XXI Escola Regional de Redes de Computadores / VIII Workshop de Segurança (ERRC). DOI: 10.5753/errc.2024.4584

ARTIGO COMPLETO/FULL PAPER

Network Anomaly Detection using Choquet Integrals through Power Measures

Detecção de Anomalias de Rede utilizando Integrais de Choquet através de Medidas de Potência

Denner Ayres • dennerayres@furg.br
Universidade Federal do Rio Grande (FURG)

Abreu Quevedo • abreurg@furg.br
Universidade Federal do Rio Grande (FURG)

Graçaliz Dimuro • gracaliz@furg.br
Universidade Federal do Rio Grande (FURG)

Giancarlo Lucca • giancarlo.lucca@ucpel.edu.br
Universidade Católica de Pelotas (UCPel)

Bruno L. Dalmazo • dalmazo@furg.br
Universidade Federal do Rio Grande (FURG)

RESUMO. Este artigo apresenta uma avaliação da detecção de anomalias utilizando integrais de Choquet e métricas de medidas de potência. O objetivo principal foi verificar a efetividade do modelo de detecção considerando diferentes dimensões de janelas deslizantes. A análise dos erros mostrou que as métricas de medição de potência apresentaram desempenho superior, especialmente em janelas deslizantes de menor dimensão. Com a implementação de um sistema de simulação de ataques, o modelo demonstrou maior precisão e eficiência em cenários com janelas reduzidas, aprimorando a detecção de ataques. Os resultados indicam que a estratégia baseada em janelas deslizantes menores é mais adequada para contextos de alta volatilidade e mudanças abruptas no tráfego.

ABSTRACT. This paper evaluates anomaly detection using Choquet integrals and power measure metrics. The main goal was to check the effectiveness of the detection model by considering different sizes of sliding windows. Error analysis showed that the power measure metrics performed better, especially with smaller sliding windows. With the implementation of an attack simulation system, the model showed higher accuracy and efficiency in scenarios with reduced windows, improving the detection of attacks. The results suggest that the strategy based on smaller sliding windows is more suitable for contexts with high volatility and sudden traffic changes.

PALAVRAS-CHAVE: Detecção de anomalias • Integrais de Choquet • Medidas de potência • Janelas deslizantes
KEYWORDS: Anomaly detection • Choquet integrals • Power measures • Sliding windows

1 Introdução

As redes de computadores desempenham um papel essencial, garantindo acesso ágil e confiável a recursos digitais, sendo fundamentais tanto para o setor empresarial quanto para atividades cotidianas. Com o crescimento exponencial de dispositivos interconectados e sensores, há um fluxo contínuo de dados provenientes de múltiplas fontes e contextos. Assim, assegurar o monitoramento e a gestão eficaz desses dados torna-se imprescindível. Diversos esforços têm sido empreendidos para aprimorar a eficiência e a segurança das informações que trafegam nas redes, tema que vem sendo amplamente estudado [1].

Na sua essência, a Internet vai além de ser apenas um meio de comunicação, ela se estabeleceu como um direito humano essencial no mundo atual. A disponibilidade de acesso à internet é fundamental para o exercício de outros direitos humanos, tais como a liberdade de expressão, o acesso à informação e a participação política. Em um mundo progressivamente digital, a ausência de acesso à Internet não só restringe as chances de interação e aprendizado, como também marginaliza pessoas e comunidades, impedindo-as de exercer plenamente seus direitos como cidadãos. Portanto, reconhecer a Internet como um direito humano significa assegurar que todos possam utilizar esse recurso essencial de forma livre, não supervisionada e sem censura, fomentando a igualdade e a dignidade humana em uma sociedade interligada [2].

Devido à sua relevância, as redes de computadores são frequentemente alvos de ataques. Muitos desses ataques deixam vestígios na rede, o que possibilita sua

detecção. Nesse contexto, diversas técnicas têm sido criadas para mitigar os impactos negativos dessas ações maliciosas, com destaque para as abordagens de detecção de anomalias, que visam identificar comportamentos atípicos e prevenir danos [3, 4]. Este trabalho compara modelos de médias móveis com fraca dependência histórica de dados, comumente utilizados para fazer previsões de tráfego de rede sob esse tipo de restrição [5]. Após avaliar as médias móveis e encontrar o modelo com menor erro na predição, utilizaremos esse modelo para detectar anomalias no tráfego de rede e comparar seu desempenho com uma função de agregação de dados baseada na integral de Choquet.

O restante desse trabalho está organizado como segue. A Seção 2 apresenta os trabalhos relacionados. A proposta sobre detecção de anomalias será apresentada na Seção 3. Os testes e avaliações serão apresentados na Seção 4. Finalizando com as considerações finais na Seção 5.

2 Trabalhos Relacionados

Esta seção apresenta alguns dos trabalhos mais relevantes no campo da detecção de anomalias em redes de computadores e integrais de Choquet, ressaltando os principais benefícios e limitações de cada abordagem [6, 7].

Os métodos convencionais de identificação de anomalias baseados em densidade, amplamente empregados, têm se mostrado eficientes em várias circunstâncias. No entanto, esses métodos têm restrições consideráveis ao serem empregados em dados incertos e imprecisos, pois presumem uma uniformidade nos dados que nem sempre é verificada na realidade. Para ultrapassar essas restrições, apresentamos um método fuzzy-rough para identificar anomalias, que combina as teorias fuzzy e rough para gerenciar incertezas e ambiguidades nos dados. A densidade fuzzy-rough é definida como o nível de agregação das instâncias ao redor de um ponto, levando em conta a incerteza inerente aos dados. A partir dessa definição, criamos um sistema de pontuação que atribui pontuações de anomalia com base em fuzzy-rough[8].

Aplicação de mais de 70 algoritmos de detecção de anomalias em mais de 900 conjuntos de dados de séries temporais possibilitou uma análise completa das várias técnicas disponíveis [3]. Foram escolhidos algoritmos de diversas famílias e métodos, assegurando uma representação integral das técnicas de identificação de anomalias. A análise dessas técnicas desvendou semelhanças interessantes, como a eficácia das médias móveis, particularmente em contextos de baixa dependência histórica [5].

Por fim, destaca-se a importância do monitoramento da infraestrutura de redes para a detecção de anomalias, bem como o papel da Função de Agregação Difusa nesse contexto. As implicações práticas deste estudo sugerem que futuras pesquisas podem explorar novos modelos e técnicas para a detecção de anomalias, especialmente em cenários de tráfego de rede. Um caminho promissor envolve o uso de integrais de choque, que mostraram potencial para melhorar a detecção de anomalias em tráfego de redes, trazendo avanços significativos para a área [9].

3 Uma proposta para detecção de anomalias utilizando uma medida de potência

Esse artigo almeja oferecer uma alternativa à detecção de anomalias clássicas, utilizando integrais de Choquet e aplicando métrica de medida de potência para atribuição de pesos. A abordagem proposta busca otimizar a detecção em diferentes tamanhos de janelas deslizantes, ajustando o modelo para aumentar a acurácia e a eficiência em diversos cenários. Nessa seção apresentamos conceitos importantes para o entendimento da nossa proposta assim como o formalismo matemático, o *dataset* utilizado na validação e detalhamos nosso modelo conceitual (Figura 1).

3.1 Modelo PMA

O modelo PMA, ou Média Móvel de Poisson (*Poisson Moving Average*), é um tipo de modelo de média móvel utilizado para a previsão do tráfego de rede. Este modelo baseia-se na distribuição de Poisson, para atribuição de pesos conforme os valores de entrada da janela deslizante. Considerando o parâmetro t = λ , a distribuição define o tamanho da janela, e os valores recentes do tráfego são ponderados por uma distribuição de Poisson.

$$PMA = \sum_{i=1}^{\lambda} P_i V_{t-1}$$

De acordo com estudos [1], o modelo PMA demonstra o menor erro na previsão do tráfego de rede em comparação a outros modelos de média móvel, como a Média Móvel Simples (SMA), a Média Móvel Ponderada (WMA) e a Média Móvel Exponencial (EMA).

3.2 Medidas Fuzzy

As medidas fuzzy [10] são um conceito central dentro da teoria da agregação e da lógica difusa, sendo utilizadas para representar e quantificar incertezas e a importância relativa entre elementos. Em particular,

essas medidas são fundamentais para a agregação de dados, onde é necessário avaliar a relação entre diferentes elementos com base em sua relevância ou contribuição para um conjunto maior. A importância de uma coalizão ou subconjunto de elementos é modelada por meio dessas medidas, permitindo a integração de diferentes níveis de importância de maneira flexível e adaptativa.

Considerando $N = \{1, \ldots, n\}$. Uma função $\mathfrak{m} : 2^N \to [0, 1]$ é chamada de medida fuzzy se, para todos $X, Y \subseteq N$, as seguintes condições forem satisfeitas:

(*m*1) Crescimento: se $X \subseteq Y$, então $\mathfrak{m}(X) \leq \mathfrak{m}(Y)$;
(*m*2) Condições de contorno: $\mathfrak{m}(\emptyset) = 0$ e $\mathfrak{m}(N) = 1$.

Neste artigo trabalhamos com a Medida de Potência (do inglês Power Measure – PM). Uma medida proposta [11] ao qual é adaptada para diferentes classes em problemas de classificação[1]. Para $A \subseteq N$, a medida utilizada é definida como:

$$\mathfrak{m}(A) = \left(\frac{|A|}{n}\right)^q, \text{ com } q > 0. \tag{1}$$

3.3 Modelo Choquet

O modelo de Integral de Choquet, é uma abordagem matemática utilizada para a agregação de dados, especialmente em contextos onde a ordenação dos valores de entrada são importantes.

$$C_m(x) = \sum_{i=1}^{n} (x_{(\mathrm{i})} - x_{(\mathrm{i\text{-}1})}).m(A_{(\mathrm{i})}), \tag{2}$$

aonde $(x_{(\mathrm{i})}, \ldots, x_{(\mathrm{n})})$ is an increasing permutation of the input x, and $A_{(\mathrm{i})} = \{(i), \ldots, (n)\}$ is the subset of weights assigned in the same way as the PMA model, but following the ordering of the subset of input values.

3.4 *Dataset*

Neste estudo, foi utilizada uma versão processada do *dataset* CIC-DDoS2019, desenvolvido pelo *Canadian Institute for Cybersecurity*. Esse *dataset* contém uma grande quantidade de dados rotulados, fundamentais para o treinamento de algoritmos de aprendizado de máquina, com uma ampla gama de ataques DDoS. O destaque deste trabalho é o uso de 29.404 amostras de tráfego de rede coletadas ao longo de 24 horas.

3.5 Modelo Conceitual

Nossa solução foca em fornecer uma abordagem sistemática para detecção de anomalias. A Figura 1 mostra a base da nossa solução, destacando seus principais componentes conceituais e suas interações. A seguir, serão apresentados detalhes sobre os componentes.

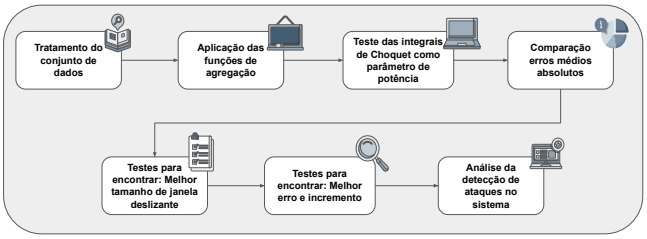


Figura 1. Modelo conceitual

- *Tratamento do conjunto de dados*: Nesse ponto os dados são pré-processados e configurados de maneira sistemática para que possam ser computados pelas funções de agregação.
- *Aplicação das funções de agregação*: Nesse ponto, as funções são testadas em seu modelo PMA e na sua metodologia de Choquet para analisar seu comportamento.
- *Teste das integrais de Choquet como parâmetro de potência*: Através dessa etapa, as funções de Choquet passam a ser testadas também com funções de parâmetro de potencia.
- *Comparação erros médios absolutos*: É realizado os teste dos erros absolutos de todas medidas de potência e em seguida é realizado a comparação das melhores medidas de potências comparada as metodologias de Choquet e PMA.
- *Testes para encontrar: Melhor tamanho de janela deslizante*: Através da obtenção dos melhores trechos de janela deslizantes, nessa etapa, o melhor trecho é testado em análise de sensibilidade de parâmetros, para encontrar o tamanho de janela deslizante ideal.
- *Testes para encontrar: Melhor erro e incremento*: A partir do tamanho ideal de janela deslizante, esse valor é fixado, e os valores de Erro e Incremento são simulados e avaliados. Através desse teste é possível encontrar o melhor parâmetro para o teste da base de dados real.
- *Análise da detecção de ataques no sistema*: Com os três parâmetros principais definidos, é possível então analisar a detecção dos ataques de forma eficaz.

4 Avaliação

Para analisar o comportamento dos resultados de acordo com os tamanhos de janela, todos os testes a

[1] Veja uma comparação de medidas fuzzy e integral de Choquet em

Tabela 1. Tabela funções para o parâmetro P

Identificador	Funções
(a)	$T_{Max}(v) = \max\{v\}$
(b)	$T_{Min}(v) = \min\{v\}$
(c)	$T_{Prd}(v) = \sum_{i=1}^{n} V_i$
(d)	$T_{Med}(v) = \frac{1}{n}\sum_{i=1}^{n} V_i$
(e)	$T_{MH}(v) = \frac{1}{\sum_{i=1}^{n} \frac{1}{V_i}}$
(f)	$T_{MG}(v) = \sqrt[n]{\prod_{i=1}^{n} v_i}$
(g)	$T_{Moda}(v) = \text{argmax}_v\, f(v)$
(h)	$T_{Mediana}(v) = \begin{cases} v_{(\frac{n+1}{2})} & \text{se ímpar} \\ \frac{v_{(\frac{n}{2})}+v_{(\frac{n}{2}+1)}}{2} & \text{se par} \end{cases}$

seguir foram realizado em análise de sensibilidade de parâmetros, ou seja, os resultados foram armazenados em vetores conforme os valores de variação das janelas deslizantes.

4.1 Erros de Predições

A primeira avaliação trata da validação do erro médio absoluto. Por se tratar de um sistema preditivo, quanto menor o erro apresentado mais precisa é a abordagem. A abordagem deste teste consiste em analisar o erro médio absoluto conforme o tamanho da janela deslizante aumenta, como mostrado na Tabela 1. O teste foi realizado utilizando 9 modelos diferentes, aplicados a diversas funções de medidas de potência. Todas as medidas de potência obtiveram resultados similares de erro. Porém, a métrica mínima foi a que demonstrou os melhores resultados através do tamanho de janela conforme o gráfico ilustrado na Figura 2a.

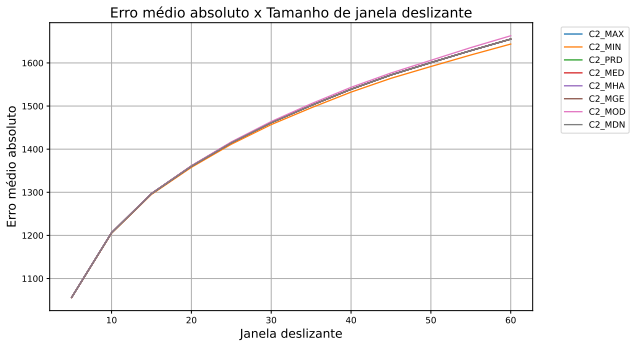
Através disso podemos observar a comparação dos erros com os modelos já propostos, que são o PMA e a integral de Choquet com *Ordered Weighted Average* (OWA). Analisando então os erros das predições entre as três propostas podemos ver através da Figura 2b.

Através desse gráfico é possível notar que a metologia de Choquet, que até então, não havia conseguido superar a metologia PMA em nenhum momento. Entretanto, através das medidas de potências, isso se torna possível para janelas deslizantes de tamanho pequeno. Ou seja, os próximos testes práticos passam a ser analisados para os menores tamanhos de janela.

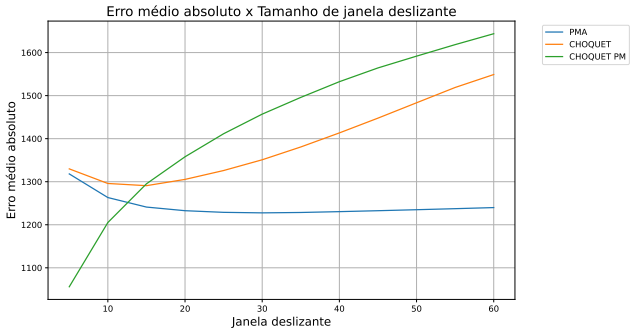
4.2 Detecção de Ataques

A metodologia adotada para determinar o *threshold* é baseada em dois parâmetros fundamentais para sua

problemas de classificação em [12].



(a) Erro das medidas de potências



(b) Erro dos modelos PMA, Choquet e Medida de potência

Figura 2. Erros de Predições

implementação: *Erro* e *Incremento*. O valor do *Erro* é empregado como um elemento chave na identificação de irregularidades, onde, quando o valor de entrada excede um certo percentual do valor original, um alerta é gerado.

O segundo indicador, o *Incremento*, estabelece a continuidade dessas alterações ao longo do tempo. Para que um agrupamento de alertas seja categorizado como um ataque, é imprescindível que eles aconteçam de maneira contínua, caracterizada pelo número de incrementos. Este parâmetro é crucial para prevenir falsos positivos, assegurando que apenas variações consistentes e com um comportamento contínuo sejam identificadas como ataques.

Através de testes realizados com erro de predições, foi identificado o tamanho ideal de janela fixa para o modelo. Utilizando essa janela fixa, diversos testes em análise de sensibilidade de parâmetros foram realizados, variando os parâmetros de *Erro* e *Incremento* para encontrar os valores mais adequados ao comportamento dos dados. Nos experimentos, o valor de *Erro* foi ajustado de 65% a 95%. Simultaneamente, o parâmetro *Incremento*, foi modificado de 3 para 12. Esta estratégia de variação possibilitou a identificação dos valores ideais de *Erro* e *Incremento* para este conjunto de dados específico.

A Tabela 2 apresenta os valores de ataques e alertas

Tabela 2. Detecções de cada função

Função	Alertas	Ataques
Geral	15594	1012
PMA	16926	893
Choquet	16942	899
Choquet PM	16080	1022

preditos em cada um dos modelos em comparação do modelo geral, que possui valores padrões do *dataset*, em relação aos demais modelos. Observamos que a metodologia que utiliza as medidas de potência é a que mais se aproximou da quantidade de ataques do sistema real. A partir desses valores, chegamos à de acurácia de 94% e uma taxa de especificidade de 96%.

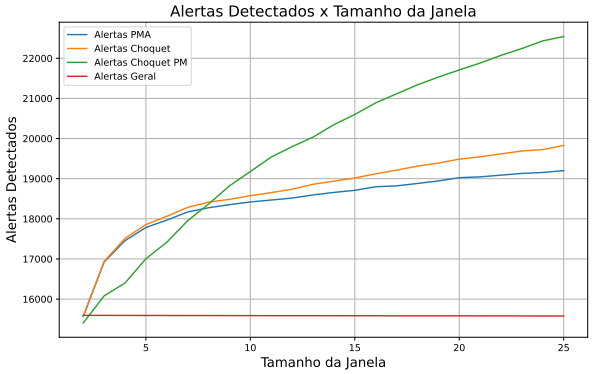
5 Considerações Finais

Os gráficos apresentados mostram os resultados da análise do comportamento da detecção ao variar o tamanho das janelas deslizantes. Conforme a Figura 3b, os dados indicam que o modelo de medidas de potência se mostra mais eficaz em janelas curtas, mostrando um desempenho superior na identificação dos eventos anômalos. Essa configuração permite que o modelo se aproxime dos valores reais, refletindo uma maior precisão na detecção.

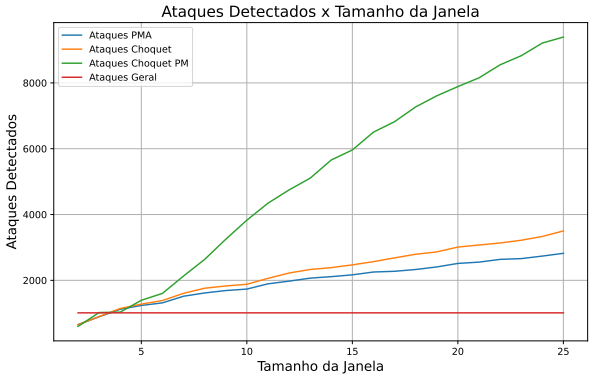
No entanto, é importante notar conforme Figura 3a, que apesar da eficiência nas janelas pequenas, o sistema tende a gerar um número excessivo de alarmes. Tal resultado pode acabar afetando a sensibilidade do sistema de predição. Nossa proposta se torna mais interessante ainda por não necessitar de treinamento prévio da base de dados, permitindo usar uma abordagem de detecção em tempo real, pois consegue se ajustar diretamente ao ambiente sem a necessidade de buscar parâmetros ideais para cada contexto. Essa característica torna nossa proposta candidata a ser utilizada onde o tempo de resposta é crítico.

Com base nos resultados obtidos, o sistema demonstrou uma acurácia de 94%, refletindo a alta precisão na classificação geral dos dados, e uma taxa de especificidade de 96%, evidenciando sua eficiência em identificar corretamente os casos negativos e reduzir falsos positivos. Esses indicadores confirmam que a metodologia utilizada, especialmente a que emprega medidas de potência, é eficaz e se aproxima de forma significativa dos dados reais, tornando-a uma abordagem promissora para detecção de ataques e alertas em sistemas de tempo real.

Em trabalhos futuros, o objetivo é melhorar a detecção de alarmes para diminuir o número de falsos positivos. Além disso, pretendemos investigar a utilização do modelo em outras situações de medida de potência, o que pode aumentar sua confiabilidade e generalização para ser empregado em variados contextos, assim como melhorar a metodologia de *threshold*.



(a) Alertas detectados



(b) Ataques detectados

Figura 3. Erros de predições

Declarações complementares

Financiamento

Os autores gostariam de agradecer à FAPERGS (24/2551-0001396-2) e à FAPERGS/CNPq (23/2551-0000126-8;23/2551-0000773-8)

Referências

1 Dalmazo, B. L.; Vilela, J. P.; Curado, M. Performance analysis of network traffic predictors in the cloud. *Journal of Network and Systems Management*, Springer, v. 25, p. 290–320, 2017.

2 Reglitz, M. The human right to free internet access. *Journal of Applied Philosophy*, Wiley Online Library, v. 37, n. 2, p. 314–331, 2020.

3 Zeufack, V. *et al.* An unsupervised anomaly detection framework for detecting anomalies in real time through network system's log files analysis. *High-Confidence Computing*, Elsevier, v. 1, n. 2, p. 100030, 2021.

4 Santo, Y. *et al.* Fault Detection on the Edge and Adaptive Communication for State of Alert in Industrial Internet of Things. *Sensors*, v. 23, n. 7, 2023. ISSN 1424-8220. DOI: 10.3390/s23073544. Disponível em: https://www.mdpi.com/1424-8220/23/7/3544.

5 Dalmazo, B. L.; Vilela, J. P.; Curado, M. Triple-Similarity Mechanism for alarm management in the cloud. *Computers & Security*, v. 78, p. 33–42, 2018. ISSN 0167-4048. DOI: https://doi.org/10.1016/j.cose.2018.05.016. Disponível em: https://www.sciencedirect.com/science/article/pii/S0167404818306515.

6 Amorim, M. *et al.* Systematic Review of Aggregation Functions Applied to Image Edge Detection. *Axioms*, v. 12, n. 4, 2023. ISSN 2075-1680. DOI: 10.3390/axioms12040330. Disponível em: https://www.mdpi.com/2075-1680/12/4/330.

7 Cardoso, F. C. *et al.* Echo state network and classical statistical techniques for time series forecasting: A review. *Knowledge-Based Systems*, v. 293, p. 111639, 2024. ISSN 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2024.111639. Disponível em: https://www.sciencedirect.com/science/article/pii/S0950705124002740.

8 Yuan, Z. *et al.* Anomaly detection based on weighted fuzzy-rough density. *Applied Soft Computing*, v. 134, p. 109995, 2023. ISSN 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2023.109995. Disponível em: https://www.sciencedirect.com/science/article/pii/S1568494623000133.

9 Ayres, D. *et al.* Comparando Médias Móveis com Integral de Choquet para Detectar Anomalias no Tráfego de Redes. *In*: ANAIS Estendidos do XXIV Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais. São José dos Campos/SP: SBC, 2024. P. 353–357. DOI: 10.5753/sbseg_estendido.2024.243381. Disponível em: https://sol.sbc.org.br/index.php/sbseg_estendido/article/view/30154.

10 Murofushi, T.; Sugeno, M.; Machida, M. Non-monotonic fuzzy measures and the Choquet integral. *Fuzzy Sets and Systems*, v. 64, n. 1, p. 73–86, 1994.

11 Barrenechea, E. *et al.* Using the Choquet Integral in the Fuzzy Reasoning Method of Fuzzy Rule-Based Classification Systems. *Axioms*, v. 2, n. 2, p. 208–223, 2013.

12 Lucca, G. *et al.* Analyzing the performance of different fuzzy measures with generalizations of the Choquet integral in classification problems. *In*: 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). Jun. 2019. P. 1–6. DOI: 10.1109/FUZZ-IEEE.2019.8858815.