

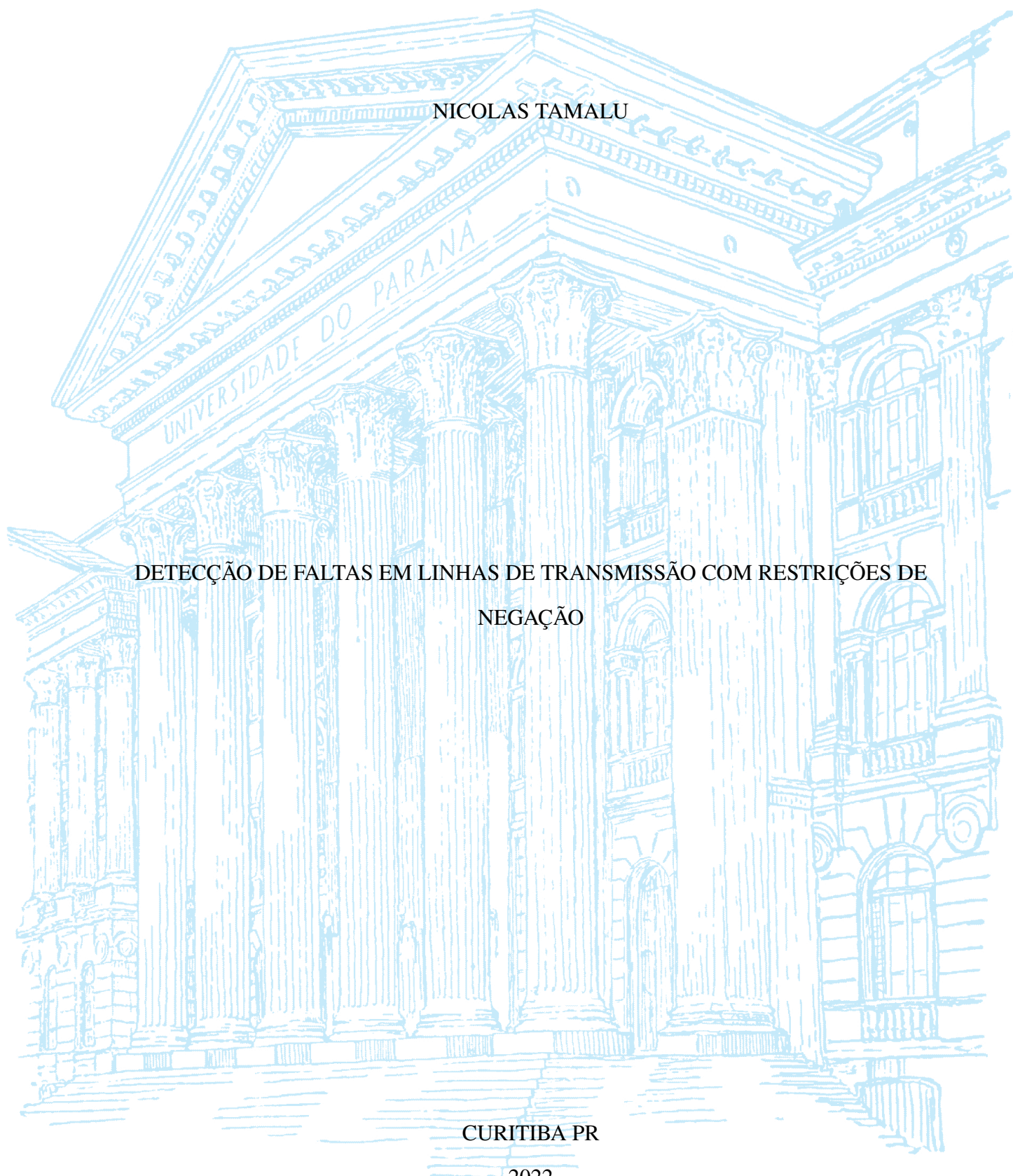
UNIVERSIDADE FEDERAL DO PARANÁ

NICOLAS TAMALU

DETECÇÃO DE FALTAS EM LINHAS DE TRANSMISSÃO COM RESTRIÇÕES DE
NEGAÇÃO

CURITIBA PR

2022



NICOLAS TAMALU

DETECÇÃO DE FALTAS EM LINHAS DE TRANSMISSÃO COM RESTRIÇÕES DE
NEGAÇÃO

Trabalho apresentado como requisito parcial à conclusão do Curso de Bacharelado em Ciência da Computação, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Prof. Dr. Eduardo Cunha de Almeida.

CURITIBA PR

2022

RESUMO

Um Sistema Elétrico de Potência (SEP) é a estrutura responsável pela geração, transmissão e distribuição de energia elétrica. Dentro dessa estrutura, as linhas de transmissões são as responsáveis por transmitir eletricidade entre usinas geradoras e subestações de distribuição. As linhas de transmissão normalmente possuem grandes extensões, o que as deixam mais vulneráveis à adversidades como condições climáticas e vandalismo, e as tornam o principal foco de falhas em um SEP. Essas adversidades podem perturbar seu estado normal de operação e ocasionar faltas (falhas) elétricas que por sua vez podem causar a interrupção do fornecimento de energia elétrica em locais críticos como hospitais e unidades de saúde. Assim, diversos estudos são realizados na área, tendo como principais objetivos a detecção, a classificação e a localização de faltas. Neste trabalho, temos como objetivo principal desenvolver um método capaz de detectar faltas utilizando restrições de negação. Restrições de negação são um formalismo lógico que expressam relações entre tuplas de um conjunto de dados. A base de dados utilizada tem simulações que contém oscilografias de corrente e tensão em períodos pré-falta e pós-falta. o método desenvolvido propõe uma maneira de dividir esses dados para realizar a descoberta de restrições de negação com o algoritmo DCFinder, destaque no estado da arte, e um modo de utilizar as restrições encontradas com janelas deslizantes para detectar faltas. Os resultados obtidos com o método foram bons, as melhores restrições de negação foram capazes de detectar falta em todas as simulações selecionadas e obtiveram uma precisão de 100% nas detecções. A divisão feita nos dados facilitou a descoberta de restrições de negação reduzindo a quantidade delas e possibilitando a execução de etapas posteriores.

Palavras-chave: Linha de Transmissão. Detecção de Falta. Restrição de Negação. Série Temporal.

LISTA DE FIGURAS

2.1	Visão geral de um SEP	11
2.2	Principais componentes de uma linha de transmissão	12
2.3	Exemplo de falta	12
2.4	<i>Framework</i> simplificado para detecção, classificação e localização de faltas . . .	14
3.1	Esquema de execução do DCFinder	22
3.2	Arquitetura da plataforma Metanome.	23
4.1	Diagrama esquemático do processo de janela deslizante	25
5.1	IEEE 9-barras modelado no ATPDraw	26
6.1	Loteamento das simulações para descoberta de restrições	30
6.2	Particionamento do espaço de predicados para a descoberta de restrições.	31
6.3	Representação da execução do DCFinder sobre uma simulação loteada e particio- nada	32
6.4	Representação do processo de avaliação com janela deslizante	34
7.1	Média do F1-score das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$	38
7.2	Média da precisão das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$	39
7.3	Média do <i>recall</i> das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$	40
7.4	Média de amostras após a falta para a primeira detecção de falta das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$	41
7.5	Média de simulações que não foram detectadas faltas das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$	42
7.6	Médias do F1-score das RNs com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ relacionadas com a média de suas coberturas	44
7.7	Médias do F1-score das RNs com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ relacionadas com a quantidade vezes que cada uma foi encontrada em lotes com falta	45
7.8	Médias do F1-score das RNs com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ relacionadas com a quantidade vezes que cada uma foi encontrada em lotes sem falta.	46

7.9	Médias do F1-score das RNs com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ relacionadas com seus comprimentos	47
-----	--	----

LISTA DE TABELAS

1.1	Exemplo de instância de um esquema de funcionários	9
5.1	Exemplo reduzido de uma simulação	27
7.1	Tempo médio para obter o grau de aproximação de cada janela e classificá-la . .	40
7.2	Médias das métricas de desempenho das 10 melhores RNs do par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ para os tempos de início das faltas	41
7.3	Médias das métricas de desempenho das 10 melhores RNs do par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ para os tipos de faltas . . .	42
7.4	Médias das métricas de desempenho da melhor RN para cada tipo de falta com o par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$	43
7.5	Médias das métricas de desempenho das 10 melhores RNs do par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ para as resistências da linha de transmissão.	43
7.6	Médias das métricas de desempenho das 10 melhores RNs do par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ para as localizações das faltas	43
A.1	Restrições de negação, suas métricas na descoberta e F1-score médio com tamanho de janela $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$	51
B.1	Restrições de negação e suas médias nas métricas de desempenho com tamanho da janela de $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$	56

SUMÁRIO

1	INTRODUÇÃO	8
1.1	OBJETIVOS	9
1.2	DESAFIOS	9
1.3	HIPÓTESE	10
1.4	ASSERTIVAS	10
1.5	ESTRUTURA DO DOCUMENTO	10
2	LINHAS DE TRANSMISSÃO	11
2.1	LINHAS DE TRANSMISSÃO	11
2.2	FALTAS EM LINHAS DE TRANSMISSÃO	12
2.2.1	Causas das Faltas	13
2.2.2	Categorias e Tipos de Faltas	13
2.2.3	Análise de Faltas de Curto-Circuito	14
2.3	CONSIDERAÇÕES FINAIS	15
3	RESTRIÇÕES DE NEGAÇÃO	16
3.1	DEFINIÇÃO	16
3.1.1	Restrição de Negação Aproximada	17
3.2	EXEMPLOS	17
3.2.1	Exemplo de Restrição de Negação Exata	18
3.2.2	Exemplo de Restrição de Negação Aproximada	18
3.2.3	Exemplo de Dependência Funcional como Restrição de Negação	18
3.2.4	Exemplo de Dependência de Ordem como Restrição de Negação	19
3.2.5	Exemplo de Combinação Única de Colunas como Restrição de Negação	19
3.3	AVALIAÇÃO DE RESTRIÇÕES DE NEGAÇÃO	20
3.3.1	Sucintez	20
3.3.2	Cobertura	20
3.4	DESCOBERTA DE RESTRIÇÕES DE NEGAÇÃO	21
3.4.1	DCFinder	22
3.4.2	Metanome	22
3.5	CONSIDERAÇÕES FINAIS	23
4	SÉRIES TEMPORAIS	24
4.1	DEFINIÇÃO	24
4.2	JANELA DESLIZANTE	24
4.2.1	Janela em Cascata	24
4.3	CONSIDERAÇÕES FINAIS	25

5	BASE DE DADOS COM SIMULAÇÕES DE FALTA	26
5.1	EXEMPLO	27
5.2	CONSIDERAÇÕES FINAIS	28
6	MÉTODO	29
6.1	DESCOBERTA DE RESTRIÇÕES DE NEGAÇÃO.	29
6.1.1	Seleção dos Dados	29
6.1.2	Loteamento dos Dados Seleccionados	29
6.1.3	Particionamento dos Atributos dos Dados Seleccionados	30
6.1.4	Descoberta de Restrições de Negação com o DCFinder	30
6.2	AVALIANDO O DESEMPENHO DAS RESTRIÇÕES DE NEGAÇÃO DESCOBERTAS	30
6.2.1	Seleção dos Dados	30
6.2.2	Janela Deslizante e Classificação	31
6.2.3	Parâmetros Variados	33
6.3	HARDWARE	35
6.4	CONSIDERAÇÕES FINAIS	35
7	RESULTADOS E DISCUSSÃO	36
7.1	RESTRIÇÕES DE NEGAÇÃO DESCOBERTAS	36
7.1.1	Processo de Descoberta de Restrições de Negação	36
7.1.2	Detalhamento das Restrições de Negação Descobertas em Lotes Sem Falta.	37
7.2	RESULTADOS DAS AVALIAÇÕES.	38
7.2.1	Variação dos Parâmetros da Janela	38
7.2.2	Variação dos Parâmetros das Simulações.	41
7.2.3	Métricas das Restrições	44
8	CONCLUSÃO	48
8.1	SUGESTÕES DE TRABALHOS FUTUROS	48
	REFERÊNCIAS	49
	APÊNDICE A – RESTRIÇÕES DE NEGAÇÃO DESCOBERTAS	51
	APÊNDICE B – RESTRIÇÕES DE NEGAÇÃO E SUAS MÉTRICAS DE DESEMPENHO NA DETECÇÃO DE FALTAS	56

1 INTRODUÇÃO

Um Sistema Elétrico de Potência (SEP) é a estrutura formada por um conjunto de instalações e equipamentos que realiza a geração, transmissão e distribuição de energia elétrica. Neste sistema, as linhas de transmissão desempenham um papel crucial, pois são os componentes responsáveis pela transmissão de energia elétrica entre usinas geradoras e subestações responsáveis pela distribuição (Bichels, 2018)

As linhas de transmissão costumam percorrer longas distâncias e, devido a isso, estão expostas à situações adversas como condições climáticas e queimadas, o que a torna muito suscetível a ocorrência de faltas (falhas), perturbações ao seu estado normal de operação. E de fato isso acontece, de acordo com Kindermann (2018), 89% das falhas em um SEP ocorrem nas linhas de transmissão.

A ocorrência de falhas pode provocar problemas como a interrupção do fornecimento de energia elétrica seguida de apagões em efeito dominó. Esses problemas são graves, pois podem afetar locais como hospitais e unidades de saúde, onde a energia elétrica é indispensável. Além disso, esses problemas geram insatisfação aos consumidores e multas podem ser aplicadas às concessionárias.

Assim, os estudos sobre faltas em linhas de transmissão possuem grande relevância. Comumente, esses estudos buscam desenvolver métodos para a realização de três atividades em particular: detecção, classificação e localização das faltas. A detecção de faltas, normalmente, atua como gatilho para a execução dos processos de classificação e localização. Além disso, as três atividades são importantes para que os operadores possam tomar decisões relacionadas ao religamento da linha de transmissão e para a avaliação posterior do funcionamento do sistema de proteção.

Os dados utilizados neste trabalho são amostras das oscilografias das correntes e tensões de uma linha de transmissão simulada com períodos pré-falta e pós-falta. No mundo real, esses dados podem ser obtidos através de Registradores Digitais de Perturbações que são capazes de capturar e armazenar os registros das oscilografias dessas grandezas elétricas.

As três principais técnicas propostas para a realização das atividades são: ondas viajantes, impedância e Inteligência Artificial (Chen et al., 2016). Diferente delas, neste trabalho propomos o uso de restrições de negação (RNs), que fornecem um formalismo lógico formado por predicados que determinam relações entre as tuplas de uma base de dados (Pena et al., 2020), para a detecção de faltas.

Como exemplo, considere a Tabela 1.1 que contém os dados de funcionários de uma empresa. Nela, a descrição de uma possível restrição de negação seria: se dois funcionários têm o mesmo cargo e o mesmo nível, então eles devem ter o mesmo salário. Analisando a tabela, notamos que essa restrição não é válida, pois o funcionário representado pela tupla t_2 têm o mesmo cargo e nível dos funcionários das tuplas t_0 e t_1 , mas um salário diferente, o que viola a restrição. Caso isso não ocorresse, diríamos que a restrição de negação descrita é válida.

Para utilizarmos RNs na detecção de falta, primeiro é necessário descobrir quais são as restrições dos dados, para isso utilizaremos o algoritmo DCFinder. A descoberta de restrições de negação é um processo computacionalmente caro (Bleifuß et al., 2017), por isso iremos dividir os dados de maneira que o processo seja mais rápido e também para tornar possível distinguir restrições de negação descobertas em períodos pré-falta e pós-falta.

Tabela 1.1: Exemplo de instância de um esquema de funcionários

	id	nome	sobrenome	cargo	nivel	horas	valor_hora	salario
t ₀	20	João	Silva	Desenvolvedor	Junior	40	15.0	2400.00
t ₁	21	Maria	Clara	Desenvolvedor	Junior	40	15.0	2400.00
t ₂	22	João	Ferreira	Desenvolvedor	Junior	44	15.0	2640.00
t ₃	23	Mateus	Santos	Desenvolvedor	Pleno	40	25.0	4000.00
t ₄	24	Gabriela	Oliveira	Desenvolvedor	Pleno	40	25.0	4000.00
t ₅	25	Pedro	Souza	Analista de RH	Estagiário	30	10.0	1200.00
t ₆	26	Clara	Santos	Analista de RH	Sênior	44	30.0	5280.00

1.1 OBJETIVOS

O principal objetivo do trabalho é propor e avaliar um método para a detecção de faltas em linhas de transmissão com restrições de negação que contém duas etapas: descobrir as restrições de negação e utilizá-las para a detecção.

Os objetivos secundários deste trabalho são:

- Propor um método de divisão dos dados que auxilie o processo de descoberta de restrições de negação;
- Propor um método que detecte faltas de maneira eficiente e independente das características da falta;
- Avaliar a relação entre o desempenho das restrições de negação na detecção de faltas e suas métricas de cobertura e sucintez.

1.2 DESAFIOS

Os principais desafios do trabalho são desenvolver um método de divisão dos dados para que se consiga realizar a descoberta das restrições de negação de maneira que possam ser utilizadas para a detecção de faltas e desenvolver um método que utilize as restrições de negação para detectar faltas com diferentes características de modo eficiente e preciso.

1.3 HIPÓTESE

As hipóteses do trabalho são que: os dados de períodos pré-falta possuem restrições de negação diferentes dos dados de períodos pós-falta e que a partir disso é possível utilizar essas restrições para a detecção de faltas.

1.4 ASSERTIVAS

As assertivas do trabalho são:

- Assertiva 1: a tarefa de detecção de falta pode ser realizada a partir das oscilografias de corrente e tensão;
- Assertiva 2: as oscilografias de corrente e tensão podem ser obtidas no mundo real através de registradores digitais de perturbação;

- Assertiva 3: a detecção de falta é primordial pois auxilia os operadores no tratamento da falha e atua como gatilho para as tarefas de classificação e localização, além disso seu desempenho repercute diretamente nessas outras duas tarefas.

1.5 ESTRUTURA DO DOCUMENTO

A partir dessa seção, o documento está estruturado da seguinte maneira:

- Capítulo 2: conceitos básicos sobre linhas de transmissão e faltas;
- Capítulo 3: definição de restrição de negação, métricas, exemplos e algoritmo utilizado para descobrir as restrições;
- Capítulo 4: conceitos básicos sobre séries temporais e maneiras de percorrê-las;
- Capítulo 5: apresentação dos dados utilizados;
- Capítulo 6: detalhamento do processo de descoberta de restrições de negação e detecção de faltas;
- Capítulo 7: análise e discussão dos resultados obtidos;
- Capítulo 8: conclusão e sugestões de trabalhos futuros.

2 LINHAS DE TRANSMISSÃO

Um Sistema Elétrico de Potência (SEP) é definido por Bichels (2018) como o conjunto de instalações e equipamentos que atuam nas áreas de geração, transmissão e distribuição de energia elétrica. Na Figura 2.1 está representada a visão geral de um SEP.

As três áreas citadas são compostas por distintos tipos de componentes:

- Geração: usinas hidrelétricas, termelétricas, eólicas e solares;
- Transmissão: linhas e subestações de transmissão;
- Distribuição: redes e subestações de distribuição.

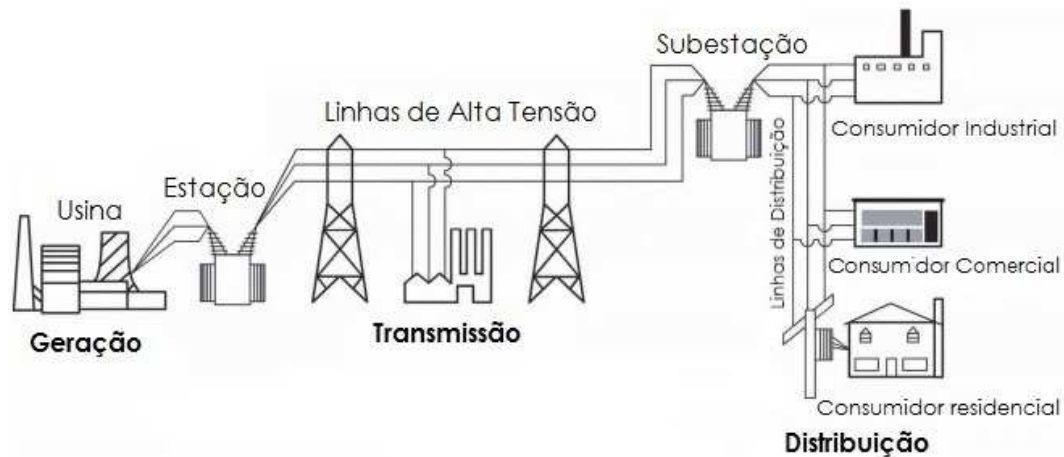


Figura 2.1: Visão geral de um SEP

Fonte: F. da Silva (2014)

2.1 LINHAS DE TRANSMISSÃO

As linhas de transmissão são os elementos responsáveis pela transmissão de energia entre os centros de produção e os de distribuição, que levarão essa energia até seus consumidores, como indústrias, comércio e residências. As tensões de operação em linhas de transmissão são elevadas, podendo chegar até 750kV, em corrente alternada, e até 800kV, corrente contínua (Bichels, 2018), por isso também podem ser conectadas diretamente a consumidores de maior porte (Fuchs, 1977).

As linhas de transmissão são predominantemente trifásicas, simétricas e equilibradas, conforme representadas pelos condutores de fase A, B e C da Figura 2.2 (Bichels, 2018). Outros componentes típicos de linhas aéreas também estão na Figura 2.2: cabos condutores, isoladores, cabos pára-raios e estruturas para a sustentação dos cabos.

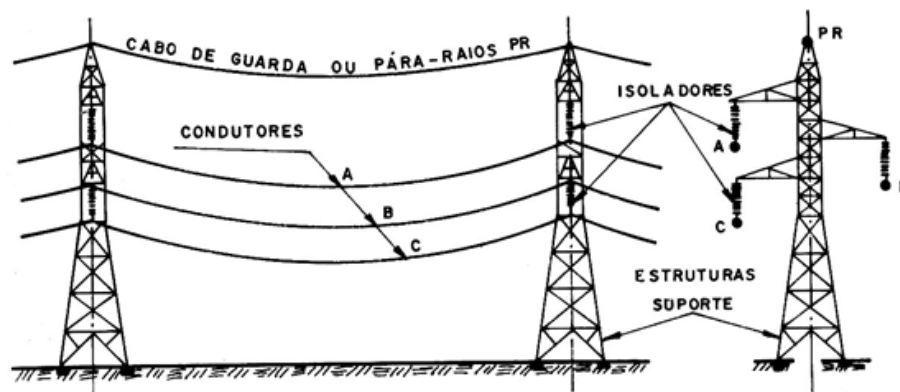


Figura 2.2: Principais componentes de uma linha de transmissão
Fonte: Labegalini et al. (1992)

2.2 FALTAS EM LINHAS DE TRANSMISSÃO

Uma falta (falha) pode ser definida como um fluxo anormal das grandezas elétricas, como tensão e corrente, nos componentes do sistema de potência (Mishra e Ray, 2018).

A Figura 2.3 está exemplificando um registro de oscilografia para o sinal da corrente em uma linha trifásica quando ocorre uma falta. No período de pré-falta (onde ainda não ocorreu uma falta) que vai do instante de tempo 0 até próximo a 0,1 segundo no exemplo, os valores para a corrente são equilibrados e simétricos, como descrito anteriormente. Em seguida, inicia-se um fluxo anormal caracterizando uma falta, prosseguindo para um período de pós-falta.

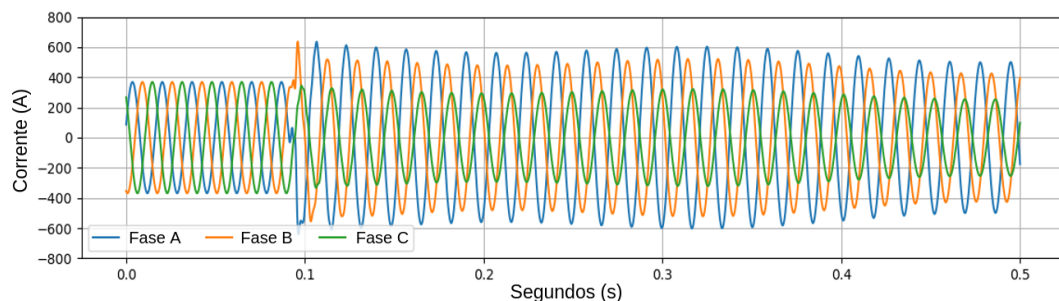


Figura 2.3: Exemplo de falta

Devido a sua exposição as diversas condições climáticas, meteorológicas e os mais variados terrenos que enfrenta, as linhas de transmissão são os componentes mais vulneráveis à falhas e, dentre todo o setor de sistema elétrico, 89% do curto-circuitos ocorrem nelas, segundo Kindermann (2018).

2.2.1 Causas das Faltas

A escolha das características dos componentes citados anteriormente, como: especificação dos cabos condutores, tipos de isoladores e dimensões da estrutura, busca satisfazer a viabilidade econômica e aspectos técnicos, tais quais: serviço contínuo e tensão e corrente uniformes e constantes (Fuchs, 1977).

Porém, mesmo projetando o sistema visando a maior qualidade possível, faltas são inevitáveis por diversos fatores, como as citadas por Kindermann (2018):

- Problemas de isolamento: projeto inadequado, desgaste do material, produto de baixa qualidade e defeitos de fabricação;
- Problemas mecânicos: ventanias, árvores, neve e contaminação;
- Problemas elétricos: descarga elétrica atmosférica, sobretensão no sistema e surtos de chaveamento;
- Problemas de natureza térmica: sobrecorrentes e sobretensão dinâmica do sistema;
- Problemas de manutenção: substituição inadequada, mão-de-obra não qualificada e inspeção deficiente;
- Problemas de outra natureza: vandalismo, queimadas e outros.

2.2.2 Categorias e Tipos de Faltas

As faltas podem ser divididas em duas categorias: circuito aberto (*open circuit*) e curto-circuito. A primeira categoria é originada pelo rompimento de um cabo condutor ou a abertura de disjuntores e isoladores. Já a segunda pode ser causada por diversos motivos como condições atmosféricas adversas e contaminação do isolador (Yadav e Dash, 2014; Raza et al., 2020).

As faltas de curto-circuito podem ser separadas em duas outras subcategorias: simétricas e assimétricas. As simétricas são aquelas que envolvem as três fases (A, B e C), podendo incluir ou não o terra (G). Já as assimétricas são aquelas envolvem uma fase e o terra, duas fases isoladamente ou, ainda, duas fases e o terra (Yadav e Dash, 2014; Raza et al., 2020).

Com isso, também podemos agrupar as faltas de acordo com o número de fases envolvidas:

- Monofásicas: AG, BG e CG;
- Bifásicas: AB, AC, BC, ABG, ACG, e BCG;
- Trifásicas: ABC e ABCG.

As faltas devido a um curto-circuito são o foco desse trabalho, pois, são as mais perigosas, visto que o aumento na intensidade da corrente durante a falha pode prejudicar as linhas de transmissão mecanicamente, devido a força imposta pelo fenômeno, deformando condutores e rompendo materiais isolantes, e termicamente, pelo calor gerado, deteriorando cabos e outros equipamentos (Yadav e Dash, 2014; Sato e Freitas, 2015).

2.2.3 Análise de Faltas de Curto-Circuito

Comumente, os métodos propostos na literatura especializada para a análise de falhas empregam oscilografias dos sinais da corrente e da tensão, os quais são obtidos a partir de registradores digitais de perturbação (Raza et al., 2020).

Os estudos realizados no campo de análise de faltas buscam desenvolver métodos para alcançar três objetivos em particular: detecção, classificação e localização. A detecção é a identificação da ocorrência de uma falta, a classificação é a especificação de qual tipo de falta ocorreu, e a localização é a determinação do local da linha de transmissão que a falta aconteceu.

Esses três objetivos se relacionam, como pode ser visto na Figura 2.4. O processo de localização pode utilizar da classificação da falta para obter um melhor desempenho. A detecção de falta, normalmente, atua como um gatilho para que os processos de classificação e localização sejam ativados, o que lhe confere grande importância. Além disso, a precisão da detecção do tempo em que a falta ocorre é de grande importância, pois repercute diretamente no desempenho dos métodos de classificação e localização.

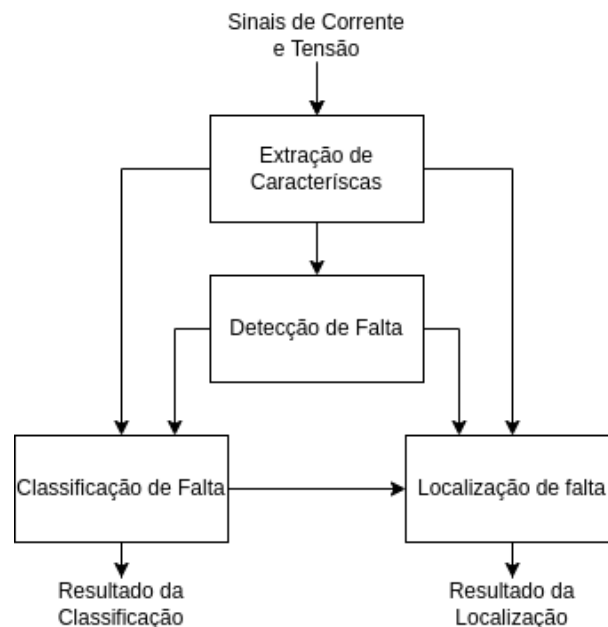


Figura 2.4: *Framework* simplificado para detecção, classificação e localização de faltas

Fonte: adaptado de (Chen et al., 2016)

2.3 CONSIDERAÇÕES FINAIS

Como vimos, as linhas de transmissão têm um papel fundamental dentro dos sistemas elétricos, também são um dos componentes que mais sofrem com curto-circuitos e estão diretamente ligadas a maior parte dos problemas de interrupção no fornecimento de energia. Todos esses fatores fazem com que estudos relacionados as faltas que ocorrem nelas tenham bastante relevância. Esses estudos se dividem em três tópicos: detecção, classificação e localização, e contam com diversas técnicas para realizar análises. O tópico a seguir será sobre restrições de negação, um formalismo lógico que será utilizado em nosso método para a detecção de faltas.

3 RESTRIÇÕES DE NEGAÇÃO

A qualidade de dados é um importante campo de estudo que investiga aspectos como consistência, deduplicação, completude de informação, atualização e acurácia de dados. De acordo com Fan (2015, tradução nossa): "A consistência de dados refere-se a validade e a integridade dos dados que representam entidades do mundo real". A verificação da consistência é, normalmente, realizada por meio da checagem de violações de dependência de dados (ou restrições de integridade). As dependências de dados são propriedades e relações que os dados devem satisfazer para que sejam uma representação válida do mundo real. A existência de registros que não estão de acordo com elas caracterizam uma violação de integridade (Fan, 2015; Pena et al., 2020).

Existem diversos tipos de dependências, cada uma delas possui uma maneira expressar propriedades e relações que os dados devem atender, por exemplo: dependências funcionais, dependências de ordem e combinações de colunas únicas. Além disso, também temos as restrições de negação (RNs), um tipo de dependência capaz de generalizar as citadas anteriormente e expressar outras restrições (Bleifuß et al., 2017; Pena et al., 2020).

Além da sua expressividade e capacidade de generalizar outras dependências, as RNs proporcionam fácil tradução para uma consulta em *Structured Query Language* (SQL) e suas aplicações em áreas como a limpeza e verificação da consistência de dados tornam essas restrições merecedoras de atenção (Chu et al., 2013).

3.1 DEFINIÇÃO

Uma restrição de negação é um formalismo lógico formado por predicados que determinam relações entre tuplas que são inconsistentes.

Considere uma instância r com n tuplas de um esquema R com atributos (colunas) $A_1, A_2, \dots, A_m$ e seja t_0 e t_1 tuplas de r , A_c e A_d atributos de R e o um operador de comparação de $O = \{<, <=, >, >=, =, \neq\}$. Dessa maneira, um predicado p tem a forma: $t_0.A_c o t_1.A_d$.

A partir disso, definimos uma restrição de negação como:

$$\varphi : \forall t_0, t_1 \in r, \neg(p_1 \wedge p_2 \wedge \dots \wedge p_m) \quad (3.1)$$

Assim, uma restrição φ é satisfeita se e somente se para todo par de tuplas $(t_0, t_1) \in r$ houver ao menos um predicado p_i falso em seu conjunto de predicados, denotado $\varphi.Predc$.

Vale observar que os atributos A_c e A_d de um predicado podem ser iguais, também é possível realizar comparações com constantes e relacionar mais de duas tuplas em uma mesma restrição, porém, nesse trabalho, estamos interessados apenas em restrições de negação variáveis (sem constantes) e com duas tuplas, devido ao algoritmo que iremos utilizar.

Outro conceito importante é o de restrição de negação mínima. Uma RN φ_1 que satisfaz uma instância r é mínima se não existe outra RN φ_2 tal que também satisfaz r e $\varphi_2.Predc \subset \varphi_1.Predc$. Ou seja, o conjunto de predicados de φ_1 não possui qualquer outro subconjunto que satisfaça r .

3.1.1 Restrição de Negação Aproximada

Uma restrição de negação aproximada é aquela que contém uma quantidade de violações abaixo de um determinado limite: se o limite máximo for igual a 0, então é uma restrição de negação exata (Pena et al., 2019).

O limite de erro é o valor máximo aceitável do grau de aproximação da restrição. O grau de aproximação pode ser medido de três maneiras diferentes (Livshits et al., 2020):

1. Calculando a quantidade de pares de tuplas que violam a restrição de negação e dividindo pela quantidade total de pares de tuplas;
2. Calculando a quantidade de tuplas distintas que violam a restrição de negação e dividindo pela quantidade total de tuplas;
3. Calculando a quantidade mínima de tuplas que devem ser removidas para que a restrição seja satisfeita e dividindo pela quantidade total de tuplas.

Neste trabalho, a primeira medida será utilizada. Assim, o grau de aproximação de uma restrição φ para instância r é dado pela Equação 3.2:

$$g(\varphi, r) = \frac{|\{(t_x, t_y) \in r \mid (t_x, t_y) \not\models \varphi\}|}{|r| \cdot (|r| - 1)} \quad (3.2)$$

onde $(t_x, t_y) \not\models \varphi$ significa que (t_x, t_y) não satisfaz φ . Da mesma maneira, $(t_x, t_y) \models \varphi$, significa que (t_x, t_y) satisfaz φ .

Ou seja, o grau de aproximação de uma RN é a quantidade de tuplas que violam a restrição dividido pela quantidade de pares de tuplas.

Assim, podemos definir: uma restrição de negação aproximada φ é válida para uma instância r se e somente se dado um ϵ , $0 \leq \epsilon < 1$, o seu grau de aproximação $g(\varphi, r)$ é menor que ϵ (Pena et al., 2019).

3.2 EXEMPLOS

Nessa seção, utilizaremos para os exemplos a instância r , representada pela Tabela 1.1, e com um esquema R que possui os atributos: *id*, *nome*, *cargo*, *nível*, *horas* e *salario*.

Para a detecção de faltas, serão utilizados registros que possuem atributos relacionados à corrente e tensão.

3.2.1 Exemplo de Restrição de Negação Exata

Seja t_x, t_y tuplas de r , uma restrição de negação dessa instância é:

$\varphi_a : \forall t_x, t_y \in r, \neg(t_x.cargo = t_y.cargo \wedge t_x.nivel = t_y.nivel \wedge t_x.horas = t_y.horas \wedge t_x.salario \neq t_y.salario)$,

ou seja, se dois empregados têm o mesmo cargo ($t_x.cargo = t_y.cargo$), o mesmo nível ($t_x.nivel = t_y.nivel$) e trabalham a mesma quantidade de horas semanais ($t_x.horas = t_y.horas$), então eles devem ter o mesmo salário. Caso tenham um salário diferente ($t_x.salario \neq t_y.salario$), a restrição estará sendo violada.

Como dito antes, uma RN pode ser traduzida para uma *query* SQL e detectar as violações. A consulta SQL correspondente a RN φ_b :

```

1 select distinct tx.*
2 from funcionarios as tx, funcionarios as ty
3 where tx.cargo = ty.cargo and
4 tx.nivel = ty.nivel and
5 tx.horas = ty.horas and
6 tx.salario != ty.salario;

```

3.2.2 Exemplo de Restrição de Negação Aproximada

Um exemplo de restrição de negação aproximada válida para essa instância para um limite de erro de 0,10 é a descrita no Capítulo 1:

$$\varphi_b : \forall t_x, t_y \in r, \neg(t_x.cargo = t_y.cargo \wedge t_x.nivel = t_y.nivel \wedge t_x.salario \neq t_y.salario)$$

Essa RN é semelhante à anterior, com a diferença de que o número de horas trabalhadas não importa. Assim, o seu grau de aproximação está abaixo do limite de erro, ou seja, é uma restrição de negação aproximada válida para aquele limite. O cálculo feito para o grau de aproximação:

$$g(\varphi_b, r) = \frac{4}{42} \approx 0,095 < 0,10,$$

onde o valor do numerador é devido aos quatro pares de tuplas: (t_0, t_2) , (t_2, t_0) , (t_1, t_2) e (t_2, t_1) que violam a restrição proposta e o denominador é a quantidade de pares de tuplas da instância: $7 * 6 = 42$.

3.2.3 Exemplo de Dependência Funcional como Restrição de Negação

A tradução de uma dependência funcional para uma RN é simples. Considere a dependência funcional $id \rightarrow nome, sobrenome$, ela expressa que id determina $nome$ e $sobrenome$, ou seja, um valor do atributo id tem um par de valores único para os atributos $nome$ e $sobrenome$ associado a ele.

Baseado nisso, construímos uma RN que é violada quando duas tuplas têm valores iguais para o atributo id e valores diferentes para os atributos $nome$ e $sobrenome$:

$$\varphi_c : \forall t_x, t_y \in r, \neg(t_x.id = t_y.id \wedge t_x.nome \neq t_y.nome \wedge t_x.sobrenome \neq t_y.sobrenome)$$

3.2.4 Exemplo de Dependência de Ordem como Restrição de Negação

Dependências de ordem são capazes de expressar relações entre atributos ordenados, algo que dependências funcionais não são capazes. Definição: seja X e Y listas de atributos de um esquema S , uma dependência de ordem $X \mapsto Y$, lida como X ordena Y , é satisfeita para uma instância s de S se, ao ordenar as tuplas com base nos atributos de X também obtemos um ordenamento dos atributos de Y . (Szlichta et al., 2018)

No exemplo dessa seção, temos a seguinte dependência de ordem: $(valor_hora, horas) \mapsto (salario)$. Ela expressa que ao ordenarmos as tuplas da instância r pelos atributos $valor_hora$ e $horas$, também iremos obter um ordenamento pelo atributo $salario$.

Essa dependência pode ser traduzida para uma restrição de negação:

$$\varphi_d : \forall t_x, t_y \in r, \neg(t_x.valor_hora < t_y.valor_hora \wedge t_x.horas < t_y.horas \wedge t_x.salario \geq t_y.salario),$$

ou seja, caso uma tupla t_x tenha valores menores que uma tupla t_y nos atributos *valor_hora* e *horas*, e um valor maior ou igual para no atributo *salario*, a restrição φ_d estará sendo violada.

Em dependências de ordem, o tipo de ordem, ascendente ou decedente, é o mesmo para todos os atributos. Quando há uma combinação dos tipos, chamamos de dependências de ordem bidirecionais. (Szlichta et al., 2018)

3.2.5 Exemplo de Combinação Única de Colunas como Restrição de Negação

Uma combinação única de colunas é um conjunto de atributos cuja a projeção em uma instância de base relacional não possui entradas duplicadas. Formalmente: uma combinação de colunas $X \in R$ é única para r , se para $t_i, t_j \in r, i \neq j : t_i[X] \neq t_j[X]$.

Combinações únicas de colunas também podem ser denotadas como chaves candidatas e em otimização de *queries* são úteis para, por exemplo, evitar eliminações duplicadas com o uso da cláusula *DISTINCT* e otimizar *joins* (Kossmann e Naumann, 2022).

Um exemplo de combinação única de colunas presente na Tabela 1.1 é conjunto com os atributos *nome* e *sobrenome*, uma vez que não há nenhum par de tuplas que possuem os mesmos valores para ambos os atributos. Essa combinação pode ser expressa como uma restrição de negação:

$$\varphi_e : \forall t_x, t_y \in r, \neg(t_x.nome = t_y.nome \wedge t_x.sobrenome = t_y.sobrenome)$$

3.3 AVALIAÇÃO DE RESTRIÇÕES DE NEGAÇÃO

Para a avaliação de RNs utilizaremos métricas que, de certo modo, são intuitivas e nos proporcionam embasamento para realizarmos comparações.

3.3.1 Sucintez

A sucintez é uma métrica que vai de acordo com o princípio da navalha de Occam e sua formalização, o princípio da descrição de comprimento mínimo. Esse princípio sugere que, dado várias hipóteses, aquela com menos variáveis e presunções, ou seja, um comprimento menor, é a melhor (Chu et al., 2013). Assim, dado um conjunto de RNs N , a sucintez de uma RN $\varphi \in N$, $Sucin(\varphi)$, é o menor valor de comprimento entre as RNs do conjunto N dividido pelo comprimento de φ , $Comp(\varphi)$, conforme definida pela Equação 3.3:

$$Sucin(\varphi) = \frac{\min \{Comp(\varphi_i) | \forall \varphi_i \in N\}}{Comp(\varphi)} \quad (3.3)$$

A equação para o cálculo de comprimento de uma RN pode variar. Nesse trabalho, utilizaremos a sugerida por Chu et al. (2013) e definida pela Equação 3.4: o comprimento de uma RN φ é a quantidade símbolos distintos que ela possui, esses símbolos são os que pertencem ao conjunto $S = \{t_0, t_1, C, O\}$, onde C é o conjunto de atributos do esquema R e O é o conjunto de operadores de comparação.

$$Comp(\varphi) = |\{s \mid s \in S \wedge s \in \varphi\}| \quad (3.4)$$

Como exemplo, considere as RNs da Seção 3.2. Seus comprimentos são: $Comp(\varphi_a) = 8$, $Comp(\varphi_b) = 7$, $Comp(\varphi_c) = 7$, $Comp(\varphi_d) = 7$ e $Comp(\varphi_e) = 5$, logo, seus valores de sucintez são: $Sucin(\varphi_a) = 0,625$, $Sucin(\varphi_b) \approx 0,71$, $Sucin(\varphi_c) \approx 0,71$, $Sucin(\varphi_d) \approx 0,71$ e $Sucin(\varphi_e) = 1,00$.

Assim, RNs com valores de sucintez maiores são mais sucintas, logo, pelo princípio da navalha de Occam, tendem a ser melhores, pois são menos ajustados aos dados, ou seja, não possuem sobre-ajuste (*overfitting*). Porém, sempre é necessário analisar a semântica de uma restrição para entender o seu real valor, porque é possível identificar uma RN trivial com comprimento pequeno, assim tendo uma sucintez alta, mas que não contém nenhuma informação valiosa ou útil. Por exemplo:

$$\varphi_f : \forall t_x, t_y \in r, \neg(t_x.\text{nome} = t_y.\text{cargo}),$$

uma restrição que diz que o nome de um funcionário não pode ser igual ao cargo dele.

3.3.2 Cobertura

A cobertura é uma métrica criada por Chu et al. (2013) que busca medir a significância estatística de uma RN, mediante de pesos dados as suas evidências. Uma evidência de uma RN φ é um par de tuplas que lhe satisfaz, ou seja, pelo menos um dos seus predicados não é falso. Com isso, estabelecemos como k – *evidencia* de uma RN φ , um par de tuplas, onde o k é o número de predicados de φ que são satisfeitos pelo par e $k \leq |\varphi.Predc| - 1$. (Chu et al., 2013). O conjunto de k – *evidencias* com um valor k é denotado como kE . E o peso de uma k – *evidencia* em uma restrição de negação φ é dado pela Equação 3.5:

$$w(k) = \frac{(k + 1)}{|\varphi.Predc|} \quad (3.5)$$

A partir disso, definimos a cobertura de uma RN φ para uma instância r como sendo a média dos pesos das suas k – *evidencias*, conforme definida pela Equação 3.6:

$$Cobertura(\varphi) = \frac{\sum_{k=1}^{\varphi.Predc-1} |kE| * w(k)}{\sum_{k=1}^{\varphi.Predc-1} |kE|} \quad (3.6)$$

Para as RNs da seção 3.2, temos as seguintes coberturas: $Cobertura(\varphi_a) \approx 0.738$, $Cobertura(\varphi_b) \approx 0.824$, $Cobertura(\varphi_c) \approx 0.968$, $Cobertura(\varphi_d) \approx 0.730$ e $Cobertura(\varphi_e) \approx 0.547$, logo, a φ_c possui uma significância estatística maior. Porém, assim como na métrica de sucintez, é necessário observar com atenção o valor semântico das RNs. Restrições triviais com um predicado, como a RN φ_f , podem obter uma cobertura de valor 1,0, porém, agregam pouco valor e são de quase nenhuma utilidade.

3.4 DESCOBERTA DE RESTRIÇÕES DE NEGAÇÃO

A descoberta automática de restrições de negação é um tema de bastante interesse, uma vez que elas possuem grande utilidade e a criação manual de RNs pode ser difícil dependendo da complexidade dos dados e dos esquemas (Bleifuß et al., 2017).

Porém, a descoberta executa duas tarefas que possuem uma alta complexidade computacional:

1. Explorar o conjunto de RNs possíveis: como o tamanho do conjunto de predicados possíveis, P , para um esquema R é: $|P| = |O| * m * (m - 1)$, onde m é a quantidade de atributos em R , e o espaço de procura de RNs, S , compreende todos os subconjuntos de predicados de P : $|S| = 2^{|P|}$, explorar todo esse conjunto é inviável. Além disso, já se foi provado por Bläsius et al. (2017) que detectar dependências funcionais é um problema $W[2]$ -complete, isso comprova complexidade computacional para se descobrir RNs, uma vez que elas generalizam dependências funcionais. Por isso, algoritmos de

descoberta de RNs buscam utilizar técnicas que diminuam esse custo e performem bem em dados do mundo real;

2. Avaliar se uma RN φ é válida: sendo n a quantidade de tuplas de uma instância r , todos os $n * (n - 1)$ pares de tuplas precisam ser testados nos predicados de φ . Portanto, essa tarefa tem um custo computacional $O(|\varphi.Predc| * n^2)$.

3.4.1 DCFinder

O algoritmo DCFinder destaca-se no estado da arte para o descobrimento de restrições de negação mínimas exatas e aproximadas. A Figura 3.1 mostra o esquema de sua execução. Resumidamente, o algoritmo passa pelas seguintes etapas:

1. Definição de um espaço de predicados;
2. Criação de listas de posições de índices (PLIs) que contém referências para tuplas que possuem atributos iguais;
3. Preparação de evidências de acordo com a seletividade dos predicados;
4. Completação das evidências com base nas PLIs;
5. Opcional: ranqueamento das RNs encontradas para descobrir quais são mais interessantes (*interestingness*).

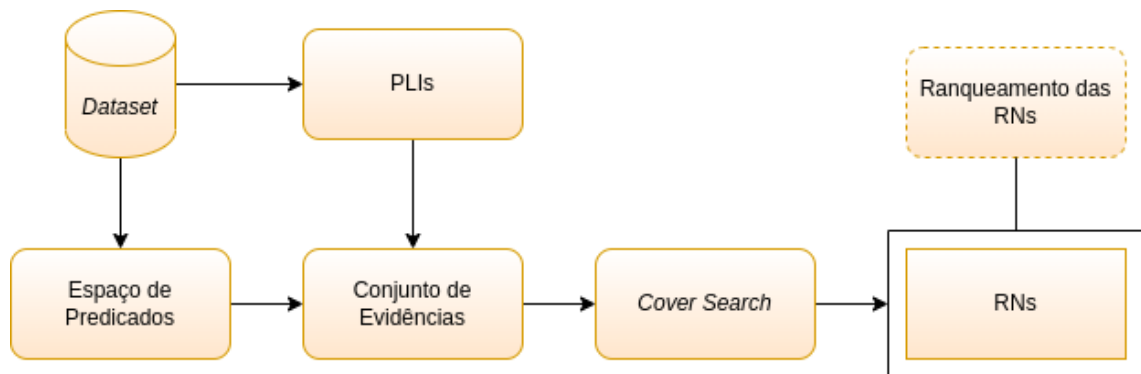


Figura 3.1: Esquema de execução do DCFinder
Fonte: adaptado de Pena et al. (2019)

3.4.2 Metanome

Nesse trabalho executaremos o código do DCFinder desenvolvido para ser utilizado via Metanome, uma plataforma criada por Papenbrock et al. (2015) com a intenção de integrar, padronizar e agrupar algoritmos relacionados a *data profiling*, fornecendo interfaces padrões para o desenvolvimento desses algoritmos, suporte a vários tipos de fontes de dados e funcionalidades como o tratamento de parametrização e a gestão de dados temporários. Sua arquitetura é demonstrada na Figura 3.2.

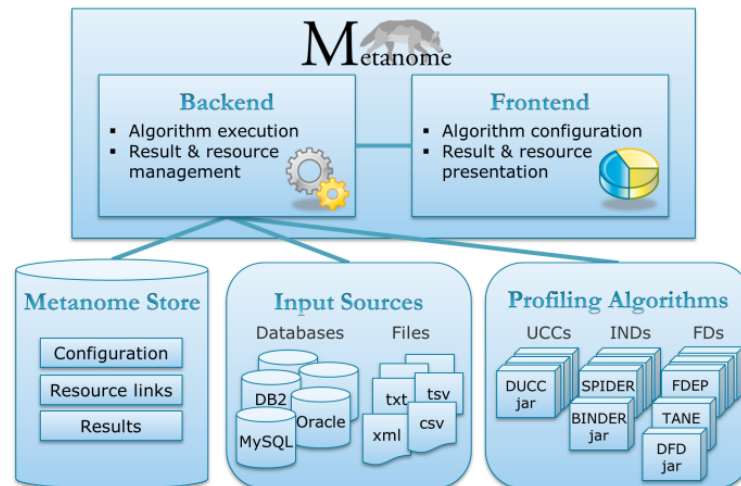


Figura 3.2: Arquitetura da plataforma Metanome
 Fonte: Papenbrock et al. (2015)

3.5 CONSIDERAÇÕES FINAIS

Nesse capítulo definimos o que é uma restrição de negação, demonstramos alguns exemplos e apresentamos métricas que irão nos ajudar na avaliação de RNs. Também apresentamos o algoritmo que iremos utilizar para a descoberta de restrições em nossos dados. No próximo capítulo falaremos sobre séries temporais e duas formas de percorrê-las.

4 SÉRIES TEMPORAIS

Séries temporais são comuns em diversas áreas do conhecimento, como: finanças, monitoramento ambiental, engenharia de aviação, astrofísica e outras. A análise dessas séries recebe muita atenção devido ao conhecimento que pode ser extraído delas (Schneider et al., 2021).

A descoberta de padrões e de eventos especiais são alguns dos objetivos dessas análises. Dentre os eventos especiais, as anomalias, subsequências com valores e padrões não usuais, possuem grande relevância, uma vez que costumam indicar a ocorrência de algum evento importante em áreas como, por exemplo, detecção de falta em oscilografias de corrente e tensão.

Neste trabalho, utilizaremos séries temporais contendo a corrente e tensão de um barramento em uma linha de transmissão.

4.1 DEFINIÇÃO

Uma série temporal T é definida como sendo uma sequência de amostras (registros) $(T_0, T_1, T_2, \dots, T_{n-1})$ ordenada cronologicamente, onde T_i é o valor da amostra na posição i em T com $0 \leq i < n$ e o comprimento de T é $|T| = n$.

Com isso, também definimos o que é uma subsequência de uma série temporal T : uma sequência de registros consecutivos de T começando em j e indo até $k - 1$, $T_{[j,k)} = (T_j, T_{j+1}, T_{j+2}, \dots, T_{k-1})$, com $0 \leq j < k \leq |T|$ (Schneider et al., 2021).

Em seguida, veremos duas maneiras de percorrer séries temporais: janela deslizando (*sliding window*) e janela em cascata (*tumbling window*), uma variante da janela deslizando.

4.2 JANELA DESLIZANTE

O processo de janela deslizando é uma maneira de percorrer séries temporais por meio de subsequências. Assim, uma janela deslizando W de tamanho l desliza do início ao fim de uma série avançando s amostras de cada vez. Por exemplo, uma janela W deslizando sobre uma série temporal T de tamanho n e avançando s amostras por vez, começa com uma janela W_1 contendo a subsequência $T_{[0,l)}$, em seguida avança para a subsequência $T_{[l,l+s)}$ na janela W_2 e assim por diante até $T_{[n-w,n)}$. A Figura 4.1 ilustra o processo (Luping e Huang, 2022).

4.2.1 Janela em Cascata

O processo de janela em cascata é uma variação da janela deslizando. Nela a quantidade de amostras s em que se avança de janela em janela é igual ao tamanho da janela, $s = l$. Assim, não há sobreposição de amostras entre as janelas (Olmezogullari e Ari, 2013).

4.3 CONSIDERAÇÕES FINAIS

Nessa seção apresentamos o conceito de séries temporais e dois processos para percorrê-las e efetuar análises. Esses processos serão utilizados cada um em uma etapa do trabalho desenvolvido. No capítulo a seguir, apresentamos os dados utilizados neste trabalho para a detecção de falta.

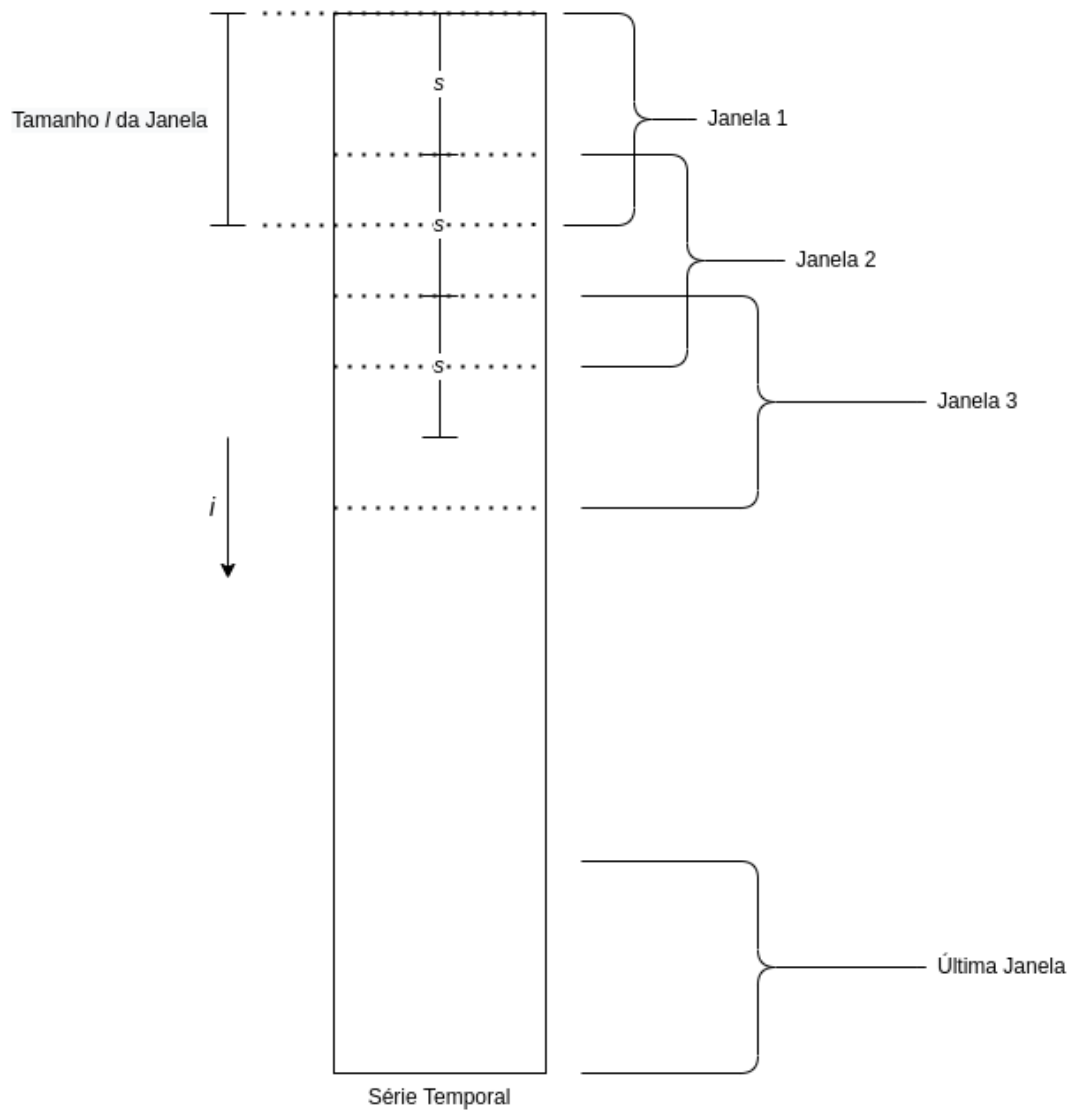


Figura 4.1: Diagrama esquemático do processo de janela deslizante
 Fonte: adaptado de Luping e Huang (2022)

5 BASE DE DADOS COM SIMULAÇÕES DE FALTA

Os dados utilizados são provenientes da *Fault Analysis Database* (FADb), criado por Ensina et al. (2022). Essa base de dados contém simulações de falta que foram geradas por meio das ferramentas ATP (European EMTP-ATP Users Group, 2020) e ATPDraw (Høidalen, 2020) utilizando o modelo IEEE 9-barras presente no segundo. Esse modelo é composto por nove barramentos, três geradores e seis linhas de transmissão, sendo uma aproximação de um sistema elétrico de potência real. Seu esquema pode ser visto na Figura 5.1.

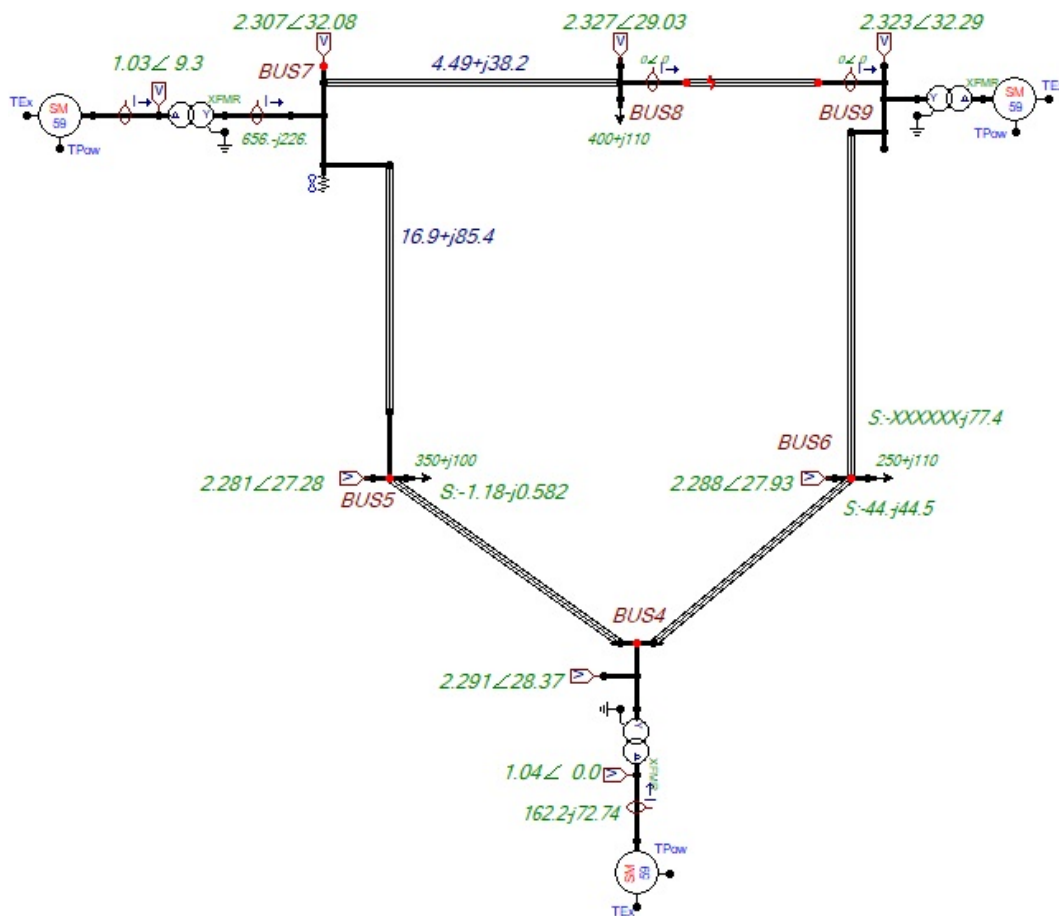


Figura 5.1: IEEE 9-barras modelado no ATPDraw

Nesse trabalho foram utilizadas apenas as simulações da linha de transmissão entre os barramentos 8 e 9 com 414km de extensão, 500kV de tensão e 60Hz de frequência. Essas simulações possuem as amostras de corrente e tensão para as três fases (A, B e C) dos dois barramentos da linha, com duração de 0,5 segundo para uma taxa de amostragem de 10kHz, o que resulta em 5000 amostras por simulação. Cada uma delas varia os seguintes parâmetros:

- Tipo da falta: AB, ABC, ABG, AC, ACG, AG, BC, BCG, BG e CG. As letras de cada tipo de falta indicam quais fases estão envolvidas na falha, bem como se há o contato com a terra (G);
- Tempo em que a falta começa: 0,091s, 0,093s, 0,095s, 0,097s, 0,099s, 0,101s, 0,103s e 0,105s;

- Resistência: de 0,01 a 200 ohms com intervalos de 10 ohms;
- Localização: de 1% a 100% com intervalos de 1%.

Assim, temos um total de 168000 simulações.

5.1 EXEMPLO

Nessa seção, um exemplo com cinco amostras de uma simulação pode ser visto na Tabela 5.1 para tornar o leitor mais familiar com simulações presentes na base dados. As colunas da Tabela 5.1 têm os seguintes significados:

- IA: corrente da fase A no barramento 8;
- IB: corrente da fase B no barramento 8;
- IC: corrente da fase C no barramento 8;
- VA: tensão da fase A no barramento 8;
- VB: tensão da fase B no barramento 8;
- VC: tensão da fase C no barramento 8.

Uma restrição de negação para esse exemplo é:

$$\varphi : \neg(t0.VC \leq t1.VC \wedge t0.IB \leq t1.IB \wedge t0.VB \leq t1.VB) \quad (5.1)$$

Tabela 5.1: Exemplo reduzido de uma simulação

IA	IB	IC	VA	VB	VC
-254,98827	316,58102	-61,59275	382069,7	-7397,7866	-374671,9
-263,0354	312,1483	-49,112877	373807,2	9072,551	-382879,75
-270,70898	307,27222	-36,563248	365013,6	25530,016	-390543,6
-277,99826	301,96	-23,961725	355701,38	41951,242	-397652,62
-284,8931	296,2193	-11,326216	345883,78	58312,914	-404196,7

5.2 CONSIDERAÇÕES FINAIS

Nesse capítulo foram descritos os dados que serão utilizados, sua origem e como foram gerados. Para maior familiaridade, um exemplo das amostras e de uma restrição de negação foi apresentado. A seguir, iremos descrever a metodologia empregada no desenvolvimento do trabalho.

6 MÉTODO

O desenvolvimento do trabalho passa por duas etapas principais: a descoberta de restrições de negação exatas que possam ser utilizadas para a detecção de faltas e a avaliação do desempenho dessas restrições descobertas. Neste capítulo vamos descrever o método proposto para detecção de faltas em linhas de transmissão e uma breve descrição do hardware utilizado.

6.1 DESCOBERTA DE RESTRIÇÕES DE NEGAÇÃO

O processo de descoberta de restrições de negação possui três fases: seleção dos dados (Subseção 6.1.1), loteamento dos dados selecionados (Subseção 6.1.2) e particionamento dos atributos (Subseção 6.1.3). Cada uma das três é justificada pois contribuem para que seja possível alcançar o objetivo principal: descobrir restrições de negação capazes de detectar faltas. Além disso, a execução dessas fases permitem calcular as métricas de cobertura e sucintez das restrições encontradas (para análise posterior), em um tempo hábil.

6.1.1 Seleção dos Dados

Os experimentos contemplaram o uso parcial das simulações da base de dados, como as que seguem:

- Tipo da falta: AB, ABC, ABG, AC, ACG, AG, BC, BCG, BG e CG;
- Tempo em que a falta começa: 0,095s, 0,097s, 0,103s e 0,105s;
- Resistência: 0,01 ohm, 50 ohms, 100 ohms, 150 ohms e 200 ohms;
- Localização: 1%, 25%, 50%, 75% e 100%.

Portanto, foram selecionadas 1000 simulações que cobrem todos os tipos de falta e diversas faixas de valores para os outros parâmetros. O uso parcial da base reduz o tempo necessário para realizar o processo e a variação dos parâmetros evita que sejam descobertas restrições específicas para apenas alguns deles.

6.1.2 Loteamento dos Dados Selecionados

As simulações foram divididas em subsequências (lotes) de 1000 amostras para melhorar o tempo de execução do cálculo da métrica de cobertura e ter a descoberta de RNs em subsequências que não possuíssem falta. Essa divisão é o mesmo que a utilização de janelas em cascata com tamanho $l = 1000$ e uma representação desse loteamento é visto na Figura 6.1. Assim, com o loteamento das simulações selecionadas, teremos, no total, 500 lotes com amostras sem falta e 4500 lotes com amostras com falta.

6.1.3 Particionamento dos Atributos dos Dados Selecionados

Como dito anteriormente, as simulações possuem os valores de corrente e tensão para as três fases dos dois barramentos da linha transmissão, ou seja, temos 12 atributos. O particionamento realizado separa os valores de cada barramento, ficando assim com seis atributos.

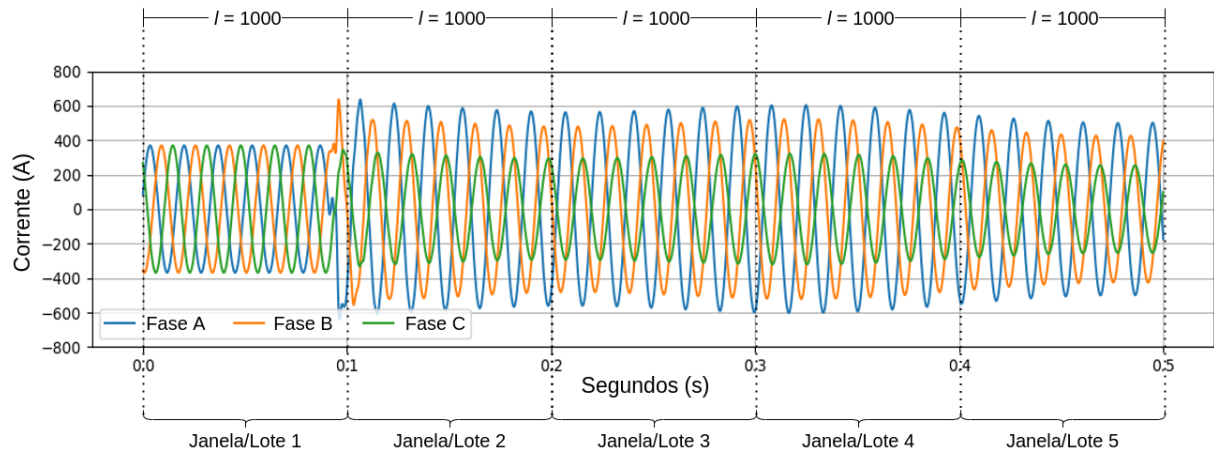


Figura 6.1: Loteamento das simulações para descoberta de restrições

Na Figura 6.2, uma ilustração do particionamento, onde, os atributos (ou colunas) selecionados estão em verde e os removidos estão em vermelho, e ao lado, uma indicação de que as amostras da série temporal estão ordenadas conforme sua posição i .

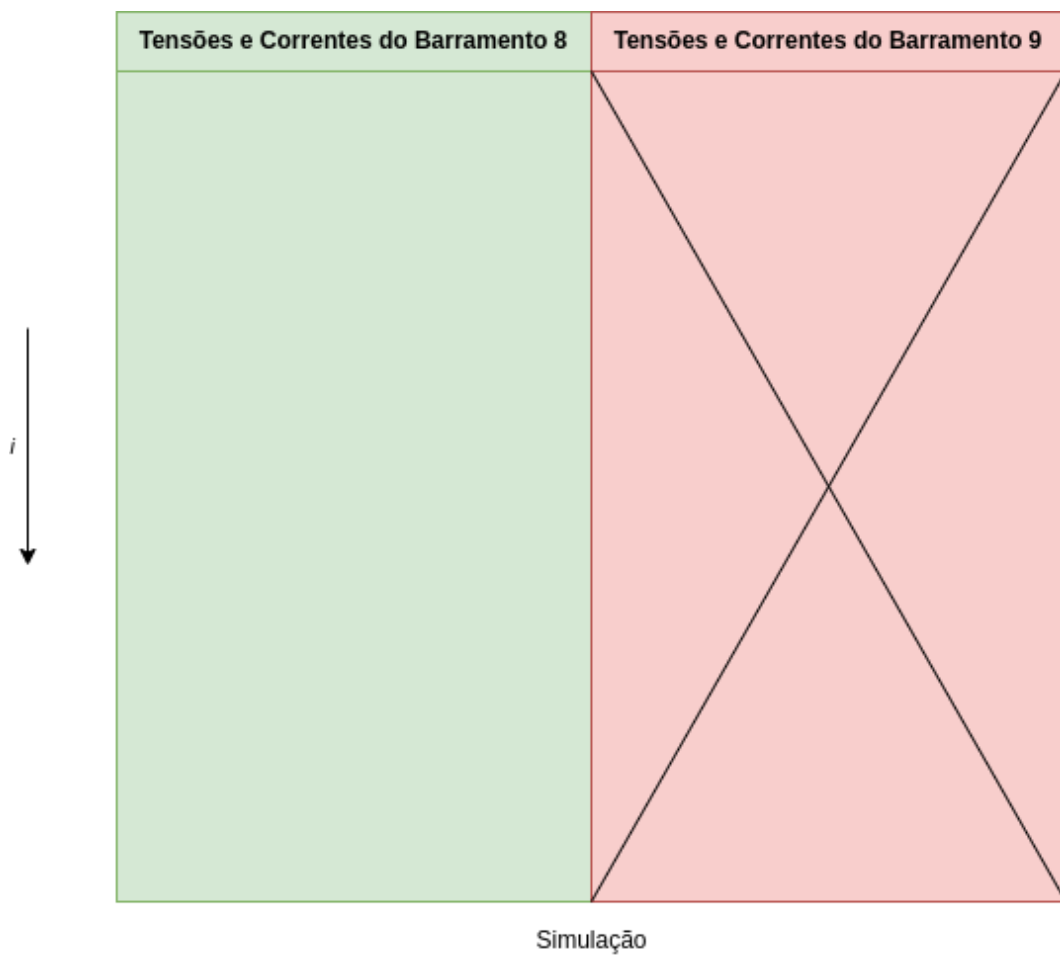


Figura 6.2: Particionamento do espaço de predicados para a descoberta de restrições

6.1.4 Descoberta de Restrições de Negação com o DCFinder

A descoberta de restrições de negação exatas ocorrerá com a execução do DCFinder para cada lote particionado de cada simulação. A Figura 6.3 representa os lotes particionados de uma simulação sobre os quais serão executados o algoritmo.

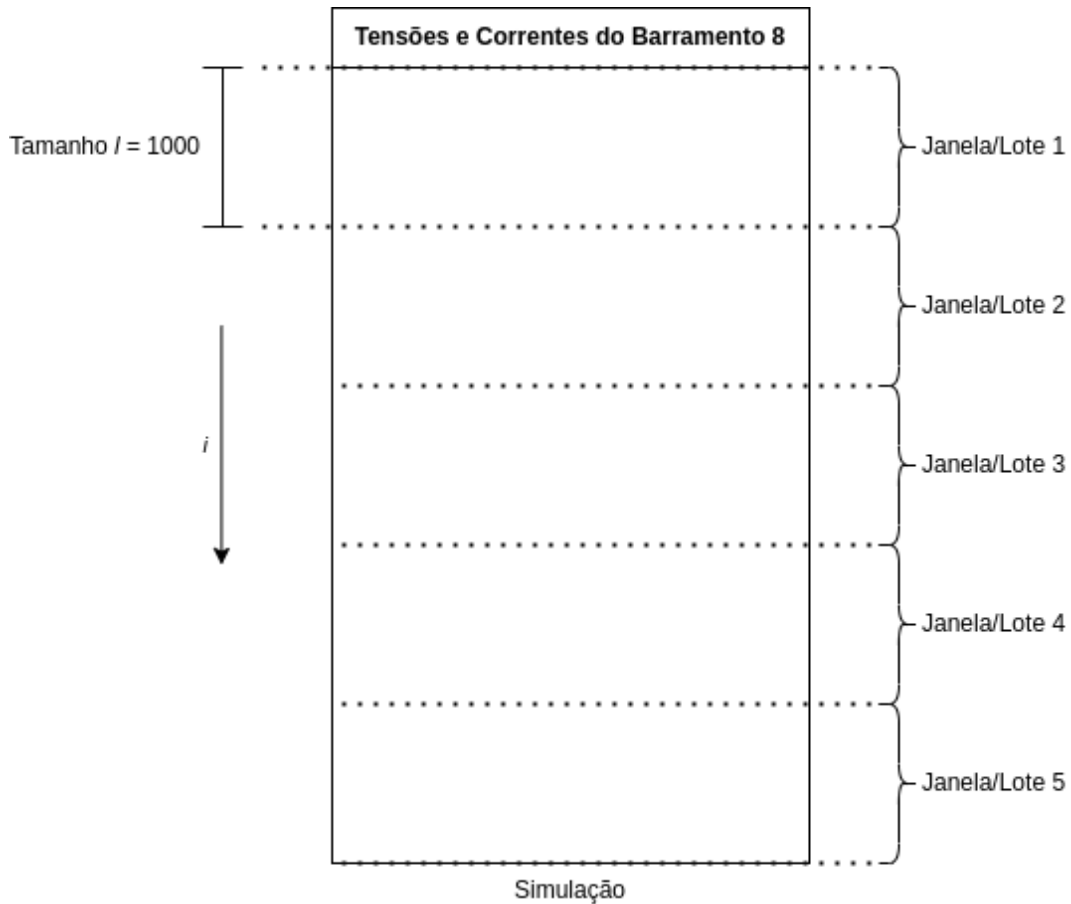


Figura 6.3: Representação da execução do DCFinder sobre uma simulação loteada e particionada

6.2 AVALIANDO O DESEMPENHO DAS RESTRIÇÕES DE NEGAÇÃO DESCOBERTAS

Para avaliar o desempenho das restrições de negação descobertas foram selecionadas outras simulações, conforme identificadas na Subseção 6.2.1, e foi desenvolvido um processo de janela deslizante que obtém grau de aproximação das janelas para uma determinada RN descoberta em lotes sem falta e, com base em um limite dado (serão experimentados alguns valores para esse limite) para o grau de aproximação, classifica as janelas quanto a presença ou não de amostras com falta. Para os erros e acertos das restrições em cada janela para o limite dado, calcula-se o F1-score, a precisão, o *recall* e a acurácia obtida de cada RN. O detalhamento do processo é visto a seguir.

6.2.1 Seleção dos Dados

O desempenho das restrições foi avaliado nas simulações com as seguintes variações de parâmetros:

- Tipo da falta: AB, ABC, ABG, AC, ACG, AG, BC, BCG, BG e CG;

- Tempo em que a falta começa: 0,091s, 0,093s, 0,099 e 0,101s;
- Resistência: 0,01 ohm, 50 ohms, 100 ohms, 150 ohms e 200 ohms;
- Localização: 1%, 25%, 50%, 75% e 100%.

Portanto, foram selecionadas 1000 simulações, todas diferentes das utilizadas na descoberta das restrições, que cobrem todos os tipos de falta e diversos valores para os outros parâmetros, o que nos permitirá realizar todo o processo em um tempo hábil e analisar a influência da variação de cada um deles na nossa avaliação.

6.2.2 Janela Deslizante e Classificação

O processo da janela deslizante com avanço $s = 1$ implementado consiste em percorrer uma simulação e, em cada etapa da janela, calcular o grau de aproximação das amostras presentes nela para uma determinada restrição.

Assim, com o grau de aproximação de todas as janelas, escolhe-se um limite para classificá-las. Caso o grau de aproximação de uma janela seja maior que esse limite, essa janela é classificada como contendo amostra com falta, ou seja, uma falta foi detectada, caso contrário, a janela é classificada como não tendo amostra com falta.

Como sabemos em que instante de tempo uma falta começa em uma simulação, também sabemos a partir de qual amostra o período faltoso começa. Assim, somos capazes de avaliar se a classificação para uma dada janela foi correta ou não. Nesse trabalho, consideramos como período faltoso, todo o período após o início da falta, e o tempo anterior ao começo da falta como sendo livre de faltas.

O resultado da classificação de uma janela é categorizado da seguinte maneira:

- Verdadeiro positivo (VP): a janela foi classificada como tendo falta e ela realmente tinha;
- Verdadeiro negativo (VN): a janela foi classificada como não tendo falta e ela realmente não tinha;
- Falso positivo (FP): a janela foi classificada como tendo falta, mas ela não tinha;
- Falso negativo (FN): a janela foi classificada como não tendo falta, mas ela tinha.

Tendo categorizado todos os resultados de todas as janelas em todas as simulações selecionadas para uma determinada RN e um limite de grau de aproximação, é realizado o cálculo do F1-score, precisão, *recall* e acurácia.

A precisão é obtida através da fórmula:

$$Precisao = \frac{VP}{VP + FP} \quad (6.1)$$

Já o *recall* é dado pela fórmula:

$$Recall = \frac{VP}{VP + FN} \quad (6.2)$$

O F1-score combina ambas as métricas anteriores:

$$F1 - score = 2 * \frac{Precisao * Recall}{Precisao + Recall} \quad (6.3)$$

Por fim, a acurácia:

$$Acuracia = \frac{vp + vn}{vp + vn + fp + fn} \quad (6.4)$$

A Figura 6.4 representa esse processo. Nela, considerando uma simulação de 10 amostras (simbolizadas pelos retângulos), sendo cinco delas com falta (pintadas de amarelo quando fora da janela deslizante e roxo quando dentro), é avaliada utilizando uma janela de tamanho $w = 4$ (pintada de azul para as amostras sem falta e de roxo para amostras com falta) e que avança de uma em uma amostra ($s = 1$). Os graus de aproximação de cada etapa da janela são mostrados abaixo e, em seguida, a classificação de cada janela utilizando como parâmetro um limite de grau de aproximação de 0,2 (valor estabelecido apenas para exemplificar). Assim, a avaliação de uma RN qualquer, com esses parâmetros da janela e para essa simulação, obteve um falso positivo, um verdadeiro negativo, dois falsos negativos e três verdadeiros positivos.

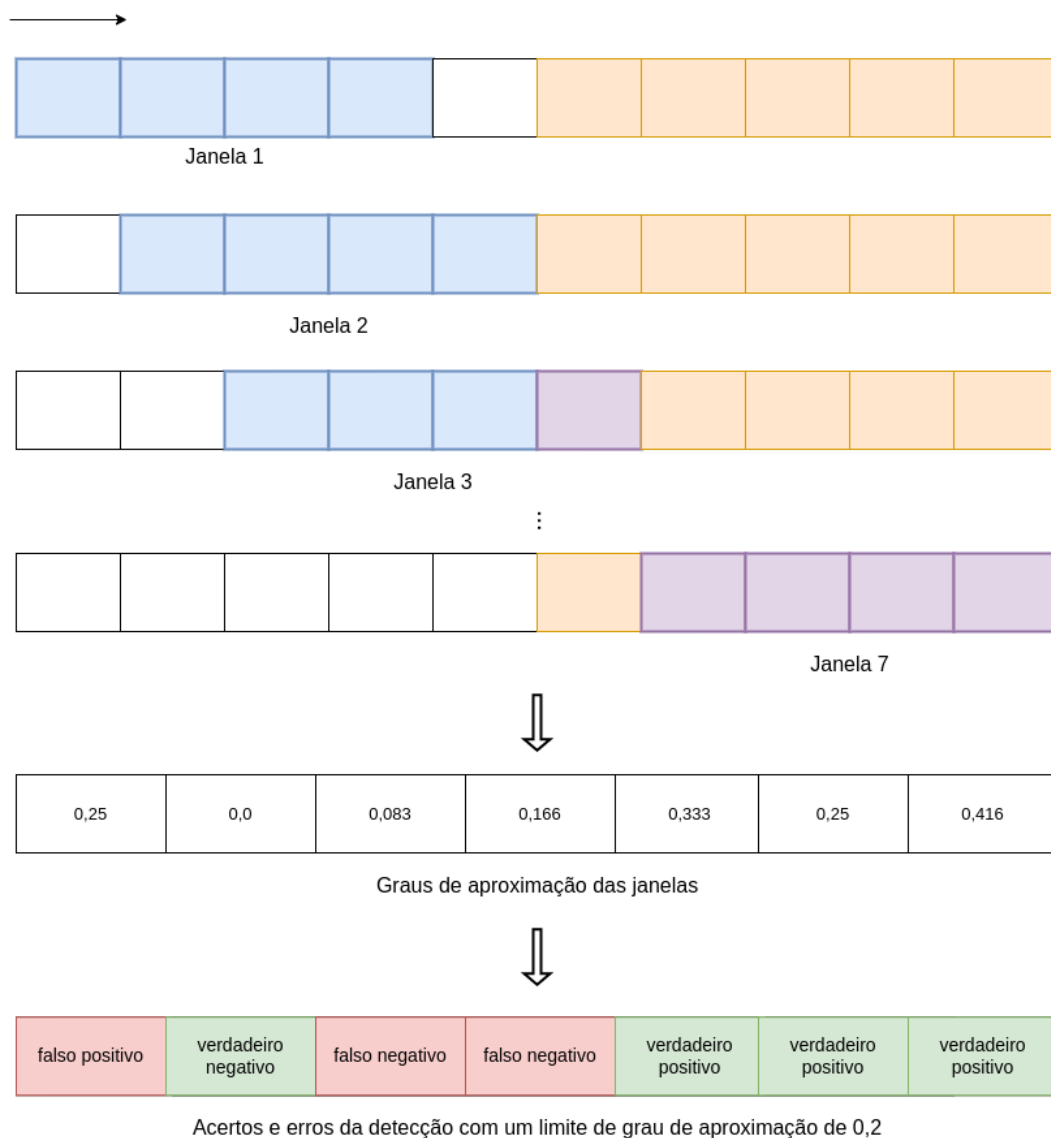


Figura 6.4: Representação do processo de avaliação com janela deslizante

Fonte: o autor

Outra métrica que desejamos obter é quanto tempo foi necessário para detectar a falta de uma simulação a partir do seu real início. Para isso, contaremos em qual amostra a partir do

início da falta uma janela detectou a anomalia pela primeira vez. Por exemplo, na Figura 6.4, a janela teve o seu primeiro verdadeiro positivo na segunda amostra com falta. Também desejamos contar em quantas simulações a janela não detectou falta.

6.2.3 Parâmetros Variados

Com o objetivo de estabelecer qual o melhor tamanho de janela e qual o melhor limite para classificar uma janela como faltosa, consideramos os seguintes valores:

- Tamanho da janela: 50, 100, 200, 400 e 800;
- Limite do grau de aproximação: 0, $1 * 10^{-5}$, $1 * 10^{-4}$, $5 * 10^{-4}$, $1 * 10^{-3}$, $5 * 10^{-3}$, $1 * 10^{-2}$, $2 * 10^{-2}$, $4 * 10^{-2}$, $6 * 10^{-2}$.

O critério adotado para a escolha do melhor par de parâmetros foi a média do F1-score das 10 RNs com os maiores F1-score obtidos em cada um. A escolha pela métrica F1-score é devido a sua combinação de precisão e *recall*, permitindo que levemos em conta tanto os erros relacionados a falsos positivos quanto aos erros de falsos negativos, além de balancear os dados, uma vez que as simulações possuem a maior parte das suas amostras em um período de falta. Já o uso da média do F1-score das 10 melhores RNs ajuda a encontrar qual é o melhor par de parâmetros de maneira geral, pois estaremos avaliando mais de uma RN o que aumenta a generalidade do efeitos da variação dos parâmetros.

6.3 HARDWARE

A execução do método proposto será realizado em uma máquina com quatro processadores AMD Opteron 8387 que possuem quatro *cores* e quatro *threads* cada e 64GB de Memória RAM.

6.4 CONSIDERAÇÕES FINAIS

Tendo descrito todo o método utilizado no desenvolvimento do trabalho, o próximo capítulo apresenta os resultados obtidos e uma análise dos mesmos.

7 RESULTADOS E DISCUSSÃO

Nesse capítulo serão apresentados os resultados obtidos em cada uma das duas etapas descritas anteriormente: descoberta de restrições de negação e avaliação do desempenho dessas RNs para detecção de faltas.

7.1 RESTRIÇÕES DE NEGAÇÃO DESCOBERTAS

Apresentamos a seguir a análise das restrições de negação descobertas e detalhes sobre cada fase do processo que justificam suas existências.

7.1.1 Processo de Descoberta de Restrições de Negação

Inicialmente, executamos o DCFinder sobre as simulações selecionadas sem loteamento e particionamento, foram encontradas 140879 RNs únicas e uma média de aproximadamente 1300 RNs por simulação. O tempo execução médio de execução do DCFinder para cada simulação foi 9,4388 segundos.

Esse número elevado de RNs inviabilizava a realização de outras etapas do processo, como a obtenção da métrica de cobertura e a posterior avaliação das restrições, devido ao tempo que seria necessário para realizar essas tarefas. Além disso, sem o loteamento dos dados, não era possível distinguir quais restrições foram encontradas em amostras sem falta, restrições essas que seriam utilizadas para a detecção de faltas.

Com o loteamento das simulações selecionadas em lotes de 1000 amostras foram descobertas 153052 RNs únicas e uma média de aproximadamente 555 RNs por lote. Um valor também alto e que implicava nas mesmas impossibilidades anteriores, mas agora era possível distinguir entre restrições presentes em amostras sem falta e com falta. O tempo médio de execução do DCFinder para cada lote foi 13,1421 segundos, um aumento em relação as simulações completas devido ao encontro de um número elevado de restrições de negação em lotes sem falta ou com poucas amostras de falta.

Por fim, aplicando o loteamento e o particionamento (uso dos atributos apenas do barramento 8) nos dados selecionados, a quantidade de RNs foi bastante reduzida, sendo encontradas 695 RNs únicas e tendo uma média de aproximadamente 18 RNs por lote, o que demonstra a relevância do número de atributos, pois ele determina o tamanho do espaço de predicados. Assim, foi possível realizar o cálculo da métrica de cobertura das RNs encontradas em aproximadamente 64 horas, uma média de 2,4577 segundos para cada cálculo, e prosseguir para a avaliação da capacidade de detecção de faltas das restrições encontradas em lotes sem falta. Ainda mais importante, o particionamento faz com que sejam descobertas restrições que possuem predicados com atributos de um único barramento, logo, a detecção de falta com essas RNs não possui a necessidade de sincronização dos dados coletados em ambos os barramentos, o que é mais condizente com um ambiente real de operação. O tempo médio de execução do DCFinder para cada lote foi 0,9534 segundo, o que evidencia a influência do número de atributos tanto na quantidade de RNs encontradas quanto no tempo necessário para realizar a descoberta das restrições.

7.1.2 Detalhamento das Restrições de Negação Descobertas em Lotes Sem Falta

Das 695 restrições encontradas, 100 foram encontradas em lotes com amostras sem falta e 694 foram encontradas em lotes com amostras com falta, ou seja, apenas uma RN foi encontrada exclusivamente em lotes sem falta, as outras 99 dos lotes sem falta também apareceram lotes com falta. Essa proporção nas quantidades de vezes em que uma RN foi encontrada em lotes sem e com falta variou bastante, por exemplo, obtivemos uma restrição que foi encontrada em todos os 500 lotes sem falta e que também foi encontrada em apenas 62 lotes com falta, já outra RN que foi encontrada em todos os lotes sem falta, também foi descoberta em 3386 lotes com falta.

A média da cobertura de cada uma das 100 RNs encontradas em lotes sem falta também variou. Restrições triviais com apenas um predicado obtiveram uma cobertura média de 1.0, outras também triviais com apenas dois predicados tiveram uma cobertura média de aproximadamente 0,5, já as não triviais variaram entre aproximadamente 0,666 e 0,833.

Os comprimentos dessas 100 restrições ficaram entre dois e seis, tendo a grande maioria (79%) cinco ou seis. Desconsiderando as restrições mais triviais (com um e dois predicados), o menor comprimento foi quatro, assim, a sucintez da maioria foi de $1 * 10^{-5}$, 666 e 0,8.

Analisando quais eram os tipos de dependência que as restrições encontradas representavam, notamos que obtivemos uma grande quantidade de dependências de ordem bidirecionais, como:

$$\varphi_e : \forall t_x, t_y \in r, \neg(t_x.IC \leq t_y.IC \wedge t_x.IA \leq t_y.IA \wedge t_x.IB \leq t_y.IB),$$

o que é coerente, pois as três fases possuem valores de corrente simétricos e equilibrados, então, o ordenamento pela corrente de uma das fases também deve ordenar os valores de outra fase, ainda que em um sentido (ascendente ou descendente) diferente. Também tivemos combinações únicas de colunas triviais:

$$\varphi_f : \forall t_x, t_y \in r, \neg(t_x.VA = t_y.VA \wedge t_x.IC = t_y.IC),$$

e outras não se adequavam a nenhuma definição conhecida, mas possuíam um padrão onde há um predicado com o operador de igualdade e outros dois com os operadores de desigualdade, como por exemplo:

$$\varphi_g : \forall t_x, t_y \in r, \neg(t_x.IC \leq t_y.IC \wedge t_x.VA \leq t_y.VA \wedge t_x.IA = t_y.IA)$$

Na Tabela A.1 do Apêndice A pode ser visto todas as restrições encontradas, assim como informações de cobertura, comprimento, quantidade lotes em que cada uma foi encontradas e outras.

7.2 RESULTADOS DAS AVALIAÇÕES

Nessa seção apresentaremos os resultados obtidos nas avaliações realizadas com o método proposto. Observaremos como a variação dos parâmetros da janela (tamanho e limite do grau de aproximação) repercutiram no desempenho e como o melhor par de parâmetros da janela se comportou com a variação dos parâmetros das simulações selecionadas. Por fim, analisaremos a relação entre o desempenho das restrições encontradas e as métricas obtidas no processo de descoberta das mesmas.

7.2.1 Variação dos Parâmetros da Janela

Como pode ser visto na Figura 7.1, o par de parâmetros com maior F1-score médio foi com o tamanho da janela de $w = 800$ e o limite do grau de aproximação de 0, que obteve um valor de 0,986228 nas suas 10 melhores RNs. Também podemos observar que o desempenho melhora quando o tamanho da janela aumenta e o limite do grau de aproximação diminui.

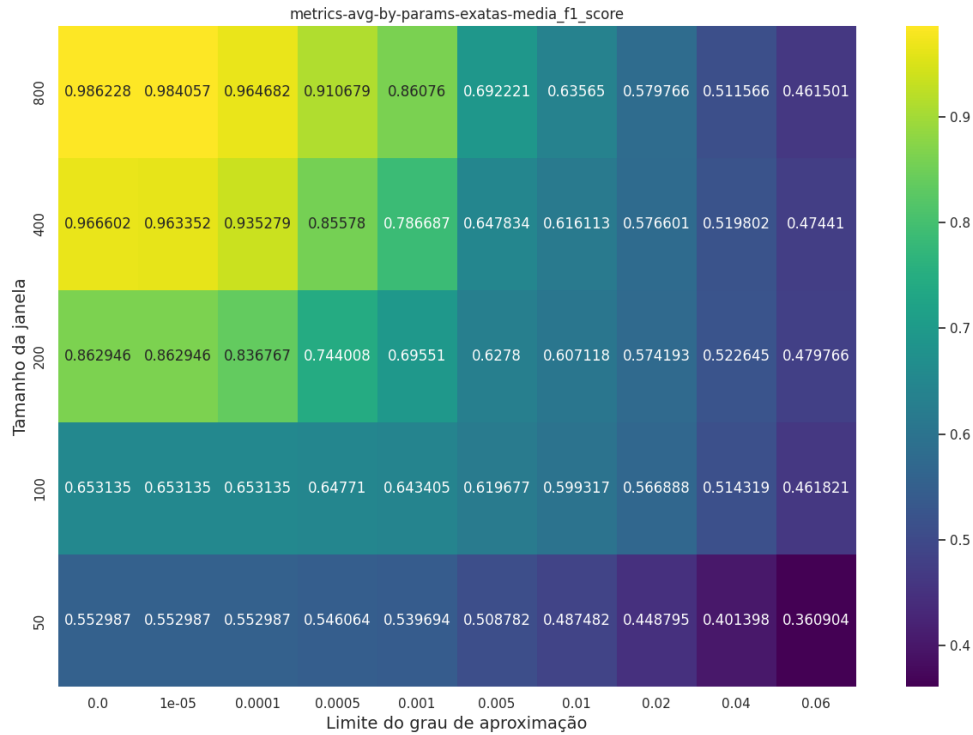


Figura 7.1: Média do F1-score das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$

Apesar desse par de parâmetros ter obtido o melhor desempenho com base no F1-score, os outros pares obtiveram uma precisão de 100% e ele foi o único que teve a ocorrência de falsos positivos, tendo uma média de precisão de 0,993708, como ser visto na Figura 7.2. Sua vantagem foi ter uma menor quantidade de falsos negativos, que pode ser visto nos valores da métrica de *recall* na Figura 7.3. Devido a essa presença de falsos positivos, que indicam a possibilidade de um *overfitting* dos parâmetros, avaliamos como sendo o melhor outro par de parâmetros: aquele com o tamanho da janela de $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$, pois ele obteve um F1-score médio aproximadamente igual (0,984057) ao com maior F1-score e uma precisão de 100%, ou seja, sem a ocorrência de falsos positivos.

Na Figura 7.4, podemos verificar que o par de parâmetros que avaliamos como sendo o melhor detectou as faltas com aproximadamente 20 amostras em média após o início de uma falha. Como a taxa de amostragem da linha de transmissão simulada é de 10kHz, o tempo médio para detectar uma falha, desconsiderando o tempo necessário para calcular o grau de aproximação da janela e classificá-la, foi de aproximadamente 20 milissegundos.

A Tabela 7.1 mostra o tempo médio necessário em cada tamanho de janela para calcular o grau de aproximação e classificar. Da maneira que o cálculo do grau de aproximação foi implementado, aproveitando as violações encontradas de uma janela anterior para calcular o grau de aproximação da janela seguinte, o desempenho se mostrou linear em relação ao tamanho da janela. Assim, as janelas com tamanho $w = 800$ levaram em média cerca de 331 milissegundos para ter seu grau de aproximação calculado e classificado.

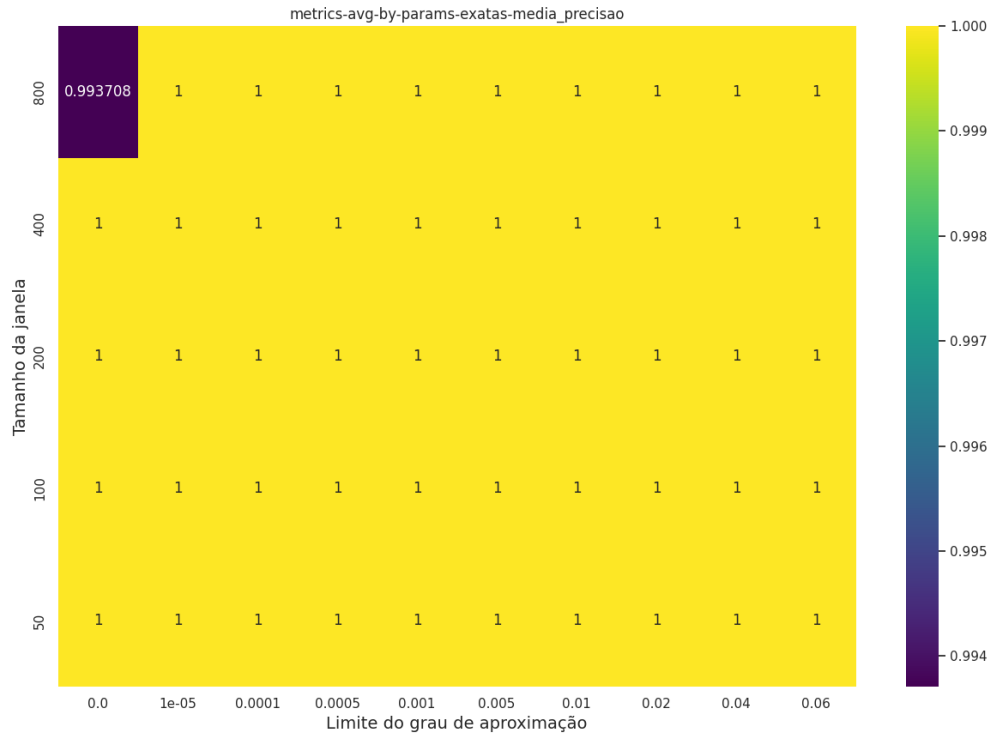


Figura 7.2: Média da precisão das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$

Tabela 7.1: Tempo médio para obter o grau de aproximação de cada janela e classificá-la

Tamanho da Janela	Tempo médio (ms)
50	20,7487
100	40,3926
200	78,3719
400	159,1655
800	331,5353

E na Figura 7.5, notamos que as janelas com tamanho maior ou igual a 200 com um limite no grau de aproximação menor ou igual a 0,0001, tiveram suas 10 melhores restrições detectando faltas em todas as simulações.

Tentando descobrir o porque de janelas maiores obterem menos falsos negativos, já que, intuitivamente, pensamos o contrário, pois janelas menores têm seus graus de aproximação mais sensíveis a mudança na quantidade de violações, buscamos realizar algumas poucas observações sobre os ocorrimto da não detecção da falta. A hipótese mais provável é que janelas maiores contemplam mais amostras e por isso possuem mais pares de tuplas que configuram uma violação, já quando a janela é menor, pares de tuplas que violariam a restrição não se encontram na mesma janela. E isso, combinado com um limite de grau de aproximação menor, reduz a quantidade de falsos negativos.

O desempenho individual de cada restrição de negação descoberta pode ser visto na Tabela B.1 do Apêndice B.

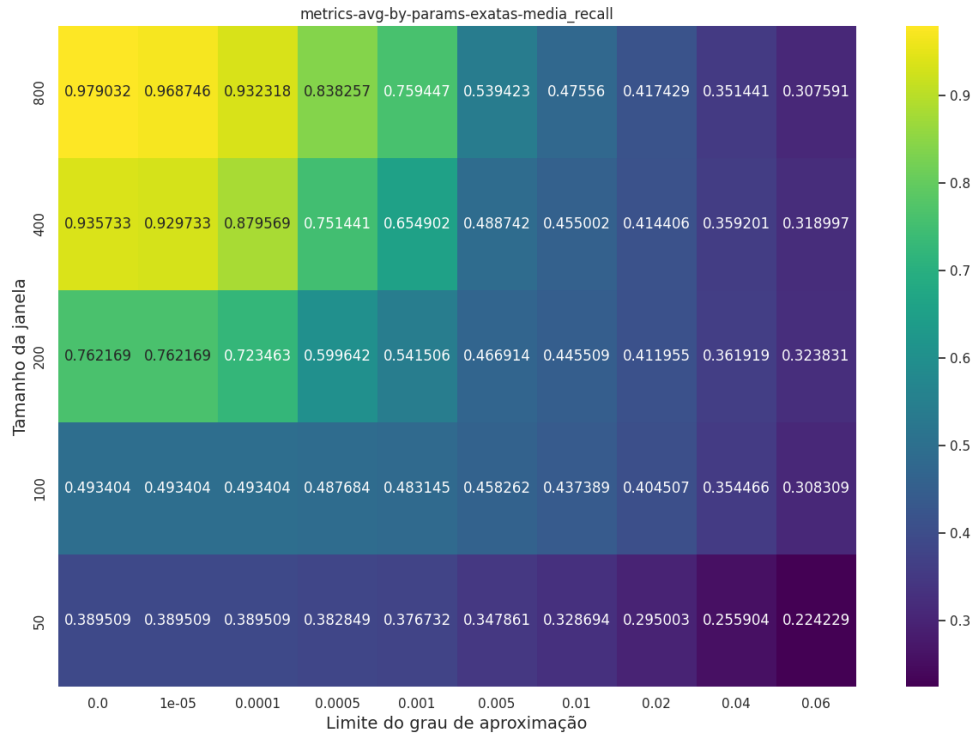


Figura 7.3: Média do *recall* das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$

7.2.2 Variação dos Parâmetros das Simulações

Nessa seção, analisaremos de modo minucioso a influência da variação dos parâmetros das simulações selecionadas, descritas na Subseção 6.2.1, no desempenho médio das 10 melhores RNs com o par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$, que, como visto anteriormente, avaliamos como sendo o melhor par.

Na Tabela 7.2, podemos observar que o tempo de início da falta não foi um parâmetro que influenciou de maneira significativa o F1-score médio das 10 melhores RNs.

Tabela 7.2: Médias das métricas de desempenho das 10 melhores RNs do par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ para os tempos de início das faltas

Tempo de Início da Falta (s)	F1-score Médio	Precisão Média	Recall Médio	Acurácia Média
0,091	0,9845	1,0000	0,9694	0,9702
0,099	0,9844	1,0000	0,9694	0,9708
0,093	0,9839	1,0000	0,9683	0,9692
0,101	0,9837	1,0000	0,9679	0,9695

Já na Tabela 7.3, que contempla os tipos de falta, notamos que houve uma flutuação maior, mas pequena, nos valores. Tendo nas simulações com o tipo de falta AC, um pior desempenho do que as outras. Porém, quando checamos apenas o desempenho da melhor restrição de negação para cada tipo de falta, notamos que os desempenhos delas se aproximam ainda mais. Isso pode ser visto na Tabela 7.4 e indica que dentre as 10 melhores RNs utilizadas anteriormente, existem aquelas que desempenham um pouco melhor para alguns tipos de falta.

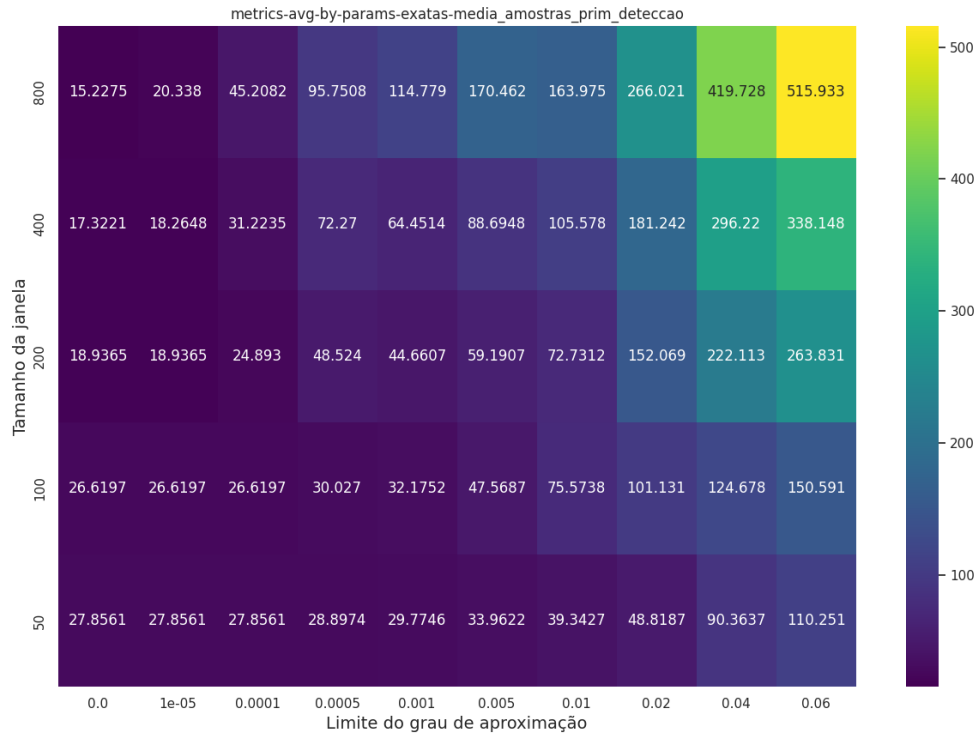


Figura 7.4: Média de amostras após a falta para a primeira detecção de falta das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$

Tabela 7.3: Médias das métricas de desempenho das 10 melhores RNs do par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ para os tipos de faltas

Tipo da Falta	F1-score Médio	Precisão Média	Recall Médio	Acurácia Média
BC	0,9972	1,0000	0,9944	0,9946
ABC	0,9962	1,0000	0,9925	0,9928
ABG	0,9951	1,0000	0,9903	0,9906
BCG	0,9922	1,0000	0,9845	0,9851
ACG	0,9899	1,0000	0,9800	0,9808
BG	0,9860	1,0000	0,9724	0,9735
AG	0,9801	1,0000	0,9610	0,9625
AB	0,9754	1,0000	0,9520	0,9538
CG	0,9665	1,0000	0,9352	0,9377
AC	0,9611	1,0000	0,9251	0,9280

Agora, na Tabela 7.5, vemos que a variação da resistência não pareceu impactar significativamente o desempenho das avaliações. Um resultado que ia contra o que prevíamos, pois, quando temos resistências maiores, as alterações nos valores de corrente e tensão devido a uma falta tendem a ser menos abruptas, o que poderia tornar a detecção das faltas mais difícil.

O efeito da variação da localização da falta na linha de transmissão também foi contra o que imaginávamos. Faltas localizadas mais próximas do barramento em que as amostras foram coletadas têm variações maiores nos valores de tensão e corrente. Isso não foi relevante no

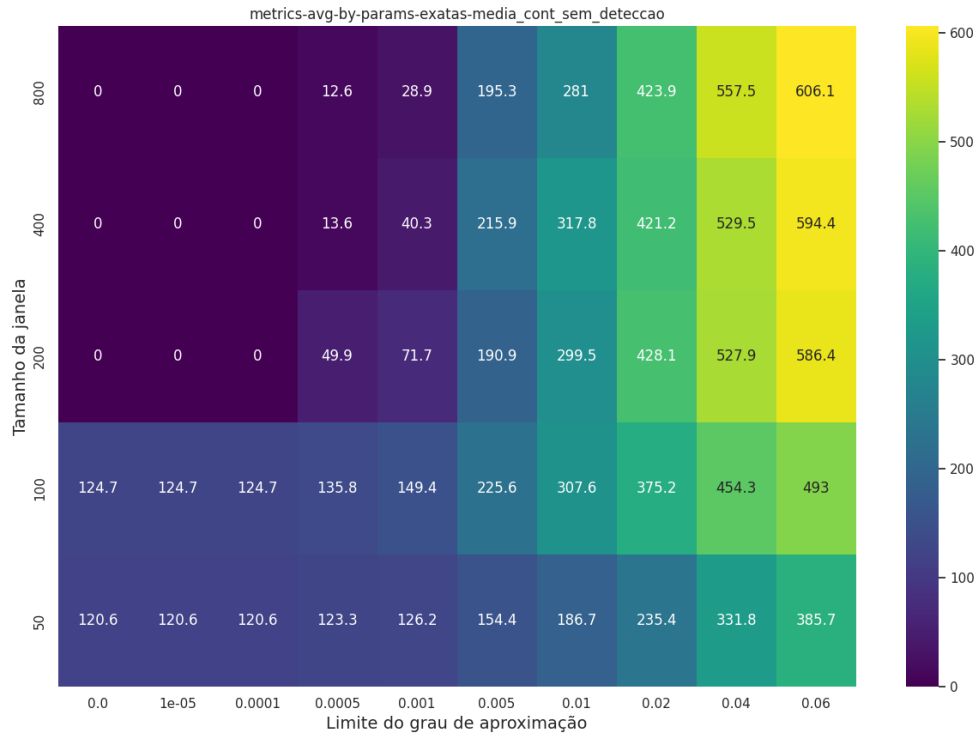


Figura 7.5: Média de simulações que não foram detectadas faltas das 10 melhores RNs para os parâmetros de tamanho da janela e limite do grau de aproximação com avanço $s = 1$

Tabela 7.4: Médias das métricas de desempenho da melhor RN para cada tipo de falta com o par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$

Tipo da Falta	F1-score Médio	Precisão Média	Recall Médio	Acurácia Média
AB	0,9985	1,0000	0,9970	0,9971
ABC	0,9985	1,0000	0,9970	0,9972
ABG	0,9983	1,0000	0,9965	0,9967
AC	0,9990	1,0000	0,9979	0,9980
ACG	0,9991	1,0000	0,9981	0,9982
AG	0,9963	1,0000	0,9926	0,9929
BC	0,9990	1,0000	0,9981	0,9981
BCG	0,9991	1,0000	0,9982	0,9982
BG	0,9973	1,0000	0,9946	0,9948
CG	0,9784	1,0000	0,9576	0,9593

desempenho, como pode ser visto na Tabela 7.6. Inclusive, a pior média foi com a localização em 0%, que é a mais próxima do barramento 8 do qual utilizamos os valores das amostras.

7.2.3 Métricas das Restrições

Nessa seção, averiguaremos a relação entre o desempenho das RNs avaliadas com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$, e as métricas obtidas: cobertura, comprimento, quantidade de lotes sem faltas e com falta em que foram encontradas.

Tabela 7.5: Médias das métricas de desempenho das 10 melhores RNs do par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ para as resistências da linha de transmissão

Resistência (ohm)	F1-score Médio	Precisão Média	Recall Médio	Acurácia Média
200	0,9936	1,0000	0,9874	0,9878
150	0,9908	1,0000	0,9818	0,9825
100	0,9831	1,0000	0,9667	0,9680
1	0,9794	1,0000	0,9597	0,9612
50	0,9734	1,0000	0,9482	0,9502

Tabela 7.6: Médias das métricas de desempenho das 10 melhores RNs do par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ para as localizações das faltas

Localização (%)	F1-score Médio	Precisão Média	Recall Médio	Acurácia Média
100	0,9935	1,0000	0,9870	0,9875
75	0,9917	1,0000	0,9835	0,9841
50	0,9863	1,0000	0,9730	0,9740
25	0,9851	1,0000	0,9707	0,9718
1	0,9635	1,0000	0,9296	0,9323

Observando a relação entre o F1-Score médio das RNs e suas coberturas médias na Tabela 7.6, é possível notar que aquelas com cobertura menor que 0,75 e aquelas com cobertura igual a um, não tiveram desempenho relevante nas avaliações. Essas restrições são as combinações de colunas únicas e as que seguem o padrão descrito na Subseção 7.1.2, ou seja, apenas as restrições equivalentes a dependências de ordem bidirecionais foram úteis para a detecção de faltas.

Agora, quando analisamos a relação entre o F1-Score médio e a quantidade de vezes que cada RN foi encontrada em lotes com falta, que pode ser vista na Tabela 7.7, é possível notar que as que obtiveram resultados relevantes também foram encontradas em menos lotes com falta, mas a diferença não é discrepante o suficiente para tirarmos conclusões.

Quanto as características de comprimento e quantidade vezes em que as RNs foram encontradas em lotes sem falta, não foi detectada nenhuma relação com o desempenho. As Tabelas 7.9 e 7.8 demonstram esses dados.

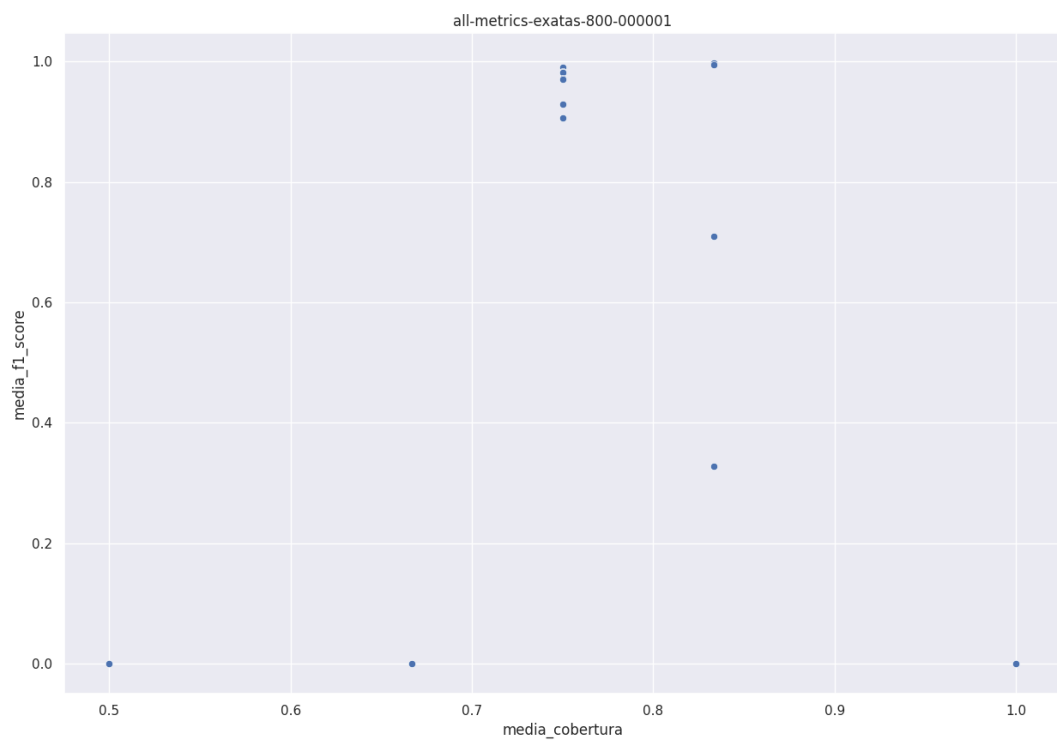


Figura 7.6: Médias do F1-score das RNs com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ relacionadas com a média de suas coberturas

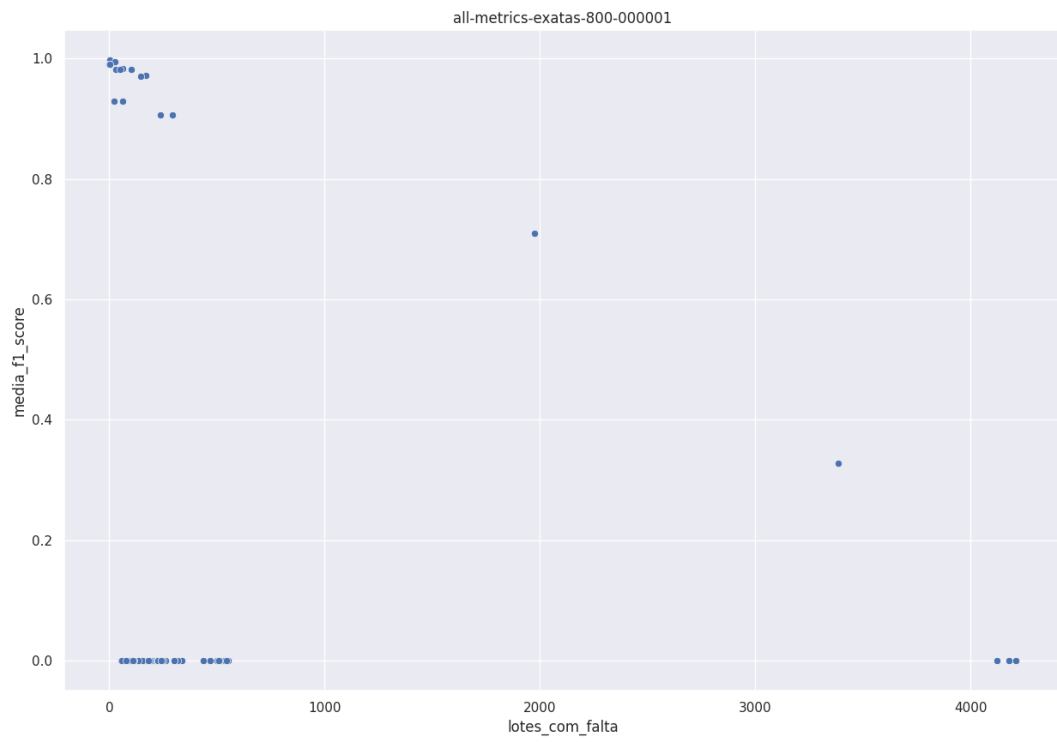


Figura 7.7: Médias do F1-score das RNs com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ relacionadas com a quantidade vezes que cada uma foi encontrada em lotes com falta

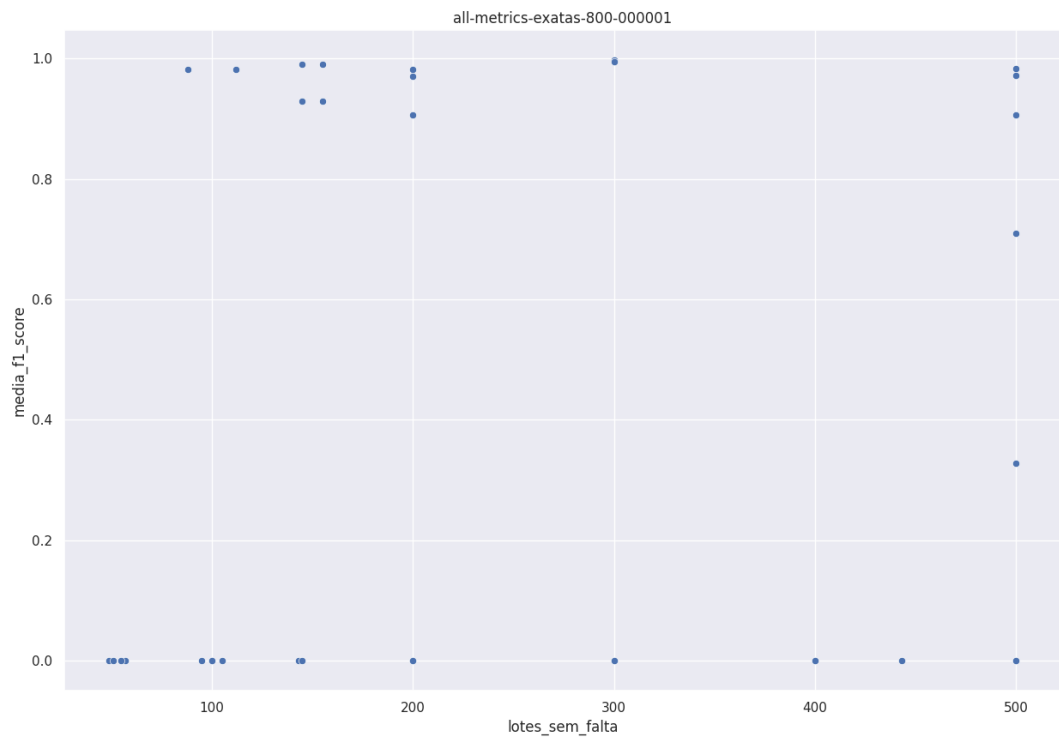


Figura 7.8: Médias do F1-score das RNs com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ relacionadas com a quantidade vezes que cada uma foi encontrada em lotes sem falta

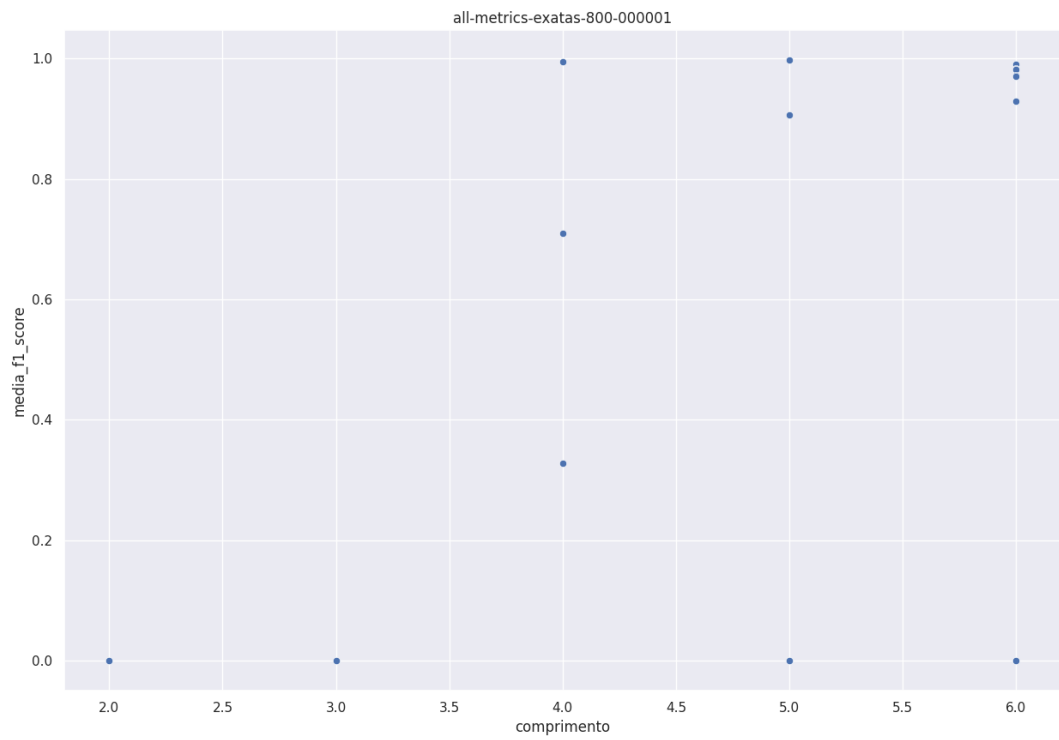


Figura 7.9: Médias do F1-score das RNs com par de parâmetros $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$ relacionadas com seus comprimentos

8 CONCLUSÃO

O uso de restrições de negação para a detecção de faltas em linhas de transmissão se mostrou possível e isso é uma constatação importante, pois é uma maneira nova de detectar faltas diferente das usuais e também é um uso de restrições de negação não habitual.

O método proposto obteve um bom desempenho na detecção de faltas. Com o melhor par de parâmetros, todas as faltas foram detectadas demonstrando a capacidade de detecção em todas as variações das simulações (tipo da falta, tempo de início da falta, localização e tempo de início). A precisão de 100% e a pequena quantidade de amostras necessária para a detecção foi um resultado expressivo e muito importante, pois a precisão na detecção da falta repercutiu diretamente no desempenho dos métodos de classificação e localização de faltas.

No processo de descoberta de restrições de negação o loteamento se mostrou útil para diferenciar as restrições de negação presentes em condições normais das em períodos pós-falta e o particionamento dos atributos teve grande impacto na redução das restrições descobertas e no tempo necessário para descobri-las. Além disso, o uso de apenas um dos barramentos na descoberta das restrições fez com que a detecção de faltas precisasse apenas das oscilografias de um barramento, o que em um ambiente real de operação dispensa a necessidade de sincronização entre os dados coletados em ambos os barramentos.

8.1 SUGESTÕES DE TRABALHOS FUTUROS

As sugestões para trabalhos futuros são:

- Avaliação do uso de restrições de negação para a classificação e localização de faltas;
- Avaliação do desempenho das restrições de negação descobertas nas simulações na detecção de faltas de casos reais;
- Aplicação e avaliação do método proposto em casos de faltas reais;
- Aplicação e avaliação do método proposto utilizando outra medida para o grau de aproximação;
- Aplicação e avaliação do método proposto utilizando o algoritmo C-FASTDC, criado por Chu et al. (2013), já que este algoritmo descobre restrições de negação com valores constantes.

REFERÊNCIAS

- Bichels, A. (2018). *Sistemas Elétricos de Potência: Métodos de Análise e Solução*. Editora da Universidade Tecnológica Federal do Paraná (EDUTFPR).
- Bläsius, T., Friedrich, T. e Schirneck, M. (2017). The Parameterized Complexity of Dependency Detection in Relational Databases. Em Guo, J. e Hermelin, D., editores, *11th International Symposium on Parameterized and Exact Computation (IPEC 2016)*, volume 63 de *Leibniz International Proceedings in Informatics (LIPIcs)*, páginas 6:1–6:13, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- Bleifuß, T., Kruse, S. e Naumann, F. (2017). Efficient denial constraint discovery with hydra. *Proc. VLDB Endow.*, 11(3):311–323.
- Chen, K., Huang, C. e Jinliang, H. (2016). Fault detection, classification and location for transmission lines and distribution systems: A review on the methods. *High Voltage*, 1:25–33.
- Chu, X., Ilyas, I. F. e Papotti, P. (2013). Discovering denial constraints. *Proc. VLDB Endow.*, 6(13):1498–1509.
- Ensina, L. A., Eduardo Soares de Oliveira, L., de Almeida, E. C., Laureano França Santos, S. e Bernardino, L. S. (2022). Fault classification in transmission lines with generalization competence. Em *IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society*, páginas 1–6.
- European EMTP-ATP Users Group (2020). About atp. Acesso em: 29/11/2020. Disponível em: <https://www.eeug.org/index.php/about-eeug/about-atp>.
- F. da Silva, R. (2014). *Trilhamento Elétrico sobre Materiais Poliméricos pela Análise da Energia Absorvida Devido às Descargas Elétricas Superficiais*. Tese de doutorado.
- Fan, W. (2015). Data quality: From theory to practice. *SIGMOD Rec.*, 44(3):7–18.
- Fuchs, R. D. (1977). *Transmissão de Energia Elétrica: Linhas Aéreas*, volume 1. Livros Técnicos e Científicos, Escola Federal de Engenharia Itajubá.
- Høidalen, H. K. (2020). Atpdraw - the graphical preprocessor to atp. Acesso em: 29/11/2020. Disponível em: <http://www.atpdraw.net>.
- Kindermann, G. (2018). *Curto-circuito*. Edição do autor, 6 edition.
- Kossmann, J. e Naumann, F. (2022). Data dependencies for query optimization: a survey. *The VLDB Journal*, 31.
- Labegalini, P. R., Labegalini, J. A., Fuchs, R. D. e de Almeida, M. T. (1992). *Projetos mecânicos das linhas aéreas de transmissão*. Edgard Blucher, 2 edition.
- Livshits, E., Heidari, A., Ilyas, I. F. e Kimelfeld, B. (2020). Approximate denial constraints. *CoRR*, abs/2005.08540.
- Luping, Z. e Huang, X. (2022). Slow time-varying batch process quality prediction based on batch augmentation analysis. *Sensors*, 22:512.

- Mishra, D. e Ray, P. (2018). Fault detection, location and classification of a transmission line. *Neural Computing and Applications*, 30.
- Olmezogullari, E. e Ari, I. (2013). Online association rule mining over fast data. páginas 110–117.
- Papenbrock, T., Bergmann, T., Finke, M., Zwiener, J. e Naumann, F. (2015). Data profiling with metanome. *Proc. VLDB Endow.*, 8(12):1860–1863.
- Pena, E. H. M., de Almeida, E. C. e Naumann, F. (2019). Discovery of approximate (and exact) denial constraints. *Proc. VLDB Endow.*, 13(3):266–278.
- Pena, E. H. M., Lucas Filho, E. R., de Almeida, E. C. e Naumann, F. (2020). Efficient detection of data dependency violations. Em *Proceedings of the 29th ACM International Conference on Information Knowledge Management*, New York, NY, USA. Association for Computing Machinery.
- Raza, A., Benrabah, A., Alquthami, T. e Akmal, M. (2020). A review of fault diagnosing methods in power transmission systems. *Applied Sciences*, 10:1312.
- Sato, F. e Freitas, W. (2015). *Análise de curto-circuito e princípios de proteção em sistemas de energia elétrica*. Elsevier.
- Schneider, J., Wenig, P. e Papenbrock, T. (2021). Distributed detection of sequential anomalies in univariate time series. *The VLDB Journal*, 30(4):579–602.
- Szlichta, J., Godfrey, P., Golab, L., Kargar, M. e Srivastava, D. (2018). Effective and complete discovery of bidirectional order dependencies via set-based axioms. *The VLDB Journal*, 27(4):573–591.
- Yadav, A. e Dash, Y. (2014). An overview of transmission line protection by artificial neural network: Fault detection, fault classification, fault location, and fault direction discrimination. *Advances in Artificial Neural Systems*, 2014.

APÊNDICE A – RESTRIÇÕES DE NEGAÇÃO DESCOBERTAS

Tabela A.1: Restrições de negação, suas métricas na descoberta e F1-score médio com tamanho de janela $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$

Restrição de Negação	F1-score Médio	Cobertura Média	Comprimento	Lotes Sem Falta	Lotes Com Falta
$t0.IA \leq t1.IA, t0.IB \leq t1.IB, t0.VA \geq t1.VA$	0,9974	0,8333	5	300	4
$t0.VC \leq t1.VC, t0.VB \leq t1.VB, t0.IB \leq t1.IB$	0,9939	0,8333	4	300	27
$t0.IC \leq t1.IC, t0.VC \geq t1.VC, t0.VA \leq t1.VA, t0.IB \geq t1.IB$	0,9897	0,7500	6	145	0
$t0.VC \leq t1.VC, t0.IC \geq t1.IC, t0.IB \leq t1.IB, t0.VA \geq t1.VA$	0,9897	0,7500	6	155	1
$t0.IC \leq t1.IC, t0.VC \geq t1.VC, t0.VA \leq t1.VA, t0.IA \leq t1.IA$	0,9829	0,7500	6	500	62
$t0.IA \leq t1.IA, t0.VA \geq t1.VA, t0.VB \leq t1.VB, t0.IB \leq t1.IB$	0,9819	0,7500	6	200	102
$t0.IC \leq t1.IC, t0.VC \geq t1.VC, t0.IB \geq t1.IB, t0.VB \geq t1.VB$	0,9814	0,7500	6	88	32
$t0.VC \leq t1.VC, t0.IC \geq t1.IC, t0.VB \leq t1.VB, t0.IB \leq t1.IB$	0,9814	0,7500	6	112	50
$t0.IC \leq t1.IC, t0.VC \geq t1.VC, t0.IA \leq t1.IA, t0.VB \geq t1.VB$	0,9723	0,7500	6	500	168
$t0.VC \geq t1.VC, t0.IA \leq t1.IA, t0.IB \leq t1.IB, t0.VA \geq t1.VA$	0,9702	0,7500	6	200	148
$t0.IC \leq t1.IC, t0.VA \leq t1.VA, t0.IB \geq t1.IB, t0.VB \leq t1.VB$	0,9297	0,7500	6	145	63
$t0.IC \geq t1.IC, t0.VB \geq t1.VB, t0.IB \leq t1.IB, t0.VA \geq t1.VA$	0,9297	0,7500	6	155	21
$t0.IC \leq t1.IC, t0.VA \leq t1.VA, t0.IA \leq t1.IA, t0.VB \leq t1.VB$	0,9067	0,7500	5	500	239
$t0.VC \leq t1.VC, t0.IA \leq t1.IA, t0.VB \leq t1.VB, t0.IB \leq t1.IB$	0,9063	0,7500	5	200	293
$t0.IC \leq t1.IC, t0.IA \leq t1.IA, t0.IB \leq t1.IB$	0,7098	0,8333	4	500	1975
$t0.VC \leq t1.VC, t0.VA \leq t1.VA, t0.VB \leq t1.VB$	0,3273	0,8333	4	500	3386
$t0.IA = t1.IA$	0,0001	1,0000	2	200	4122
$t0.IA \leq t1.IA, t0.VC = t1.VC, t0.VA \geq t1.VA$	0,0000	0,6667	6	200	184
$t0.IA \leq t1.IA, t0.VC = t1.VC, t0.VB \leq t1.VB$	0,0000	0,6667	5	200	255
$t0.IB = t1.IB$	0,0000	1,0000	2	300	4211
$t0.IB = t1.IB, t0.VB = t1.VB$	0,0000	0,5000	3	200	202
$t0.IC = t1.IC$	0,0000	1,0000	2	200	4178
$t0.IC \leq t1.IC, t0.IA \leq t1.IA, t0.VA = t1.VA$	0,0000	0,6667	5	200	252

Restrição de Negação	F1-score Médio	Cobertura Média	Comprimento	Lotes Sem Falta	Lotes Com Falta
t0.IC <= t1.IC, t0.IA = t1.IA, t0.VB <= t1.VB	0,0000	0,6667	5	100	133
t0.IC <= t1.IC, t0.IA = t1.IA, t0.VB >= t1.VB	0,0000	0,6667	6	100	117
t0.IC <= t1.IC, t0.IA >= t1.IA, t0.VB = t1.VB	0,0000	0,6667	6	95	172
t0.IC >= t1.IC, t0.IA <= t1.IA, t0.VB = t1.VB	0,0000	0,6667	6	105	53
t0.IC <= t1.IC, t0.IA <= t1.IA, t0.VC = t1.VC	0,0000	0,6667	5	200	254
t0.IC = t1.IC, t0.IB = t1.IB	0,0000	0,5000	3	200	152
t0.IC <= t1.IC, t0.IB >= t1.IB, t0.VA = t1.VA	0,0000	0,6667	6	57	75
t0.IC = t1.IC, t0.IB <= t1.IB, t0.VA >= t1.VA	0,0000	0,6667	6	200	172
t0.IC <= t1.IC, t0.IB <= t1.IB, t0.VB = t1.VB	0,0000	0,6667	5	400	465
t0.IC <= t1.IC, t0.VA <= t1.VA, t0.IA = t1.IA	0,0000	0,6667	5	300	338
t0.IC <= t1.IC, t0.VA <= t1.VA, t0.IB = t1.IB	0,0000	0,6667	5	200	236
t0.IC >= t1.IC, t0.VA = t1.VA, t0.IB <= t1.IB	0,0000	0,6667	6	143	158
t0.IC <= t1.IC, t0.VA <= t1.VA, t0.VB = t1.VB	0,0000	0,6667	5	400	463
t0.IC <= t1.IC, t0.VA <= t1.VA, t0.VC = t1.VC	0,0000	0,6667	5	500	548
t0.IC = t1.IC, t0.VB = t1.VB	0,0000	0,5000	3	300	252
t0.IC <= t1.IC, t0.VB >= t1.VB, t0.IB = t1.IB	0,0000	0,6667	6	200	223
t0.IC = t1.IC, t0.VB <= t1.VB, t0.IB <= t1.IB	0,0000	0,6667	5	100	134
t0.IC <= t1.IC, t0.VB >= t1.VB, t0.VA = t1.VA	0,0000	0,6667	6	200	216
t0.IC = t1.IC, t0.VB <= t1.VB, t0.VA >= t1.VA	0,0000	0,6667	6	100	61
t0.IC <= t1.IC, t0.VB >= t1.VB, t0.VC = t1.VC	0,0000	0,6667	6	500	525
t0.IC = t1.IC, t0.VC = t1.VC	0,0000	0,5000	3	300	250
t0.IC <= t1.IC, t0.VC >= t1.VC, t0.IA = t1.IA	0,0000	0,6667	6	200	214
t0.IC <= t1.IC, t0.VC <= t1.VC, t0.IB = t1.IB	0,0000	0,6667	5	200	239
t0.IC <= t1.IC, t0.VC = t1.VC, t0.IB <= t1.IB	0,0000	0,6667	5	500	549
t0.IC <= t1.IC, t0.VC <= t1.VC, t0.VA = t1.VA	0,0000	0,6667	5	200	233
t0.IC <= t1.IC, t0.VC >= t1.VC, t0.VB = t1.VB	0,0000	0,6667	6	400	435
t0.VA <= t1.VA, t0.IA <= t1.IA, t0.IC = t1.IC	0,0000	0,6667	5	200	185
t0.VA <= t1.VA, t0.IA <= t1.IA, t0.VB = t1.VB	0,0000	0,6667	5	200	233
t0.VA <= t1.VA, t0.IA = t1.IA, t0.VB <= t1.VB	0,0000	0,6667	5	100	153

Restrição de Negação	F1-score Médio	Cobertura Média	Comprimento	Lotes Sem Falta	Lotes Com Falta
t0.VA = t1.VA, t0.IB = t1.IB	0,0000	0,5000	3	200	200
t0.VA >= t1.VA, t0.IB <= t1.IB, t0.VB = t1.VB	0,0000	0,6667	6	500	529
t0.VA = t1.VA, t0.IC = t1.IC	0,0000	0,5000	3	300	250
t0.VA <= t1.VA, t0.IC = t1.IC, t0.VB <= t1.VB	0,0000	0,6667	5	100	194
t0.VA = t1.VA, t0.VB = t1.VB	0,0000	0,5000	3	500	503
t0.VA <= t1.VA, t0.VB <= t1.VB, t0.IB = t1.IB	0,0000	0,6667	5	200	263
t0.VA = t1.VA, t0.VC = t1.VC	0,0000	0,5000	3	500	502
t0.VB >= t1.VB, t0.IA = t1.IA, t0.IB <= t1.IB	0,0000	0,6667	6	100	118
t0.VB >= t1.VB, t0.IC = t1.IC, t0.IB <= t1.IB	0,0000	0,6667	6	200	168
t0.VB >= t1.VB, t0.VA = t1.VA, t0.IB <= t1.IB	0,0000	0,6667	6	500	529
t0.VC <= t1.VC, t0.IA = t1.IA, t0.IB <= t1.IB	0,0000	0,6667	5	200	226
t0.VC >= t1.VC, t0.IA <= t1.IA, t0.IB = t1.IB	0,0000	0,6667	6	200	215
t0.VC >= t1.VC, t0.IA <= t1.IA, t0.IC = t1.IC	0,0000	0,6667	6	200	169
t0.VC >= t1.VC, t0.IA <= t1.IA, t0.VA = t1.VA	0,0000	0,6667	6	443	467
t0.VC <= t1.VC, t0.IA = t1.IA, t0.VB <= t1.VB	0,0000	0,6667	5	200	232
t0.VC >= t1.VC, t0.IA <= t1.IA, t0.VB = t1.VB	0,0000	0,6667	6	200	217
t0.VC >= t1.VC, t0.IA = t1.IA, t0.VB <= t1.VB	0,0000	0,6667	6	49	77
t0.VC = t1.VC, t0.IB = t1.IB	0,0000	0,5000	3	200	200
t0.VC = t1.VC, t0.IB <= t1.IB, t0.IA >= t1.IA	0,0000	0,6667	6	145	163
t0.VC <= t1.VC, t0.IB = t1.IB, t0.VA >= t1.VA	0,0000	0,6667	6	200	223
t0.VC = t1.VC, t0.IB <= t1.IB, t0.VA >= t1.VA	0,0000	0,6667	6	500	515
t0.VC <= t1.VC, t0.IB <= t1.IB, t0.VB = t1.VB	0,0000	0,6667	5	200	537
t0.VC <= t1.VC, t0.IC = t1.IC, t0.IB <= t1.IB	0,0000	0,6667	5	200	188
t0.VC <= t1.VC, t0.IC = t1.IC, t0.VB <= t1.VB	0,0000	0,6667	5	200	188
t0.VC <= t1.VC, t0.VA <= t1.VA, t0.IA = t1.IA	0,0000	0,6667	5	200	250
t0.VC <= t1.VC, t0.VA = t1.VA, t0.IA >= t1.IA	0,0000	0,6667	6	57	79
t0.VC <= t1.VC, t0.VA = t1.VA, t0.IB <= t1.IB	0,0000	0,6667	5	500	550
t0.VC <= t1.VC, t0.VA <= t1.VA, t0.IC = t1.IC	0,0000	0,6667	5	100	93
t0.VC = t1.VC, t0.VB = t1.VB	0,0000	0,5000	3	500	501

Restrição de Negação	F1-score Médio	Cobertura Média	Comprimento	Lotes Sem Falta	Lotes Com Falta
t0.VC <= t1.VC, t0.VB >= t1.VB, t0.IA = t1.IA	0,0000	0,6667	6	51	60
t0.VC <= t1.VC, t0.VB <= t1.VB, t0.IB = t1.IB	0,0000	0,6667	5	100	236
t0.VC = t1.VC, t0.VB <= t1.VB, t0.IB <= t1.IB	0,0000	0,6667	5	200	508
t0.IA = t1.IA, t0.IB = t1.IB	0,0000	0,5000	3	100	103
t0.IA <= t1.IA, t0.IB = t1.IB, t0.VA >= t1.VA	0,0000	0,6667	6	100	226
t0.IA = t1.IA, t0.IB <= t1.IB, t0.VA >= t1.VA	0,0000	0,6667	6	100	316
t0.IA <= t1.IA, t0.IB <= t1.IB, t0.VB = t1.VB	0,0000	0,6667	5	200	243
t0.IA <= t1.IA, t0.IB >= t1.IB, t0.VC = t1.VC	0,0000	0,6667	6	55	77
t0.IA = t1.IA, t0.IC = t1.IC	0,0000	0,5000	3	200	154
t0.IA <= t1.IA, t0.IC = t1.IC, t0.VB <= t1.VB	0,0000	0,6667	5	200	183
t0.IA = t1.IA, t0.VA = t1.VA	0,0000	0,5000	3	300	301
t0.IA <= t1.IA, t0.VA = t1.VA, t0.IB <= t1.IB	0,0000	0,6667	5	200	551
t0.IA <= t1.IA, t0.VA = t1.VA, t0.VB <= t1.VB	0,0000	0,6667	5	500	545
t0.IA = t1.IA, t0.VB = t1.VB	0,0000	0,5000	3	300	301
t0.IA <= t1.IA, t0.VB <= t1.VB, t0.IB = t1.IB	0,0000	0,6667	5	200	240
t0.IA = t1.IA, t0.VB <= t1.VB, t0.IB <= t1.IB	0,0000	0,6667	5	100	136
t0.IA <= t1.IA, t0.VB >= t1.VB, t0.IC = t1.IC	0,0000	0,6667	6	100	113
t0.IA = t1.IA, t0.VB <= t1.VB, t0.VA >= t1.VA	0,0000	0,6667	6	100	111
t0.IA = t1.IA, t0.VC = t1.VC	0,0000	0,5000	3	300	300

APÊNDICE B – RESTRIÇÕES DE NEGAÇÃO E SUAS MÉTRICAS DE DESEMPENHO NA DETECÇÃO DE FALTAS

Tabela B.1: Restrições de negação e suas médias nas métricas de desempenho com tamanho da janela de $w = 800$ e limite do grau de aproximação de $1 * 10^{-5}$

Restrição de Negação	F1-score	Precisão	Recall	Acurária	Amostras 1ª Detecção	Não Detectou Falta
$t0.IA \leq t1.IA, t0.IB \leq t1.IB, t0.VA \geq t1.VA$	0,9974	1,0000	0,9948	0,9950	12,3820	0
$t0.VC \leq t1.VC, t0.VB \leq t1.VB, t0.IB \leq t1.IB$	0,9939	1,0000	0,9878	0,9883	11,9510	0
$t0.IC \leq t1.IC, t0.VC \geq t1.VC, t0.VA \leq t1.VA, t0.IB \geq t1.IB$	0,9897	1,0000	0,9795	0,9803	10,8840	0
$t0.VC \leq t1.VC, t0.IC \geq t1.IC, t0.IB \leq t1.IB, t0.VA \geq t1.VA$	0,9897	1,0000	0,9795	0,9803	10,8840	0
$t0.IC \leq t1.IC, t0.VC \geq t1.VC, t0.VA \leq t1.VA, t0.IA \leq t1.IA$	0,9829	1,0000	0,9664	0,9676	22,8070	0
$t0.IA \leq t1.IA, t0.VA \geq t1.VA, t0.VB \leq t1.VB, t0.IB \leq t1.IB$	0,9819	1,0000	0,9644	0,9657	19,5090	0
$t0.IC \leq t1.IC, t0.VC \geq t1.VC, t0.IB \geq t1.IB, t0.VB \geq t1.VB$	0,9814	1,0000	0,9634	0,9648	22,7340	0
$t0.VC \leq t1.VC, t0.IC \geq t1.IC, t0.VB \leq t1.VB, t0.IB \leq t1.IB$	0,9814	1,0000	0,9634	0,9648	22,7330	0
$t0.IC \leq t1.IC, t0.VC \geq t1.VC, t0.IA \leq t1.IA, t0.VB \geq t1.VB$	0,9723	1,0000	0,9461	0,9482	32,0750	0
$t0.VC \geq t1.VC, t0.IA \leq t1.IA, t0.IB \leq t1.IB, t0.VA \geq t1.VA$	0,9702	1,0000	0,9422	0,9444	37,4210	0
$t0.IC \leq t1.IC, t0.VA \leq t1.VA, t0.IB \geq t1.IB, t0.VB \leq t1.VB$	0,9297	1,0000	0,8686	0,8736	11,6340	0
$t0.IC \geq t1.IC, t0.VB \geq t1.VB, t0.IB \leq t1.IB, t0.VA \geq t1.VA$	0,9297	1,0000	0,8686	0,8736	11,6340	0
$t0.IC \leq t1.IC, t0.VA \leq t1.VA, t0.IA \leq t1.IA, t0.VB \leq t1.VB$	0,9067	1,0000	0,8294	0,8359	37,0410	0
$t0.VC \leq t1.VC, t0.IA \leq t1.IA, t0.VB \leq t1.VB, t0.IB \leq t1.IB$	0,9063	1,0000	0,8286	0,8351	36,1750	0
$t0.IC \leq t1.IC, t0.IA \leq t1.IA, t0.IB \leq t1.IB$	0,7098	1,0000	0,5501	0,5672	18,1483	400
$t0.VC \leq t1.VC, t0.VA \leq t1.VA, t0.VB \leq t1.VB$	0,3273	1,0000	0,1957	0,2263	18,4600	400
$t0.IA = t1.IA$	0,0001	1,0000	0,0000	0,0381	259,0000	999
$t0.IA \leq t1.IA, t0.VC = t1.VC, t0.VA \geq t1.VA$	0,0000	0,0000	0,0000	0,0381		1000
$t0.IA \leq t1.IA, t0.VC = t1.VC, t0.VB \leq t1.VB$	0,0000	0,0000	0,0000	0,0381		1000
$t0.IB = t1.IB$	0,0000	0,0000	0,0000	0,0381		1000
$t0.IB = t1.IB, t0.VB = t1.VB$	0,0000	0,0000	0,0000	0,0381		1000
$t0.IC = t1.IC$	0,0000	0,0000	0,0000	0,0381		1000
$t0.IC \leq t1.IC, t0.IA \leq t1.IA, t0.VA = t1.VA$	0,0000	0,0000	0,0000	0,0381		1000

Restrição de Negação	F1-score	Precisão	Recall	Acurácia	Amostras 1ª Detecção	Não Detectou Falta
t0.IC <= t1.IC, t0.IA = t1.IA, t0.VB <= t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.IA = t1.IA, t0.VB >= t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.IA >= t1.IA, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC >= t1.IC, t0.IA <= t1.IA, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.IA <= t1.IA, t0.VC = t1.VC	0,0000	0,0000	0,0000	0,0381		1000
t0.IC = t1.IC, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.IB >= t1.IB, t0.VA = t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IC = t1.IC, t0.IB <= t1.IB, t0.VA >= t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.IB <= t1.IB, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VA <= t1.VA, t0.IA = t1.IA	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VA <= t1.VA, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC >= t1.IC, t0.VA = t1.VA, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VA <= t1.VA, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VA <= t1.VA, t0.VC = t1.VC	0,0000	0,0000	0,0000	0,0381		1000
t0.IC = t1.IC, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VB >= t1.VB, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC = t1.IC, t0.VB <= t1.VB, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VB >= t1.VB, t0.VA = t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IC = t1.IC, t0.VB <= t1.VB, t0.VA >= t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VB >= t1.VB, t0.VC = t1.VC	0,0000	0,0000	0,0000	0,0381		1000
t0.IC = t1.IC, t0.VC = t1.VC	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VC >= t1.VC, t0.IA = t1.IA	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VC <= t1.VC, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VC = t1.VC, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VC <= t1.VC, t0.VA = t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IC <= t1.IC, t0.VC >= t1.VC, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VA <= t1.VA, t0.IA <= t1.IA, t0.IC = t1.IC	0,0000	0,0000	0,0000	0,0381		1000
t0.VA <= t1.VA, t0.IA <= t1.IA, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VA <= t1.VA, t0.IA = t1.IA, t0.VB <= t1.VB	0,0000	0,0000	0,0000	0,0381		1000

Restrição de Negação	F1-score	Precisão	Recall	Acurácia	Amostras 1ª Detecção	Não Detectou Falta
t0.VA = t1.VA, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VA >= t1.VA, t0.IB <= t1.IB, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VA = t1.VA, t0.IC = t1.IC	0,0000	0,0000	0,0000	0,0381		1000
t0.VA <= t1.VA, t0.IC = t1.IC, t0.VB <= t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VA = t1.VA, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VA <= t1.VA, t0.VB <= t1.VB, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VA = t1.VA, t0.VC = t1.VC	0,0000	0,0000	0,0000	0,0381		1000
t0.VB >= t1.VB, t0.IA = t1.IA, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VB >= t1.VB, t0.IC = t1.IC, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VB >= t1.VB, t0.VA = t1.VA, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.IA = t1.IA, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC >= t1.VC, t0.IA <= t1.IA, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC >= t1.VC, t0.IA <= t1.IA, t0.IC = t1.IC	0,0000	0,0000	0,0000	0,0381		1000
t0.VC >= t1.VC, t0.IA <= t1.IA, t0.VA = t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.IA = t1.IA, t0.VB <= t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC >= t1.VC, t0.IA <= t1.IA, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC >= t1.VC, t0.IA = t1.IA, t0.VB <= t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC = t1.VC, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC = t1.VC, t0.IB <= t1.IB, t0.IA >= t1.IA	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.IB = t1.IB, t0.VA >= t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.VC = t1.VC, t0.IB <= t1.IB, t0.VA >= t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.IB <= t1.IB, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.IC = t1.IC, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.IC = t1.IC, t0.VB <= t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.VA <= t1.VA, t0.IA = t1.IA	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.VA = t1.VA, t0.IA >= t1.IA	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.VA = t1.VA, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.VA <= t1.VA, t0.IC = t1.IC	0,0000	0,0000	0,0000	0,0381		1000
t0.VC = t1.VC, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000

Restrição de Negação	F1-score	Precisão	Recall	Acurácia	Amostras 1ª Detecção	Não Detectou Falta
t0.VC <= t1.VC, t0.VB >= t1.VB, t0.IA = t1.IA	0,0000	0,0000	0,0000	0,0381		1000
t0.VC <= t1.VC, t0.VB <= t1.VB, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.VC = t1.VC, t0.VB <= t1.VB, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA = t1.IA, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA <= t1.IA, t0.IB = t1.IB, t0.VA >= t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IA = t1.IA, t0.IB <= t1.IB, t0.VA >= t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IA <= t1.IA, t0.IB <= t1.IB, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA <= t1.IA, t0.IB >= t1.IB, t0.VC = t1.VC	0,0000	0,0000	0,0000	0,0381		1000
t0.IA = t1.IA, t0.IC = t1.IC	0,0000	0,0000	0,0000	0,0381		1000
t0.IA <= t1.IA, t0.IC = t1.IC, t0.VB <= t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA = t1.IA, t0.VA = t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IA <= t1.IA, t0.VA = t1.VA, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA <= t1.IA, t0.VA = t1.VA, t0.VB <= t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA = t1.IA, t0.VB = t1.VB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA <= t1.IA, t0.VB <= t1.VB, t0.IB = t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA = t1.IA, t0.VB <= t1.VB, t0.IB <= t1.IB	0,0000	0,0000	0,0000	0,0381		1000
t0.IA <= t1.IA, t0.VB >= t1.VB, t0.IC = t1.IC	0,0000	0,0000	0,0000	0,0381		1000
t0.IA = t1.IA, t0.VB <= t1.VB, t0.VA >= t1.VA	0,0000	0,0000	0,0000	0,0381		1000
t0.IA = t1.IA, t0.VC = t1.VC	0,0000	0,0000	0,0000	0,0381		1000