011). Multivariable gaussian evolving fuzzy modeling system. *IEEE Transactions on Fuzzy Systems*, 19(1), 91–104.
- Liu, Q. Liu, Q. (2002). Approximate reasoning based on granular computing in granular logic. In *Proceedings. International Conference on Machine Learning and Cybernetics*, volume 3, 1258–1262 vol.3.
- Liu, S., Pedrycz, W., Gacek, A., Dai, Y. (2017). A two-phase method of forming a granular representation of signals. *Signal Processing*, 141, 1 – 15.
- Liu, S., Pedrycz, W., Gacek, A., Dai, Y. (2018). Development of information granules of higher type and their applications to granular models of time series. *Engineering Applications of Artificial Intelligence*, 71, 60 – 72.
- Mencar, C. Fanelli, A. (2008). Interpretability constraints for fuzzy information granulation. *Information Sciences*, 178(24), 4585 – 4618.
- Moshtaghi, M., Bezdek, J.C., Leckie, C., Karunasekera, S., Palaniswami, M. (2015). Evolving fuzzy rules for anomaly detection in data streams. *IEEE Transactions on Fuzzy Systems*, 23(3), 688–700.
- Moshtaghi, M., Leckie, C., Bezdek, J.C. (2016). Online clustering of multivariate time-series. In *Proceedings of the 2016 SIAM International Conference on Data Mining*, 360–368.
- Odgaard, P.F. Stoustrup, J. (2015). A benchmark evaluation of fault tolerant wind turbine control concepts. *IEEE Transactions on Control Systems Technology*, 23(3), 1221–1228. doi:10.1109/TCST.2014.2361291.
- Odgaard, P.F., Stoustrup, J., Kinnaert, M. (2013). Fault-tolerant control of wind turbines: A benchmark model. *IEEE Transactions on Control Systems Technology*, 21(4), 1168–1182.
- Pedrycz, W. Wang, X. (2016). Designing fuzzy sets with the use of the parametric principle of justifiable granularity. *IEEE Transactions on Fuzzy Systems*, 24(2), 489–496.
- Pedrycz, W. Gomide, F. (2007). *Fuzzy systems engineering: toward human-centric computing*. John Wiley & Sons.
- Pedrycz, W. Homenda, W. (2013). Building the fundamentals of granular computing: A principle of justifiable granularity. *Applied Soft Computing*, 13(10), 4209 – 4218.
- Ren, H., Li, X., Li, Z., Pedrycz, W. (2018). Data representation based on interval-sets for anomaly detection in time series. *IEEE Access*, 6, 27473–27479.
- Silva, A.M., Caminhas, W., Lemos, A., Gomide, F. (2014). A fast learning algorithm for evolving neo-fuzzy neuron. *Applied Soft Computing*, 14, 194 – 209. *Evolving Soft Computing Techniques and Applications*.
- Wang, G., Shi, P., Wang, B., Zhang, J. (2014). Fuzzy  $n$ -ellipsoid numbers and representations of uncertain multichannel digital information. *IEEE Transactions on Fuzzy Systems*, 22(5), 1113–1126.
- Wang, X., Pedrycz, W., Gacek, A., Liu, X. (2016). From numeric data to information granules: A design through clustering and the principle of justifiable granularity. *Knowledge-Based Systems*, 101, 100 – 113.
- Yao, Y. (2018). Three-way decision and granular computing. *International Journal of Approximate Reasoning*, 103, 107 – 123.
- Yin, S., Jiang, Y., Tian, Y., Kaynak, O. (2017). A data-driven fuzzy information granulation approach for freight volume forecasting. *IEEE Transactions on Industrial Electronics*, 64(2), 1447–1456.
- Zhu, X., Pedrycz, W., Li, Z. (2017). Granular data description: Designing ellipsoidal information granules. *IEEE Transactions on Cybernetics*, 47(12), 4475–4484.