

# Método otimizado para localização de faltas em redes Inteligentes de distribuição de energia elétrica

José Carlos Filho<sup>1</sup>

<sup>1</sup>Ciência da Computação – Universidade Federal do Piauí (UFPI)  
Teresina – PI – Brazil

jcarloslimafilho@ufpi.edu.br

**Resumo.** *O sistema de energia elétrica vem sofrendo grandes transformações saindo do fluxo de potência unidirecional do sistema convencional, partindo para o fluxo bidirecional, onde a potência e a informação fluem entre os lados de geração e distribuição, migrando para um sistema de rede inteligente o SmartGrid. Para a construção de uma estrutura de coleta de dados e gerenciamento de informações, há uma implantação em larga escala de medidores inteligentes, em especial o uso difundido de tecnologias de detecção e medição, como unidades de medição síncronas (PMUs). A quantidade de dados gerados na rede inteligente está na 4ª ordem de magnitude maior do que na rede tradicional. Esses dados são diversos em estrutura, complexos em fonte e têm características típicas de Big Data conhecidas como "4V", ou seja, grande volume, variedade, valor e velocidade. Um dos problemas que utiliza essa estrutura é a localização de falhas. O problema de localização tem como objetivo indicar a área ou região ou distância da subestação que está conectada à fonte causadora da falha e consiste em uma tarefa fundamental para garantir uma boa qualidade de energia elétrica. Este trabalho tem como objetivo desenvolver um método para localização de faltas em redes inteligentes de distribuição de energia elétrica de forma rápida e precisa se preocupando características de um recurso muito utilizado nos métodos atuais a Big Data.*

## 1. Introdução

O sistema de energia elétrica vem sofrendo grandes transformações saindo do fluxo de potência unidirecional do sistema convencional, partindo para o fluxo bidirecional, onde a potência e a informação fluem entre os lados de geração e distribuição, migrando para um sistema de rede inteligente o SmartGrid (SG), esses sistemas são baseados em tecnologias de comunicação e informação [Ghorbanian et al. 2019, Hossain et al. 2019]. Esse novo modelo de rede exige novas abordagens de Operações e Planejamento, acompanhando melhorias na qualidade do serviço. É de extrema importância a atuação do Sistema de Gerenciamento da Distribuição, por meio de seu Sistema de Gerenciamento de Falhas, para garantir a confiabilidade da rede. Neste sentido, é fundamental saber onde acontece uma falha para auxiliar a tomada de decisão gerando uma ação adequada para reação deste sistema [Personal et al. 2016]. As falhas em uma rede elétrica diminuem a qualidade da energia e afetam o desempenho dos motores de indução mais utilizados pelas indústrias. A perturbação mais comum, a queda de tensão, é uma redução na tensão com duração de algumas centenas de milissegundos que pode interromper os processos

de produção. Equipamentos sensíveis em uso industrial moderno são propensos a falhas, levando a perdas econômicas e exigindo ainda o procedimento de reinicialização. [Huchche et al. 2017].

A detecção de anomalias em redes inteligentes depende principalmente dos dados coletados pelos medidores inteligentes. Devido à capacidade de comunicação rápida dos medidores, as empresas de energia podem detectar anomalias elétricas rapidamente e tomar medidas relevantes para evitar perdas econômicas desnecessárias, já localização rápida e precisa da posição da falta reduz o tempo de restauração, o tempo de interrupção do fornecimento e geralmente melhora a confiabilidade do sistema [Arsoniadis et al. 2021, Wen et al. 2018, Galijasevic and Abur 2002]

Para a construção dessa estrutura de coleta de dados e gerenciamento de informações, há uma implantação em larga escala de medidores inteligentes, em especial o uso difundido de tecnologias de detecção e medição, como unidades de medição síncronas (PMUs). De acordo com [Wang et al. 2020] os resultados de pesquisas mostram que a quantidade de dados gerados na rede inteligente está na 4ª ordem de magnitude maior do que na rede tradicional. Esses dados são diversos em estrutura, complexos em fonte e têm características típicas de Big Data conhecidas como "4V", ou seja, grande volume, variedade, valor e velocidade [Wang et al. 2020, Khan et al. 2021]. Usar esses dados para fornecer tomada de decisão científica para o desenvolvimento e controle de operação da rede elétrica não é apenas uma necessidade urgente para o desenvolvimento de redes inteligentes, mas também uma maneira de alcançar a força, autocura, compatibilidade, economia, integração e otimização de redes inteligentes [Zhao and Li 2016, Wei-jing et al. 2019].

### **1.1. Revisão da literatura**

O problema de localização tem como objetivo indicar a área ou região ou distância da subestação que está conectada à fonte causadora da falha e consiste em uma tarefa fundamental para garantir uma boa qualidade de energia elétrica. No entanto, a localização não é uma tarefa trivial, pois as falhas geralmente ocorrem em um curto intervalo de tempo e se propagam por todo o sistema [Miranda Filho et al. 2016, Chang et al. 2007, Zheng et al. 2020]. Na literatura foram encontradas várias estratégias utilizadas para determinar a localização das fontes causadoras de falhas no sistema elétrico entre eles métodos baseados em impedância, baseados em ondas viajantes e métodos baseados em aprendizado de máquina [Dashti et al. 2021].

Os métodos baseados em impedância pode ser afetado por diversos fatores, incluindo tipo de falha, cargas desequilibradas, heterogeneidade de linhas aéreas e erros de medição, quando se usa as informações de corrente de carga nas fases sem falha, a distância de falha calculada torna-se imprecisa, pois a corrente de carga nas fases com falha é desprezada. Já os métodos baseados em ondas viajantes requerem altas taxas de amostragem e sobrecarga de comunicação dos dispositivos de medição. Além disso é necessário o uso de sistema como o de posicionamento global (GPS) para a sincronização de tempo em sinais multiterminais. [Chen et al. 2019, Mohammadi et al. 2017, Mohammadi and Leborgne 2020] Já os métodos baseados em aprendizado de máquina aparece como a ferramenta necessária que de acordo com [Hossain et al. 2019] por uma questão óbvia que surge da geração de Big Data a par-

tir de Smart Grid são as formas mais eficientes de analisá-los para extrair informações valiosas.

No trabalho de [Wu et al. 2021] é projetado um sistema de localização de faltas baseado em máquina de vetores de suporte (SVM). A forma da onda viajante após a ocorrência da falha é coletada como um recurso e o mecanismo de regressão do SVM é utilizado para obter a localização da falha. O artigo analisa primeiro as características da forma de onda das falhas de alta impedância. A seguir, três etapas são propostas para reduzir a influência da resistência de aterramento na localização da falta. O sistema de localização proposto só precisa obter uma forma de onda de tensão de falha de terminação única em uma frequência de registro de falha de 20 kHz para obter uma localização precisa de falhas de uma fase à terra para diferentes valores de resistência de aterramento. Contudo no contexto de Big Data, devido ao tamanho dos dados, a probabilidade de ocorrer desequilíbrio de classes é alta. Além disso, por causa dos problemas complexos embutidos em tais dados, os efeitos potenciais do desequilíbrio de classes no aprendizado de máquina são graves [L'heureux et al. 2017].

Uma estratégia que vem crescendo é a utilização de aprendizado profundo que se baseia em redes neurais artificiais, especificamente redes neurais convolucionais. O trabalho de [Zhang et al. 2021] é desenvolvido um novo método de aprendizado profundo com tradução de imagem para imagem (Pix2Pix) para localização de falhas baseadas em ondas viajantes. O método proposto pode traduzir a imagem do componente fornecida pelos dados da PMU de baixa frequência para outras de alta frequência por meio do Pix2Pix. Isso nos permite melhorar significativamente a precisão da localização da falha. A utilização de aprendizagem profunda no contexto de Big Data, é um grande trunfo, pois torna os algoritmos menos suscetíveis a desafios de veracidade, como dados sujos, com ruído incertos. No entanto, existem desvantagens na imagem, pois esses procedimentos não são projetados para aprender progressivamente, tornando-os vulneráveis ao problema de velocidade de dados [L'heureux et al. 2017]. O trabalho de [Chen et al. 2019] desenvolve uma nova estrutura de rede convolucional de grafos (GCN) para localização de faltas em redes de distribuição de energia. A abordagem proposta integra várias medições em diferentes barramentos, levando em consideração a topologia do sistema. Os resultados da simulação mostram uma precisão de localização de falha muito alta. Além disso, a abordagem proposta é robusta para erros de medição de ruído e perda de dados uma outra abordagem e vista em [Li et al. 2019] que propõe um método de localização de linhas com defeito baseado em um classificador de rede neural convolucional (CNN) usando tensões de barramento. O classificador proposto é baseado em recursos com interpretações físicas que melhoram a robustez do desempenho de localização. Para melhorar ainda mais o desempenho da localização, uma estratégia conjunta de colocação de unidades de medição fasorial é proposta. É possível observar que os métodos que utilizam CNN se apresentam de forma eficiente em relação a precisão da localização porem todos utilizam todas as barras no sistema para extrair as informações o que pode gerar problemas característicos da big data como volume e velocidade, além disso um sistema complexo como um sistema elétrico de potencia com um grande número de pontos pode exigir um modelo CNN mais profundo. Com isso, a escalabilidade é difícil de manusear para selecionar um modelo eficiente com um número ideal de parâmetros treináveis [Tikariha et al. 2021].

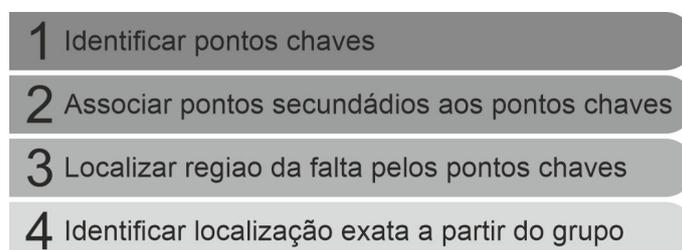
## 1.2. Contribuições

De acordo com as exposições das limitações dos estudos revisados acima, este trabalho tem como objetivo desenvolver um método para localização de faltas em redes inteligentes de distribuição de energia elétrica de forma otimizada reduzindo o grande volume de informação, aprimorando a velocidade do processamento, mantendo a qualidade do valor da informação que trafega na rede e se adaptando a variedade dos dados.

Entre os trabalhos revisados observa-se que os métodos necessitam para seu bom funcionamento um grande volume de informações do sistema utilizando todos os pontos da rede em todos os momentos de medição e na proposta apresentada dividir a rede em regiões, ou seja, continua usando todos os pontos da rede porém nos momentos necessários e para que isso aconteça o primeiro passo é identificar os nós (medidores) chaves que refletem o resultado de uma falta em uma determinada região, para essa identificação foi utilizado um algoritmo genético onde a função de escolha reflete dados característicos como RMS, valor de pico, juntamente com dados estatístico. Em seguida é associado os demais nós chamados de nós secundários aos nós chaves definindo a relação de todos os nós por região, finalizando com a aplicação de uma rede neural convolucional que consegue extrair automaticamente características dos dados garantindo a variedade, aplicação essa com somente nós pertencentes da região atingida pela falta, reduzindo o volume de dados, com menos dados maior velocidade de processamento, sem contar com outros benefícios que a redução dos dados podem proporcionar não abordados, como aplicação de técnicas de compressão de dados e criptografia. Comparações com a acurácia de outros métodos de localização da falta no sistema elétrico serão analisados para garantir o valor dos dados.

## 2. Método Proposto

O método proposto divide o sistema de distribuição em grupos, ocorrendo uma falta identificada em qual dos grupos apresentou o distúrbio, depois é realizada uma busca profunda entre as barras no sistema quem compõem o grupo indicando a localização exata, para isso o método segue as etapas como mostra a figura 1.



**Figure 1. Etapas da metodologia**

É importante salientar que a primeira etapa sub-sessão 2.1 e a segunda etapa sub-sessão 2.2 não é necessário executar sempre que é identificada uma falta, identificado os pontos chaves e os secundários só é necessário executar novamente caso aconteça alguma alteração na topologia da rede. Nas sub sessões a seguir iremos detalhar cada uma das etapas.

## 2.1. Identificar pontos chaves

Na primeira etapa do método são selecionados os pontos chaves. Quando acontece uma falta no sistema os efeitos da falta são propagados por todo o sistema como mostra a Figura 2 alterando a leitura dos dados em todos os pontos, utilizando algumas características do evento como RMS, Valor de Pico, Fator de Pico, Valor Médio e Distância do ponto à Subestação é possível identificar pontos estratégicos que representam uma região.

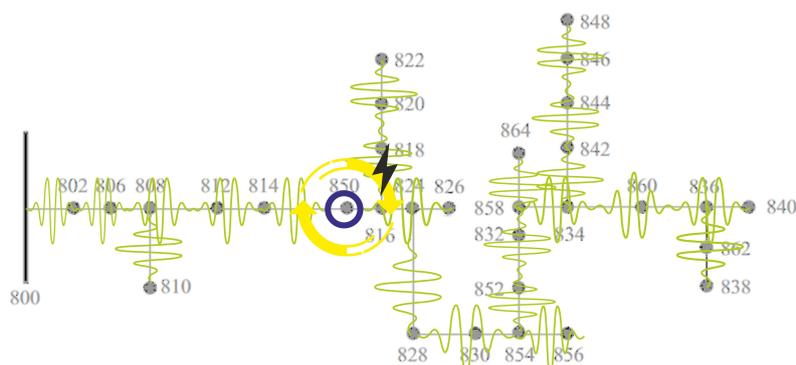


Figure 2. Propagação da falha a partir do ponto 850

A figura 3 mostra um exemplo da seleção dos pontos estratégicos que representa uma região, no exemplo mostra um sistema de teste IEEE de 34 barras onde são selecionados os pontos 802, 824, 828 e 852 como pontos chaves. De acordo com o tamanho do sistema a quantidade de grupo pode variar devido a quantidade mínima de pontos por grupo para execução do algoritmo genético, no caso do sistema de 34 barras a quantidade máxima de grupos são 3. Como a quantidade de pontos chaves são 4 e o número máximo de grupos são 3 podemos observar que um grupo/região pode ter mais de um ponto chave, onde o ponto 802 está circulado em vermelho representando o grupo 1, os pontos 824, 828 circulos de azul representam o grupo 2 e por fim o ponto 852 em laranja representa o grupo 3. A próxima etapa identifica quais são os pontos que compõem um grupo.



Figure 3. Identificação dos pontos chaves

Para a identificação desses pontos é utilizado um algoritmo genético que em sua implementação considera as fases típicas que são: Inicialização (População), Função Objetivo ou Aptidão, Seleção, Crossover e Mutação como mostra a Figura 4.

### 2.1.1. Algoritmo Genético

O Algoritmo Genético (GA) é um dos primeiros algoritmos estocásticos de base populacional proposto na história. Os principais operadores de GA são seleção, cruzamento e mutação.

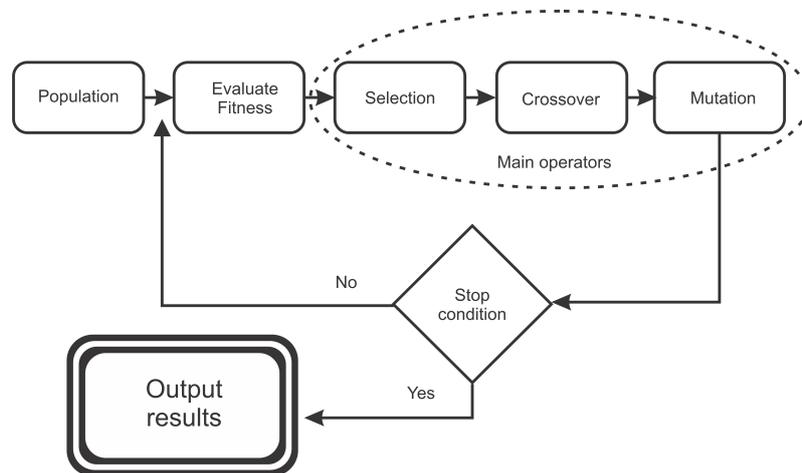


Figure 4. Fases típicas do algoritmo genético

GA é um algoritmo baseado em população. Cada solução corresponde a um cromossomo e cada parâmetro representa um gene. O GA avalia a aptidão de cada indivíduo na população usando uma função de aptidão (objetivo). Para melhorar as soluções pobres, as melhores soluções são escolhidas aleatoriamente com um mecanismo de seleção (por exemplo, roda de roleta). É mais provável que esse operador escolha as melhores soluções, pois a probabilidade é proporcional ao fitness (valor objetivo). [Mirjalili 2019].

### 2.1.2. Implementação

Na inicialização a criação da população é realizada de forma aleatória onde cada indivíduo é representado por um vetor. Cada barra do sistema é representado por uma célula deste vetor, com valores entre 0 e 1, onde 0 indica que a barra não foi selecionada e 1 que a barra foi selecionada como mostra a figura 5, o tamanho do vetor vai depender do tamanho da rede.

Individual (Vector of the selected bars)

Bars	802	806	808	812	814	816	824	828	830	832	834	836	840	842	844	846	848	850	852	854	858	860	862
Individual	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0

Figure 5. Vetor das barras selecionadas.

### 2.1.3. Função-objetivo

A função objetivo é aplicado a cada indivíduo para obter a acurácia da localização da falha, a estratégia de localização é dividida em dois momentos. O primeiro momento é

responsável por obter e analisar os clusters gerados, o algoritmo de clusterização é responsável por fazer inferências sobre a propagação do evento na rede de distribuição, dividindo os dados de entrada não rotulados em grupos de acordo com a similaridade entre os distúrbios evidenciados da rede e independente da topologia do sistema. O segundo momento corresponde à aplicação do sistema baseado em regras de decisão, que é responsável por analisar as características de cada um dos clusters gerados e identificar qual deles fornece a região de localização da falta. O processo de avaliação da função objetivo se caracteriza por:

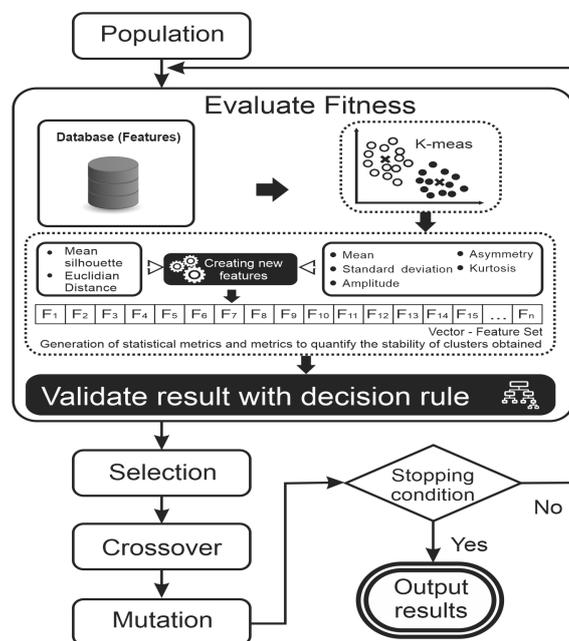
- Os dados processados são apenas das barras selecionadas
- No final da segunda fase 2 se obtém a acurácia subtraída por um peso ( $p$ ):

$$p = bs * \left(\frac{1}{tb}\right) \quad (1)$$

onde  $bs$  é a quantidade de barras selecionadas e  $tb$  e o total de barras, ou seja, o tamanho do vetor.

#### 2.1.4. Peso da função-objetivo

A necessidade da utilização do peso na função-objetivo é para gerar o equilíbrio entre a quantidade de barras e a acurácia da localização, eliminando situações como dois ou mais indivíduos apresentem a melhor acurácia encontrada porem com quantidade de barras selecionadas bem diferentes, o peso garante a escolha da menor quantidade de barras, quanto menor for o valor do peso melhor será a avaliação.

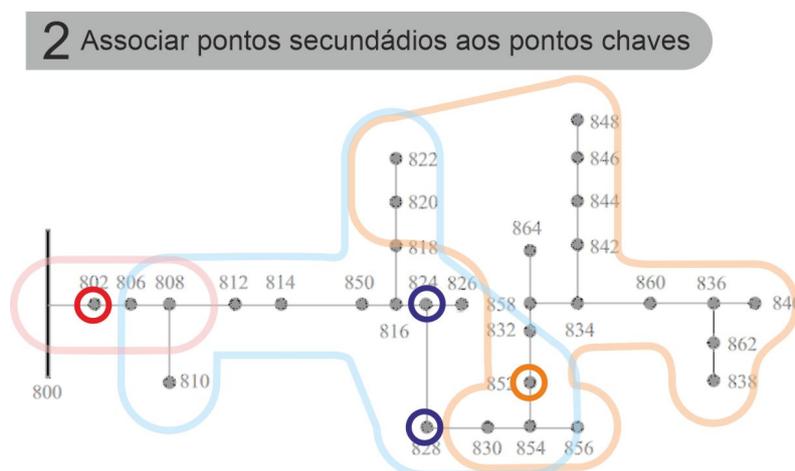


**Figure 6. Estrutura básica de um algoritmo genético com função para avaliar aptidão proposta.**

Todo o processo se repete como mostra a figura 6 até que atinja uma determinada quantidade de gerações (condição de parada do algoritmo utilizado), finalizando com a indicação da quantidade e quais são as barras que devem ser selecionadas.

## 2.2. Associar pontos secundários aos pontos chaves

A segunda etapa ocorre em paralelo com a primeira etapa, como o processo do algoritmo genético em sua função de aptidão (objetivo) utiliza o algoritmo de agrupamento k-means os grupos já são formados, porém como a falta pode ocorrer em qualquer ponto do sistema os pontos pertencentes a uma região pode sobre algumas alterações, com isso alguns pontos secundários podem fazer parte de mais de uma região.



**Figure 7. Divisão das regiões com seus pontos chaves e secundários (Primeira região vermelha, Segunda região azul e a terceira região laranja)**

A figura 7 mostra o sistema IEEE 34 barras a divisão das regiões com seus pontos chaves e secundários. A primeira região é está envolta com uma linha rosa claro e seu ponto chave em vermelho, a segunda região marcada por uma linha azul claro e dois pontos chaves em azul e a última região circundada por uma linha laranja claro e o ponto chave na cor laranja.

## 2.3. Localizar região da falta pelos pontos chaves

A terceira etapa acontece quando é identificado uma falha no sistema, através dos valores obtidos dos pontos chaves no momento do evento aplicado a um algoritmo de classificação é identificado o ponto que melhor reflete as características de uma falha.

Como mostra a Figura 8 o ponto que melhor reflete a falta é o ponto 824, o ponto identificado na figura está circulado com a cor azul, ou seja, a segunda região pode ser visto na Figura 7 é a região onde ocorreu a falta.

## 2.4. Identificação da falta

Depois da terceira etapa descrita na subseção 2.3 podemos isolar a região identificada e aplicar um outro algoritmo somente com os pontos selecionados para encontrar o ponto onde ocorreu a falha. Nessa etapa não existe uma diferenciação entre os pontos Chaves ou secundário.

Na literaturas entre os métodos de localização de falta tem se destacado a utilização da estrutura baseada em Deep-Learning [Ozcanli et al. 2020, Lv et al. 2016]. A rede neural convolucional (CNN) é o modelo de aprendizado profundo mais comum que tem sido amplamente explorado para tarefas de classificação. [Albawi et al. 2017,

### 3 Localizar região da falta pelos pontos chaves

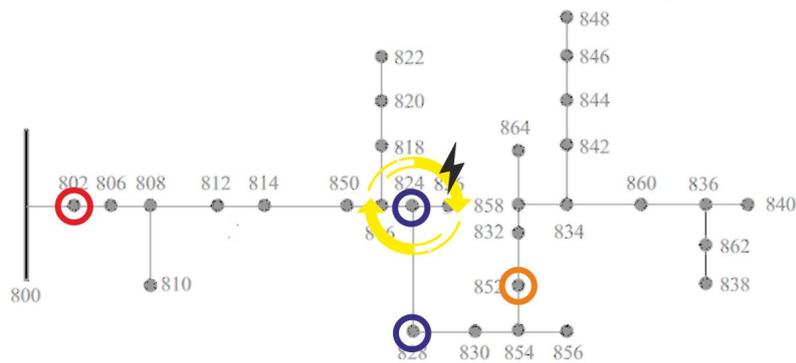


Figure 8. Identificação da Região B através do ponto chave 824

Li et al. 2016], porem todos os métodos encontrados utilizam todos os pontos da rede, neste trabalho vamos utilizar uma rede neural convolucional utilizando apenas os pontos da região isolada, tendo vários ganhos para o problema de big data, como redução do volume de dados e processamento.

a classificação de falhas em sistemas de rede elétrica complexos, com um grande número de pontos e diferentes tipos de cargas pode exigir um modelo CNN mais profundo. Portanto, em tais casos, a escalabilidade é difícil de manusear para selecionar um modelo eficiente com um número ideal de parâmetros treináveis [Tikariha et al. 2021]. Visto isso, quando aumenta o tamanho da rede a quantidade de região também aumenta, mantendo equalizado a quantidade de pontos a ser analisado.

### 4 Identificar localização exata a partir do grupo

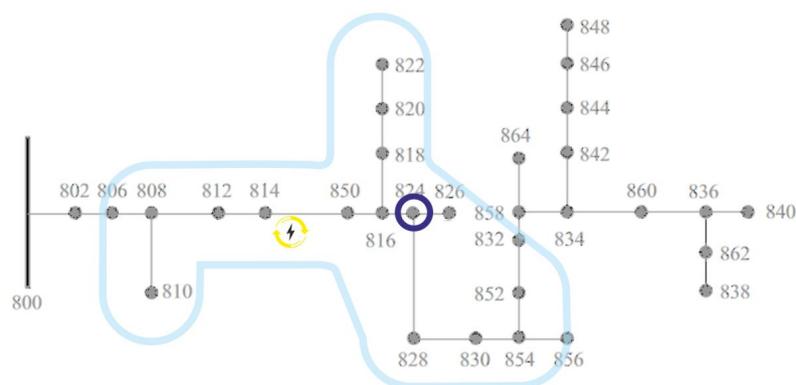


Figure 9. Localização da falta no ponto 814

Como exemplo podemos visualizar na Figura 9 que nessa ultima etapa de todo o sistema iremos utilizar informações de apenas 15 pontos. Como na etapa anterior descrita na subseção 2.3 esses 15 pontos fazem parte da região da falta, e com o algoritmo de redes neurais convolucionais será identificado o ponto 814 com o ponto onde ocorreu a falta.

### 3. Experimentos e resultados

Até o momento foi desenvolvido e simulado até a terceira etapa. Para as simulações e avaliação dessas etapas, foi necessário modelar e simular o Alimentador de Teste IEEE 34-Bus [Feeders 1991] através do software ATP (Alternative Transients Program).

Tendo modelado o sistema, as simulações foram realizada alterando o comportamento do sistema para gerar faltas com diferentes durações e magnitudes. Por isso, foi possível obter um banco de dados para validar a proposta.

Vale ressaltar que todos os afundamentos de tensão foram gerados como resultado de curtos-circuitos trifásicos com duração de falha de 4 ciclos, diferentes impedâncias de falha (10, 20, 30, 40 e 50  $\Omega$ ), com ângulo de incidência de 0.

O curto-circuito é simulado em cada uma das 23 barras trifásicas do sistema, considerando uma taxa de amostragem de 256 pontos/ciclo. Isso dá um total de 115 simulações. Após cada simular os valores de tensão medidos em cada nó do sistema são armazenados. Assim, um banco de informações necessário para testar o algoritmo proposto é obtido. O processo de construção do banco de dados é realizado por meio de comunicação entre ATP® e MATLAB®.

As implementações das rotinas de programação foram desenvolvidas em linguagem python utilizando a biblioteca Pymoo [Blank and Deb 2020] para GA com as seguintes configurações: tamanho da população 30, inicialização aleatória, eliminação de indivíduos duplicados e condição de parada 50 gerações. Para a função de avaliação, o algoritmo de agrupamento utilizado foi o k-means com k igual a 3, a escolha desta configuração foi baseada no desempenho do método apresentado em [Lima Filho et al. 2020], a biblioteca sklearn [Pedregosa et al. 2011] foi utilizada para realizar o algoritmo k-means e obter o valor de acurácia.

#### 3.1. Identificação dos grupos

Neste momento será discutidos os resultados obtidos utilizando a proposta de otimização (GA) para selecionar dentre todas as barras existente no sistema uma quantidade reduzida de barras e identificando locais estratégicos dessas barras (pontos chaves) com intuito de diminuir o volume de dados que será tanto transmitido quanto analisado.

A tabela 1 mostra os melhores resultados de cada geração (melhor indivíduo) com destaque para as gerações 12, 27 e 45 que representam as melhores configurações em relação a quantidade e posições das barras escolhidas, como os critérios de escolha estão empatados nas três gerações tendo como resultado 0.818 para função objetivo (avaliação), 99,17% de acurácia e 0,17391 de peso, o GA indica como resultado o primeiro resultado encontrado no caso o da 12 geração.

Ainda observando a tabela 1 podemos verificar que a posição das barras selecionadas tem influencia no resultado, como exemplo a 5ª geração apesar de ter 4 como quantidade de barras selecionadas o valor da acurácia é de 89,55% bem abaixo do resultado indicado pelo GA no valor de 99,17% que também tem 4 barras selecionadas.

O resultado da otimização foi comparado com uma simulação do algoritmo não otimizado, na tabela 2 é possível visualizar o melhor desempenho do método não otimizado (K-means with k=3 and PART) com 99,42% de acurácia porem são utilizados

**Table 1. Melhores resultados por geração**

Geração	Barras	Retorno OF	Acurácia (%)	Peso
1	8	0,600	94,77%	0,34783
2	8	0,635	98,26%	0,34783
3	6	0,686	94,70%	0,26087
4	6	0,677	93,79%	0,26087
5	4	0,722	89,55%	0,17391
6	4	0,765	93,94%	0,17391
7	4	0,757	93,11%	0,17391
8	4	0,783	95,68%	0,17391
9	4	0,809	98,26%	0,17391
10	4	0,809	98,33%	0,17391
11	4	0,781	95,45%	0,17391
12	4	0,818	99,17%	0,17391
13	4	0,800	97,42%	0,17391
14	4	0,809	98,26%	0,17391
15	4	0,809	98,26%	0,17391
16	4	0,800	97,35%	0,17391
17	4	0,800	97,35%	0,17391
18	4	0,800	97,42%	0,17391
19	4	0,792	96,59%	0,17391
20	4	0,799	97,27%	0,17391
21	4	0,809	98,26%	0,17391
22	4	0,800	97,42%	0,17391
23	4	0,800	97,42%	0,17391
24	4	0,800	97,42%	0,17391
25	4	0,793	96,67%	0,17391
26	4	0,799	97,27%	0,17391
27	4	0,818	99,17%	0,17391
28	4	0,775	94,92%	0,17391
29	4	0,791	96,52%	0,17391
30	4	0,801	97,50%	0,17391
31	4	0,800	97,42%	0,17391
32	4	0,800	97,42%	0,17391
33	4	0,801	97,50%	0,17391
34	4	0,817	99,09%	0,17391
35	4	0,800	97,35%	0,17391
36	4	0,800	97,35%	0,17391
37	4	0,792	96,59%	0,17391
38	4	0,790	96,44%	0,17391
39	4	0,783	95,68%	0,17391
40	4	0,809	98,33%	0,17391
41	4	0,790	96,44%	0,17391
42	4	0,800	97,42%	0,17391
43	4	0,817	99,09%	0,17391
44	4	0,791	96,52%	0,17391
45	4	0,818	99,17%	0,17391
46	4	0,817	99,09%	0,17391
47	4	0,790	96,44%	0,17391
48	4	0,809	98,33%	0,17391
49	4	0,801	97,50%	0,17391
50	4	0,801	97,50%	0,17391

os dados das 23 barras trifásicas da rede. Já a acurácia da metodologia otimizada tem uma queda muito pequena de apenas 0,26%, totalizando 99,16% de acuratividade, porem com a utilização de dados de apenas 4 barras tendo um ganho em não precisar enviar e analisar os dados de 19 barras do sistema.

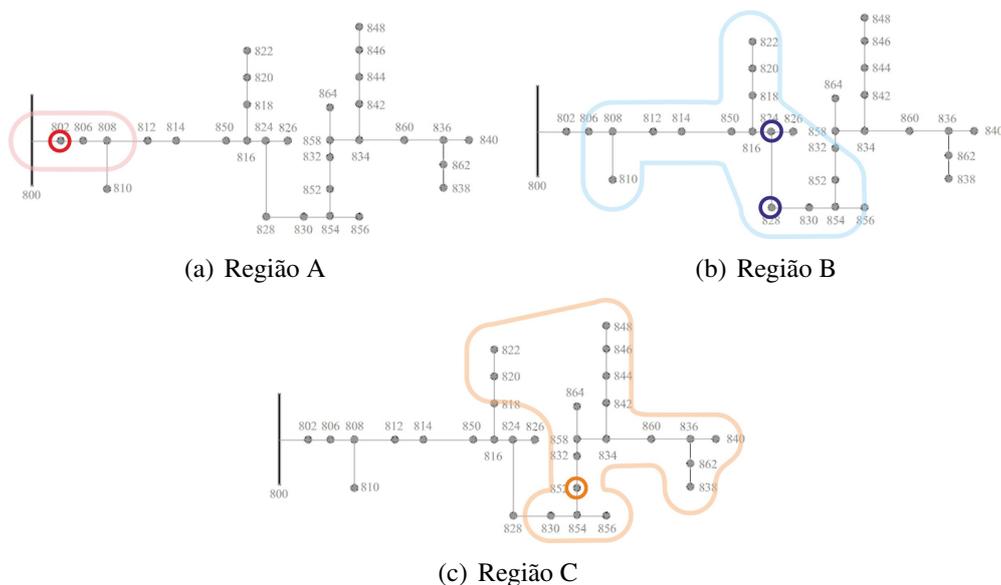
**Table 2. Acurácia do método não otimizado e otimizado**

Algoritmo	Acurácia (%)	Barras
K-means with k=3 and PART	99,42	23
Genetic Algorithm	99,16	4

Como a falta simulada é uma falha no sistema de curta duração a velocidade da interpretação dos dados e um fator importante a se considerar, o processo de otimização tem uma pequena perda de acuratividade contudo um volume de dados a ser analisado é bem inferior, o que credencia a uma maior velocidade na análise da informação. Na próxima sessão é analisado e validado os grupos que formam cada região.

### 3.2. Validação dos Grupos

Como uma falta pode ocorrer em qualquer ponto e os efeitos da falha se propagam alterando o comportamento da rede é preciso validar a frequência em que os pontos são relacionados em uma mesma região. A Figura 10 e a Tabela 3 representa as 115 simulações realizadas com a configuração indicada pelo algoritmo genético com 4 pontos chaves indicados na tabela pela coluna "Chave" a tabela também mostra a quantidade de vezes em que cada ponto da rede é alocado em um determinado grupo, ou seja, em uma região.



**Figure 10. Divisão dos grupos com a configuração indicada pelo algoritmo genético**

Na Figura 10(a) é demonstrado um grupo pequeno (com poucos pontos) considerando a subestação como um ponto(800) essa região contempla 4 pontos 800, 802,

806 e 808 que na simulação realizada é quantidade mínima necessária para execução do algoritmo genético em seu processo de treinamento. Devido a isso a divisão de grupos para o sistema IEEE de 34 pontos é restrita a 3, evitando a criação de vários grupos pequenos que pode comprometer a eficácia do método.

**Table 3. Quantidade de ocorrências dos ponto nas divisões dos grupos**

Barra	Chave	A	B	C	Total
800		115			115
802	A	115			115
806		115			115
808		110	5		115
812			115		115
814			115		115
816			115		115
818			115		115
820			107	8	115
822			95	20	115
824	B		115		115
828	B		115		115
830			113	2	115
832				115	115
834				115	115
836				115	115
840				115	115
842				115	115
844				115	115
846				115	115
848				115	115
850			115		115
852	C		10	105	115
854			113	2	115
858				115	115
860				115	115
862				115	115
864				115	115
888				115	115
890			21	94	115

É possível visualizar na Tabela 3 que alguns pontos nas 115 simulações realizadas podem aparecer em um grupo diferente, contudo é visível a coerência dos grupos pela quantidades de vezes que um ponto está no mesmo grupo. Para os pontos que aparecem em dois grupos diferente, mesmo que a quantidade de vezes em que esse ponto se apresenta em um determinado grupo seja pequena como exemplo o ponto 830 com 2 simulações no grupo C e as demais 113 simulações no grupo B esse ponto será considerado nos dois grupos.

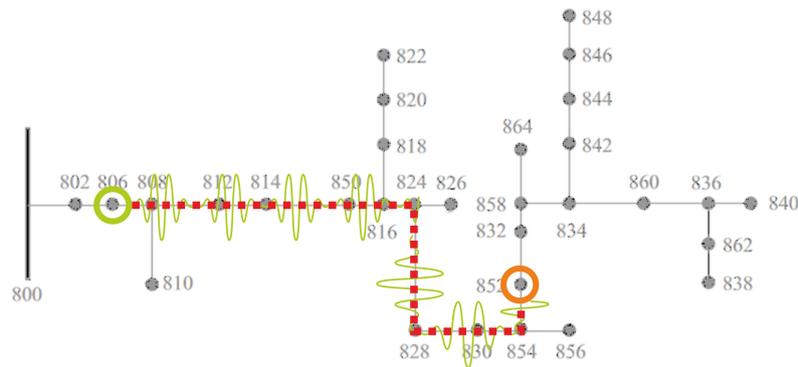
Outra situação que pode ser visualizada na Tabela 3 é que o ponto chave da região C o ponto 852 em 10 simulações faz parte do grupo B, como visto anteriormente um ponto pode está em dois grupos diferentes e como na ultima etapa não se faz distinção entre pontos chaves e pontos secundários o ponto 852 será analisado se o grupo B for isolado ou se o grupo isolado for o grupo C, para um melhor entendimento a Tabela 4 mostra os pontos em que foi gerado a falha para as 10 simulações que incluem o ponto 852 no grupo B e as demais 105 simulações que incluem o ponto 852 no grupo C.

**Table 4. Origem da falha para alocação do ponto 852 nas regiões B e C**

Ocorrencia Falta	B	C
802	5	0
806	5	0
808	0	5
812	0	5
814	0	5
816	0	5
824	0	5
828	0	5
830	0	5
832	0	5
834	0	5
836	0	5
840	0	5
842	0	5
844	0	5
846	0	5
848	0	5
850	0	5
852	0	5
854	0	5
858	0	5
860	0	5
862	0	5
<b>Total</b>	<b>10</b>	<b>105</b>

As falhas que incluem no grupo B foram geradas nos pontos 802 e 806 com a quantidade de 5 simulações cada, ao visualizar a figura 11 é perceptível a distancia entre o ponto 852 para os pontos 802 e o 806, são os pontos mais distantes em relação a subestação e com a propagação do evento as características registradas no ponto 852 o inclui no grupo B. A partir do ponto 808 apesar da distancia em relação ao ponto 852 não ser tão diferente aos pontos anteriores os valores alterados pela propagação incluem o ponto 852 no grupo C.

Visto isso, vale ressaltar que a inclusão de um ponto em dois grupos diferentes não deverá afetar a precisão da classificação realizada pela CNN.



**Figure 11. Propagação da falta no pontos 806 até o ponto 852 no Sistema IEEE de 34 barras**

## 4. Pesquisa Futura

Neste trabalho foi implementado e analisado até a 3ª etapa: identificação dos pontos chaves, divisão dos pontos em grupos gerando regiões e localização da região em que ocorreu a falha. Para continuidade do trabalho será realizado os seguintes pontos:

- Implementar CNN e aplicar ao método.
- Realizar comparação com CNN que usam todo o sistema e CNN com grafo.
- Implementar e gerar simulações com um sistema IEEE de 123 Barras.

### 4.1. Rede Neural Convolutacional

A escolha da Rede Neural Convolutacional para implementação da quarta etapa se deve pelas seguintes razões: A operação de convolução na CNN gera conectividade esparsa entre as entradas de diferentes camadas, que é mais adequada do que a operação totalmente conectada para o vetor esparsa. A CNN revela as características locais. O pequeno tamanho dos *kernels* fortalece a CNN para aprender os recursos locais em vez dos recursos globais. A CNN tem a vantagem de aprender a propriedade da vizinhança a partir da entrada [Li et al. 2019].

### 4.2. Comparações e novas simulações

Depois de implementado CNN será realizada comparações com outros metodo de localização da falta no sistema elétrico para avaliar o desempenho do método em relação a diminuição do grande volume de informação, aprimorando a velocidade do processamento, mantendo a qualidade do valor da informação que trafega na rede e se adaptando a variedade dos dados. A implementação de novas simulação com o sistema IEEE 123 node test feeder será importante para avaliação da escalabilidade do método..

## 5. Conclusão

Este trabalho tem como objetivo desenvolver um método para localização de faltas em redes inteligentes de distribuição de energia elétrica de forma rápida e precisa, mas diferente dos outros métodos, também se preocupa com um recurso muito utilizado nos métodos atuais a Big Data, problemas típicos característicos de Big Data como grande volume, variedade, valor e velocidade. Para mitigar esses problemas o método utiliza 4 etapas:

Identificar pontos chaves; Associar pontos secundários aos pontos chave; Localizar região da falta pelos pontos chaves e por fim identifica do ponto que ocorreu a falta.

Nas simulações realizadas o sistema é dividido em 3 regiões, onde é possível identificar qual região se encontra a falta com apenas 4 pontos chaves e uma acurácia de 99,13%, em seguida a região é isolada e será aplicada uma CNN com apenas os pontos pertencentes a região, com isso é possível atenuar os problemas de volume e velocidade. Também foi verificado a robustez da divisão dos grupo garantindo o valor da informação.

Como continuidade do trabalho será implementado uma rede neural convolucional, sua característica de ter como entrada um vetor esparso contribui com a aplicação na metodologia proposta por utilizar apenas os pontos isolados da região.

## References

- Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6. Ieee.
- Arsoniadis, C. G., Apostolopoulos, C. A., Georgilakis, P. S., and Nikolaidis, V. C. (2021). A voltage-based fault location algorithm for medium voltage active distribution systems. *Electric Power Systems Research*, 196:107236.
- Blank, J. and Deb, K. (2020). pymoo: Multi-objective optimization in python. *IEEE Access*, 8:89497–89509.
- Chang, G., Chao, J., Chu, S., and Chen, C. (2007). A new procedure for tracking the source location. In *2007 IEEE Power Engineering Society General Meeting*, pages 1–4. IEEE.
- Chen, K., Hu, J., Zhang, Y., Yu, Z., and He, J. (2019). Fault location in power distribution systems via deep graph convolutional networks. *IEEE Journal on Selected Areas in Communications*, 38(1):119–131.
- Dashti, R., Daisy, M., Mirshekali, H., Shaker, H. R., and Aliabadi, M. H. (2021). A survey of fault prediction and location methods in electrical energy distribution networks. *Measurement*, 184:109947.
- Feeders, R. D. T. (1991). Ieee distribution planning working group report. *IEEE Trans. Power Syst*, 6(3):975–985.
- Galijasevic, Z. and Abur, A. (2002). Fault location using voltage measurements. *IEEE Transactions on Power Delivery*, 17(2):441–445. DOI: 10.1109/61.997915.
- Ghorbanian, M., Dolatabadi, S. H., and Siano, P. (2019). Big data issues in smart grids: A survey. *IEEE Systems Journal*, 13(4):4158–4168.
- Hossain, E., Khan, I., Un-Noor, F., Sikander, S. S., and Sunny, M. S. H. (2019). Application of big data and machine learning in smart grid, and associated security concerns: A review. *Ieee Access*, 7:13960–13988.
- Huchche, V., Patne, N., and Junghare, A. (2017). Computation of energy loss in an induction motor during unsymmetrical voltage sags—a graphical method. *IEEE Transactions on Industrial Informatics*, 14(5):2023–2030.

- Khan, M. F., Azam, M., Khan, M. A., Algarni, F., Ashfaq, M., Ahmad, I., and Ullah, I. (2021). A review of big data resource management: Using smart grid systems as a case study. *Wireless Communications and Mobile Computing*, 2021.
- Li, W., Deka, D., Chertkov, M., and Wang, M. (2019). Real-time faulted line localization and pmu placement in power systems through convolutional neural networks. *IEEE Transactions on Power Systems*, 34(6):4640–4651.
- Li, Y., Hao, Z., and Lei, H. (2016). Survey of convolutional neural network. *Journal of Computer Applications*, 36(9):2508–2515.
- Lima Filho, J., da Silva Borges, F. A., Rabelo, R. d. A. L., Silva, I. S., Junior, R. P. T., and de Carvalho Filho, A. O. (2020). Methods for voltage sag source location by cluster algorithm and decision rule labeling with a comparative approach of k-means and dbscan clustering algorithms. In *2020 5th International Conference on Smart and Sustainable Technologies (SpliTech)*, pages 1–8. IEEE.
- Ly, F., Wen, C., Bao, Z., and Liu, M. (2016). Fault diagnosis based on deep learning. In *2016 American control conference (ACC)*, pages 6851–6856. IEEE.
- L’heureux, A., Grolinger, K., Elyamany, H. F., and Capretz, M. A. (2017). Machine learning with big data: Challenges and approaches. *Ieee Access*, 5:7776–7797.
- Miranda Filho, J., de Carvalho Filho, J. M., Paiva, A. P., de Souza, P. V. G., and Tomasin, S. (2016). A pca-based approach for substation clustering for voltage sag studies in the brazilian new energy context. *Electric Power Systems Research*, 136:31–42.
- Mirjalili, S. (2019). Genetic algorithm. In *Evolutionary algorithms and neural networks*, pages 43–55. Springer.
- Mohammadi, Y. and Leborgne, R. C. (2020). A new approach for voltage sag source relative location in active distribution systems with the presence of inverter-based distributed generations. *Electric Power Systems Research*, 182:106222.
- Mohammadi, Y., Moradi, M. H., and Leborgne, R. C. (2017). Employing instantaneous positive sequence symmetrical components for voltage sag source relative location. *Electric Power Systems Research*, 151:186–196.
- Ozcanli, A. K., Yaprakdal, F., and Baysal, M. (2020). Deep learning methods and applications for electrical power systems: A comprehensive review. *International Journal of Energy Research*, 44(9):7136–7157.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830.
- Personal, E., García, A., Parejo, A., Larios, D. F., Biscarri, F., and León, C. (2016). A comparison of impedance-based fault location methods for power underground distribution systems. *Energies*, 9(12):1022.
- Tikariha, A., Londhe, N. D., Bag, B., and Raj, R. (2021). Classification of faults in an ieee 30 bus transmission system using fully convolutional network. *International Transactions on Electrical Energy Systems*, 31(11):e13134.

- Wang, S., Sun, X., Geng, J., Han, Y., Zhang, C., and Zhang, W. (2020). Application and analysis of big data technology in smart grid. In *Journal of Physics: Conference Series*, volume 1639, page 012043. IOP Publishing.
- Wei-jing, Z., He-chun, Z., Shi-ying, Y., and Tong, L. (2019). A homomorphic encryption-based privacy preserving data aggregation scheme for smart grid. In *2019 15th International Conference on Computational Intelligence and Security (CIS)*, pages 315–319. IEEE.
- Wen, L., Zhou, K., Yang, S., and Li, L. (2018). Compression of smart meter big data: A survey. *Renewable and Sustainable Energy Reviews*, 91:59–69.
- Wu, J.-Y., Lan, S., Xiao, S.-J., and Yuan, Y.-B. (2021). Single pole-to-ground fault location system for mmc-hvdc transmission lines based on active pulse and ceemdan. *IEEE Access*, 9:42226–42235.
- Zhang, J., Gong, Q., Zhang, H., Wang, Y., and Wang, Y. (2021). A novel pix2pix enabled traveling wave-based fault location method. *Sensors*, 21(5):1633.
- Zhao, H. and Li, N. (2016). Performance evaluation for sustainability of strong smart grid by using stochastic ahp and fuzzy topsis methods. *Sustainability*, 8(2):129.
- Zheng, Z., Qi, L., Wang, H., Zhu, M., and Chen, Q. (2020). Recognition method of voltage sag causes based on bi-lstm. *IEEE Transactions on Electrical and Electronic Engineering*, 15(3):418–425.