



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS  
DEPARTAMENTO DE ELETRÔNICA E SISTEMAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

ANTONY JÚNIO BARBOSA DE SOUZA

**PROCESSAMENTO DE SINAIS SOBRE GRAFOS COM UMA APLICAÇÃO  
À DETECÇÃO DE SUPERAQUECIMENTO EM QUADROS DE ENERGIA ELÉTRICA**

Recife

2022

ANTONY JÚNIO BARBOSA DE SOUZA

**PROCESSAMENTO DE SINAIS SOBRE GRAFOS COM UMA APLICAÇÃO  
À DETECÇÃO DE SUPERAQUECIMENTO EM QUADROS DE ENERGIA ELÉTRICA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de mestre em Engenharia Elétrica.

Área de concentração: Comunicações.

Orientador: Prof. Dr. Juliano Bandeira Lima

Coorientador: Prof. Dr. José Rodrigues de Oliveira Neto

Recife

2022

Catálogo na fonte:  
Bibliotecária Sandra Maria Neri Santiago, CRB-4 / 1267

S729p

Souza, Antony Júnio Barbosa de.

Processamento de sinais sobre grafos com uma aplicação à detecção de superaquecimento em quadros de energia elétrica / Antony Júnio Barbosa de Souza. – 2022.

60 f.: il., fig.

Orientador: Prof. Dr. Juliano Bandeira Lima.

Coorientador: Prof. Dr. José Rodrigues de Oliveira Neto.

Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Elétrica. Recife, 2022.

Inclui referências e apêndice.

1. Engenharia elétrica. 2. Processamento de sinais. 3. Grafos. 4. Redes de sensores. 5. Detecção de anomalias. 6. Instalações elétricas. I. Lima, Juliano Bandeira (Orientador). II. Oliveira Neto, José Rodrigues de (Coorientador). III. Título.

UFPE

621.3 CDD (22. ed.)

BCTG/2021-36

ANTONY JÚNIO BARBOSA DE SOUZA

**PROCESSAMENTO DE SINAIS SOBRE GRAFOS COM UMA APLICAÇÃO  
À DETECÇÃO DE SUPERAQUECIMENTO EM QUADROS DE ENERGIA ELÉTRICA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para obtenção do título de mestre em Engenharia Elétrica. Área de concentração: Comunicações.

Aprovada em:  22  /  12  /  2022  .

**BANCA EXAMINADORA**

---

Prof. Dr. Juliano Bandeira Lima (Orientador)  
Universidade Federal de Pernambuco

---

Prof. Dr. José Rodrigues de Oliveira Neto (Examinador Interno)  
Universidade Federal de Pernambuco

---

Prof. Dr. Fernando José Ribeiro Sales (Examinador Externo)  
Universidade Federal de Pernambuco

---

Prof. Dr. Felipe Alberto Barbosa Simão Ferreira (Examinador Externo)  
Universidade Federal Rural de Pernambuco

Dedico este trabalho a Deus, que sempre me colocou nos melhores caminhos e esteve comigo nos momentos mais difíceis.

## **AGRADECIMENTOS**

Gostaria de agradecer a algumas pessoas que me acompanharam nesses últimos anos de mestrado e que sem elas certamente este trabalho não teria se concretizado. Em primeiro lugar, agradeço aos meus pais, o sr. Antônio e a sra. Cristina, que sempre foram os meus principais alicerces e sempre me incentivaram a seguir adiante com a minha educação, sem medir esforços para isso. Também não posso deixar de agradecer ao meu orientador, Professor Doutor Juliano Bandeira Lima, e ao meu coorientador, Professor Doutor José Rodrigues de Oliveira Neto, por toda paciência, empenho e toda instrução excelente que recebi. Gostaria de agradecer também a todos os meus colegas de mestrado, que estiveram comigo nessa caminhada e sempre nos ajudamos mutuamente para que pudéssemos seguir adiante. Por último, quero agradecer também aos meus amigos e colegas de trabalho pelo incentivo a iniciar essa caminhada.

## RESUMO

Nas últimas décadas, diversos avanços tecnológicos têm permitido a produção e o armazenamento massivo de dados relacionados a uma variedade de cenários do mundo real. Muitos desses cenários podem ser representados como redes sobre as quais os respectivos dados são distribuídos. É o caso, por exemplo, de uma rede de sensores espacialmente dispostos dedicada à realização de medições que se inter-relacionam de alguma forma, compondo o que seria um sinal sobre a rede em questão. Em geral, cenários como o exemplificado podem ser modelados como um grafo cujos nós estão associados às amostras de um sinal. Para lidar com sinais dessa natureza, foi proposto o processamento de sinais sobre grafos (GSP), que, em linhas gerais, busca estender para o chamado domínio dos vértices conceitos e operações do processamento de sinais clássico, dedicado a analisar sinais em domínios como o tempo discreto. Nesta dissertação, é apresentado um estudo revisivo sobre os fundamentos do processamento de sinais sobre grafos, com ênfase em aplicações da respectiva teoria em problemas relacionados a redes de sensores. Como contribuição original deste trabalho, é realizado um estudo de caso que consiste em aplicar o GSP à detecção de superaquecimento em quadros de energia elétrica. O referido estudo considera a rede elétrica de um grande hospital da região metropolitana do Recife, modelando-a como um grafo cujos nós correspondem aos quadros de energia. Feitas as devidas ressalvas, relativas principalmente à quantidade limitada de dados disponíveis para realização do estudo, os resultados obtidos sugerem que o GSP pode ser uma ferramenta útil para a aplicação em questão, provendo indícios satisfatórios sobre o aparecimento de pontos quentes na rede analisada.

Palavras-chave: processamento de sinais; grafos; redes de sensores; detecção de anomalias; instalações elétricas.

## ABSTRACT

In recent decades, several technological advances have allowed the massive production and storage of data related to a variety of real-world scenarios. Many of these scenarios can be represented as networks over which the respective data is distributed. This is the case, for example, of a network of spatially arranged sensors dedicated to performing measurements that are interrelated in some sense, composing what would be a signal over the network in question. In general, scenarios like the one exemplified can be modeled as a graph whose nodes are associated with samples of a signal. To deal with signals of this nature, graph signal processing (GSP) was proposed, which seeks to extend to the so-called vertex domain concepts and operations of classical signal processing, dedicated to analyzing signals in domains such as discrete-time only. In this dissertation, a review on the fundamentals of graph signal processing is presented, with emphasis on applications of the respective theory in problems related to sensor networks. As an original contribution of this work, a case study is carried out, which consists of applying GSP to detect overheating in electrical energy switchboards. This study considers the electrical network of a large hospital in the metropolitan region of Recife, modeling it as a graph whose nodes correspond to the energy switchboards. Despite limitations mainly related to the restricted amount of data available for the study, the results obtained suggest that the GSP can be a useful tool for the application in question, providing satisfactory evidence about the appearance of hot spots in the analyzed network.

Keywords: signal processing; graphs; sensor networks; anomaly detection; electrical installations.

## LISTA DE FIGURAS

Figura 1 – Representação de domínios em que se encontram sinais discretos (a) no tempo e (b) no espaço. . . . .	10
Figura 2 – Representação de domínios em que se encontram sinais discretos no tempo e no espaço. . . . .	11
Figura 3 – Exemplo de matriz de adjacência e o grafo . . . . .	17
Figura 4 – Exemplo da representação gráfica de grafos direcionados e não direcionados. . . . .	18
Figura 5 – Exemplos de Grafos . . . . .	19
Figura 6 – Grafo rede de sensores . . . . .	20
Figura 7 – Comparação da transformada de Fourier sobre o grafo de dois sinais com ordenamentos diferentes . . . . .	21
Figura 8 – Magnitude dos Coeficientes Espectrais. . . . .	30
Figura 9 – Filtro de dois canais . . . . .	31
Figura 10 – Estimativa de Exposição a Poluição de edifícios na cidade de Nova York. . . . .	35
Figura 11 – Grafo e Matriz de Adjacência . . . . .	38
Figura 12 – Temperatura média ao longo do ano. . . . .	39
Figura 13 – Grafo com sinal ideal. . . . .	41
Figura 14 – Alguns sinais sobre grafos e respectivas GFTs. . . . .	43
Figura 15 – Percentual de acerto de nós anômalos e falsos positivos $\times$ variação do valor de gama. . . . .	44
Figura 16 – Percentual de acerto médio por sinal. . . . .	44

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> . . . . .	<b>10</b>
1.1	OBJETIVOS . . . . .	13
1.2	MATERIAIS E MÉTODOS . . . . .	14
1.3	CONTRIBUIÇÕES . . . . .	15
1.4	ESTRUTURA DA DISSERTAÇÃO . . . . .	15
<b>2</b>	<b>PROCESSAMENTO DE SINAIS SOBRE GRAFOS</b> . . . . .	<b>17</b>
2.1	CONCEITOS BÁSICOS . . . . .	17
2.2	SINAIS SOBRE GRAFOS . . . . .	20
<b>2.2.1</b>	<b>Matrizes de Grau, Adjacência e Laplaciana</b> . . . . .	<b>21</b>
<b>2.2.2</b>	<b>Deslocamento de Sinais sobre Grafos</b> . . . . .	<b>23</b>
2.3	TRANSFORMADA DE FOURIER SOBRE GRAFOS . . . . .	24
2.4	FILTRAGEM . . . . .	26
2.5	OUTRAS OPERAÇÕES REALIZADAS SOBRE GRAFOS . . . . .	27
<b>2.5.1</b>	<b>Deslocamentos no domínio da frequência</b> . . . . .	<b>27</b>
<b>2.5.2</b>	<b>Convolução de Sinais sobre Grafos</b> . . . . .	<b>28</b>
<b>3</b>	<b>APLICAÇÕES DE GSP EM REDES DE SENSORES</b> . . . . .	<b>29</b>
3.1	DETECÇÃO DE MAL FUNCIONAMENTO EM REDES DE SENSORES .	29
3.2	DETECÇÃO E LOCALIZAÇÃO DE ANOMALIAS EM REDES DE SENSORES . . . . .	30
3.3	MODELAGEM DE POLUIÇÃO DO AR EM AMBIENTES URBANOS . .	33
<b>4</b>	<b>ESTUDO DE CASO</b> . . . . .	<b>36</b>
4.1	CARACTERÍSTICAS DO SINAL E MONTAGEM DO GRAFO . . . . .	36
4.2	ANÁLISE DA GFT . . . . .	41
4.3	DETECÇÃO E LOCALIZAÇÃO DE ANOMALIAS . . . . .	42
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b> . . . . .	<b>46</b>
5.1	TRABALHOS FUTUROS . . . . .	47
	<b>REFERÊNCIAS</b> . . . . .	<b>49</b>
	<b>APÊNDICE A – CÓDIGOS UTILIZADOS PELO AUTOR EM PYTHON</b>	<b>53</b>

## 1 INTRODUÇÃO

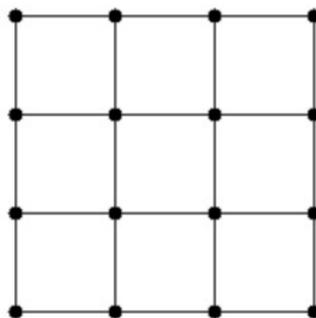
Nos últimos anos, o processamento de sinais sobre grafos (GSP, do inglês *graph signal processing*) vem se tornando uma área de pesquisa de crescente interesse da Academia e da Indústria (ORTEGA et al., 2018). Nesse contexto, têm havido desenvolvimentos teóricos relevantes e têm sido propostas aplicações em diferentes cenários práticos (JABLONSKI, 2017; SANDRYHAILA; MOURA, 2014). Um estímulo particular à evolução do GSP e à sua consolidação enquanto ferramenta útil para processamento e análise de sinais é a crescente naturalidade com que dados dos mais diferentes tipos e origens têm sido armazenados e disponibilizados para uso e extração de informações. Esses dados são oriundos, por exemplo, do acesso a redes sociais (ZHANG; MOURA, 2014; PEROZZI; AL-RFOU; SKIENA, 2014), da verificação de uma previsão de tempo, da realização de um exame médico (HUANG et al., 2018; ANIRUDH; THIAGARAJAN, 2017; MASOUMI; REZAEI; HAMZA, 2019) etc. Do ponto de vista de representação, alguns desses dados não se adequam a modelos corriqueiros de sinais que se associam a variações de um valor ao longo de um eixo temporal (unidimensional) ou a mudanças de intensidade no espaço bidimensional (caso de uma imagem).

Se os domínios citados forem do tipo discreto, as amostras dos respectivos sinais são dispostas de forma equispaçada, conforme ilustrado na Figura 1, o que dá suporte a noções com significado físico bastante intuitivo e casadas com ferramentas matemáticas bem estabelecidas. Tais noções incluem, por exemplo, a de amostra passada ou futura (no caso de sinal temporal), vizinhança vertical, horizontal ou diagonal (no caso de sinal no espaço), variação menos ou mais abrupta da intensidade do sinal, o que, por sua vez, associa-se à ideia de frequência baixa ou frequência alta, respectivamente, e assim por diante.

Figura 1 – Representação de domínios em que se encontram sinais discretos no tempo e no espaço.



(a)

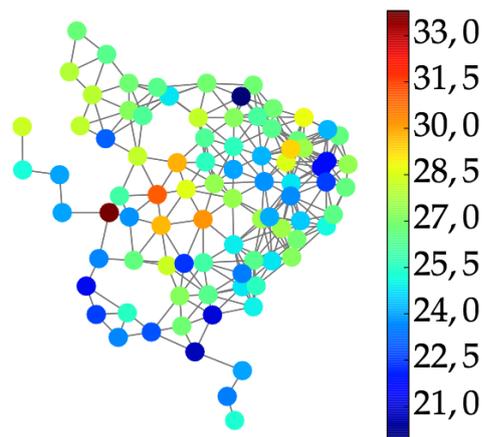


(b)

Fonte: O Autor (2022).

Extrapolando as situações que se acabou de descrever, o paradigma do GSP contempla sinais que se encontram sobre domínios modeláveis como um grafo arbitrário, isto é, um conjunto de vértices conectados por arestas (SHUMAN et al., 2013). Em contraposição ao domínio temporal e ao espacial, isso permite que se leve em conta irregularidades (ou não-uniformidades) do ponto de vista de disposição das amostras e da forma como essas amostras se relacionam (STANKOVIC; DAKOVIC; SEJDIC, 2019). Um exemplo sugestivo dessa possibilidade é ilustrado na Figura 2, em que é ilustrado um grafo cujos vértices representam cidades do Nordeste brasileiro. Nesse grafo, a conexão (ou não) de um par de cidades por uma aresta pode ser determinada por algum critério de distância. A cada uma dessas cidades está associada uma amostra que representa o valor da temperatura medida naquele local num dia específico. A cor com que cada amostra é pintada é função da escala apresentada no lado direito da imagem.

Figura 2 – Um grafo das cidades do Nordeste do Brasil, sobre o qual é definido um sinal de temperatura em 1º de fevereiro de 2012, obtido do Banco de Dados Meteorológicos para Ensino e Pesquisa, disponível em: <http://www.inmet.gov.br/portal/index.php?r=bdmep/bdmep>.



Fonte: Ribeiro; Lima (2018)

Um modelo (ou representação) como o da última figura pode ser empregado em vários outros tipos de redes de sensores e também em contextos mais diversos (ORTEGA et al., 2018). Por exemplo, linhas de transmissão de energia elétrica podem ser modeladas considerando as subestações como vértices (ou nós) e as interconexões entre estas subestações como as arestas de um grafo (RAMAKRISHNA; SCAGLIONE, 2021). No caso de redes sociais, os vértices representam os membros daquela rede e as arestas as relações de amizade (que podem, inclusive, indicar reciprocidade ou não) (AKELLA et al., 2009). No contexto de redes biológicas, grafos podem ser utilizados para modelar as interdependências (arestas) físicas ou funcionais das diversas regiões (vértices) do cérebro; o sinal estaria relacionado a estímulos aplicados sobre essas regiões (GOLDSBERRY et al., 2017; HUANG et al., 2016; RABIEI et al., 2019) ou ainda utilizados para mostrar a interdependência entre genes (PIRAYRE et al., 2015). Redes de transporte podem ser modeladas por grafos em que os cruzamentos são os vértices, as vias

são as arestas e fenômenos como a retenção de veículos ou os números de subidas e descidas em táxis (nos cruzamentos) correspondem ao sinal (MOHAN et al., 2014). Também podem ser caracterizados como sinais sobre grafos números de citações (ou outro índice de produtividade) em redes de colaboração científica, números relacionados ao espalhamento de uma doença em cidades cujas populações se conectam de alguma forma (HOSSEINALIPOUR et al., 2020), atributos de pontos pertencentes a uma nuvem tridimensional de pontos (*3D point cloud*) (CHEN, 2016), entre outros.

Modelos como os exemplificados acima, que consideram que o sinal se encontra no chamado domínio dos vértices, podem ser hibridizados, de maneira que se considere também uma variação temporal dos valores das amostras sobre cada um dos respectivos nós; semelhante a uma estrutura bidimensional: uma dimensão é a dos vértices e a outra é a do tempo. Nos últimos anos, essa possibilidade tem sido levada em conta por diversos autores e encontrado aplicabilidade em diversos cenários práticos (ORTEGA et al., 2018). O fato é que, à medida em que as tecnologias avançam e se tornam mais acessíveis, a demanda pelo processamento de grandes volumes de dados aumenta e a necessidade de desenvolver novas ferramentas teóricas nesse contexto se mostra de modo mais clara.

Numa perspectiva ainda mais específica da utilidade do processamento de sinais sobre grafos, alguns trabalhos relativamente recentes podem ser mencionados. Em (JAIN; MOURA; KONTOKOSTA, 2014), por exemplo, os autores adotam, para uma região da cidade de Nova York, um modelo que considera edifícios como vértices e suas emissões de poluentes no ar como sinal; a sua ideia é verificar os locais mais suscetíveis a maiores concentrações de poluentes na região considerada. A metodologia e os resultados obtidos sugerem que é possível prever locais onde as concentrações de poluentes no ar seriam maiores. Com o modelo utilizado, foi possível verificar que nem sempre locais com maiores emissões de poluentes estão associados aos locais com a maior concentração de poluentes. Este tipo de processamento pode ser utilizado para orientar e otimizar a instalação de sensores de qualidade e monitoramento do ar, por exemplo, reduzindo a quantidade de sensores que precisam ser instalados. Outra aplicação relacionada ao posicionamento de sensores utiliza a teoria de amostragem de sinais sobre grafos para maximizar a frequência de corte do sinal e selecionar os nós com a maior quantidade de informação disponível (SAKIYAMA et al., 2016).

O GSP também já foi utilizado para a detectar anomalias em redes de sensores (SANDRYHAILA; MOURA, 2014; EGILMEZ; ORTEGA, 2014), estimar a temperatura em algumas estações climáticas (SPELTA; MARTINS, 2018) e para detectar padrões não triviais em redes de sensoriamento complexas em (JABLONSKI, 2017). Outras aplicações em GSP têm como objetivo otimizar a velocidade de processamento de dados. Em (BLONDEL et al., 2008), por exemplo, é feita a aplicação do algoritmo de Louvain, utilizado para calcular a modularidade de uma rede, para a detecção rápida de comunidades em grandes redes. Este tipo de processamento visa ao agrupamento de dados semelhantes e de grandes comunidades antes de realizar outros

tipos de análise, otimizando o tempo utilizado para o seu processamento. O GSP também encontra aplicações em técnicas de *machine learning* e *deep learning* (DONG et al., 2020; CHEUNG et al., 2020; GRIPON; ORTEGA; GIRAULT, 2018).

Com o avanço das tecnologias e o aumento da demanda por cidades cada vez mais conectadas, é importante que novas técnicas de processamento para os mais variados tipos de dados que estão sendo gerados possam ser desenvolvidas. Nesse sentido, buscar novas formas de aplicação de técnicas já consolidadas na literatura também se torna válido.

No contexto de quadros de distribuição de energia elétrica, a inspeção e a manutenção periódica destes quadros são importantes para garantir que o fornecimento de energia elétrica vai ocorrer de forma adequada. Falhas nestes equipamentos quase sempre resultam em parada total ou parcial do fornecimento de energia, causando diversos transtornos para os usuários. Uma das principais técnicas utilizadas para se inspecionar quadros elétricos é a utilização de câmeras termográficas, que são capazes de mapear a temperatura nas diferentes regiões do quadro elétrico. A partir deste mapeamento é possível definir se o quadro está funcionando dentro das suas condições normais de operação ou se há a formação de pontos quentes neste quadro.

Em grandes edifícios a distribuição de energia é feita através de diversos quadros elétricos. Alguns chegam a ter centenas de quadros de distribuição. Inspecionar e analisar as imagens termográficas feitas em todos estes quadros demanda bastante tempo. Encontrar uma forma de processar as leituras de temperatura destes quadros de modo que seja possível prever quais quadros estão com problemas de super-aquecimento pode reduzir bastante o tempo levado para realizar estas inspeções e processar as informações levantadas. Naturalmente, alguns destes quadros têm algum grau de relação uns com os outros. Eles podem compartilhar um mesmo ponto de alimentação ou distribuir energia para um mesmo ambiente. Desta forma é possível montar um grafo com as informações destes quadros e, a partir do sinal de temperatura aplicado sobre este grafo, aplicar técnicas de detecção de anomalias que possam detectar o superaquecimento em algum quadro.

A variedade e a representatividade de áreas que podem ser contempladas com avanços teóricos e práticos do GSP sugerem que contribuições a essa linha de pesquisa podem ser relevantes, o que justifica o desenvolvimento de uma dissertação nessa direção. Nesse escopo, encontrou-se motivação para a realização do presente trabalho, cujos objetivos são delineados a seguir.

## 1.1 OBJETIVOS

Esta dissertação tem como objetivo geral a realização de um estudo sobre os fundamentos do processamento de sinais sobre grafos, visando à sua aplicabilidade à detecção de variações de temperatura em quadros de energia elétrica das instalações de um hospital na cidade do Recife.

### Objetivos específicos:

Para alcançar o objetivo geral deste projeto, esta dissertação possui os seguintes objetivos específicos:

- Apresentar uma revisão acerca dos conceitos basilares do processamento de sinais sobre grafos (GSP);
- Investigar, a partir de trabalhos disponíveis na literatura, algumas das principais metodologias de GSP empregadas em aplicações de redes de sensores;
- Criar um modelo de grafo utilizando dados de temperatura de quadros elétricos de um grande hospital da Região Metropolitana do Recife-PE;
- Aplicar um algoritmo de identificação e localização de anomalias nas leituras de temperaturas dos quadros elétricos e comparar a precisão do modelo gerado com os dados de manutenções destes quadros.

## 1.2 MATERIAIS E MÉTODOS

O material de revisão apresentado nos próximos capítulos desta dissertação foi produzido a partir de artigos de referência publicados, sobretudo, em periódicos internacionais. Esses artigos foram obtidos em portais como o IEEEXplore (IEEE)<sup>1</sup> e o Science Direct (Elsevier)<sup>2</sup>, os quais acumulam uma parte significativa dos trabalhos na linha de processamento de sinais sobre grafos.

As atividades relacionadas a experimentos computacionais foram realizadas utilizando Python, uma linguagem de programação de sintaxe relativamente simples, de fácil implementação e que contém várias bibliotecas de comandos disponíveis na Internet. Essa linguagem foi empregada, particularmente, para modelar o grafo do estudo de caso realizado. Mais especificamente, os cálculos envolvendo a transformada de Fourier sobre o grafo (GFT, do inglês *graph Fourier transform*), a respectiva transformada inversa, a aplicação de filtros e a geração das representações gráficas foram feitos utilizando-se a biblioteca *PyGSP* (DEFFERRARD et al., ). A Biblioteca *NetworkX* (HAGBERG; SCHULT; SWART, 2008) também foi utilizada para a geração de figuras. Além destas bibliotecas, também foram utilizadas nos códigos gerados ao longo da pesquisa as bibliotecas *Numpy*, *Matplotlib*, *Scipy*, *Sys* e *Statistics*. Todas estas bibliotecas utilizadas estão disponíveis gratuitamente no site <<https://www.python.org/>>.

Os dados utilizados para a modelagem do grafo e as medições de temperatura empregadas como sinal, cuja análise é realizada no Capítulo 4, foram obtidos pelo autor a partir dos relatórios

<sup>1</sup> IEEEXplore: <<https://ieeexplore.ieee.org/>>.

<sup>2</sup> Science Direct: <<https://www.sciencedirect.com/>>.

de inspeções termográficas realizadas nos últimos cinco anos no hospital em questão. As inspeções são realizadas periodicamente e utilizam um termovisor modelo *Fluke Ti10*, adequado para capturar imagens de termografia por meio de um sensor de infravermelho, conforme descrito no manual do usuário. As imagens são geradas no formato *is2* e são analisadas manualmente pelo operador com o auxílio do software *Fluke Connect*, disponibilizado no site fabricante<sup>3</sup>.

Cada inspeção gera um relatório de análise termográfica dos quadros, em que são registrados os dados de temperatura máxima dos componentes dos quadros, observações em relação aos tipos de materiais com desvio de temperatura e recomendações de manutenções que devem ser realizadas. Em complemento a isso, também foram utilizados os dados de manutenções preventivas, preditivas e corretivas que foram executadas ao longo dos últimos cinco anos, a fim de definir as anomalias térmicas na época em que os dados foram coletados. Os dados de anomalias térmicas são utilizados para comparar e definir a precisão do algoritmo utilizado.

### 1.3 CONTRIBUIÇÕES

A partir da realização deste trabalho, pôde-se validar, para o cenário prático considerado, alguns dos processos de construção de grafos e algumas das técnicas de processamento revisadas, como a identificação de anomalias em sinais com domínios diferentes do processamento de sinais clássicos. Ao longo do Capítulo 4, é possível verificar a aplicação de técnicas de detecção de anomalias aplicadas em GSP para detectar a formação de pontos quentes em quadros elétricos.

Os resultados obtidos foram satisfatórios, no sentido de que a formação de pontos quentes foi detectada com precisão aceitável do ponto de vista prático e baixa indicação de falsos positivos, isto é, quadros que não tinham problemas sendo indicados como anômalos.

Outro ponto a se destacar é que a construção do grafo se mostrou importante no processo, pois foi possível diferenciar quadros com anomalias de quadros considerados normais a partir da aplicação do algoritmo, mesmo em casos em que as variações de temperatura em alguns quadros eram similares. Isso demonstra que a topologia e o algoritmo empregados estavam adequados à abordagem, tornando possível a validação dos critérios utilizados para a definição das arestas dos grafos.

### 1.4 ESTRUTURA DA DISSERTAÇÃO

Após este capítulo introdutório, esta dissertação se encontra estruturada como a seguir:

- No Capítulo 2, são apresentados os conceitos basilares do GSP, com ênfase nas ferramentas matemáticas mais importantes para o estudo prático realizado;

<sup>3</sup> Fluke Connect: <<https://connect.fluke.com/>>.

- O Capítulo 3 relata, de forma resumida, algumas possibilidades de aplicação do GSP em redes de sensores;
- No Capítulo 4, é descrito o estudo de caso conduzido nesta dissertação, o qual contempla a aplicação do GSP à detecção de anomalias (de temperatura) nos quadros de energia elétrica de um hospital da Região Metropolitana do Recife;
- As considerações finais deste trabalho e as possibilidades para desenvolvimento de pesquisas futuras são elencadas no Capítulo 5.

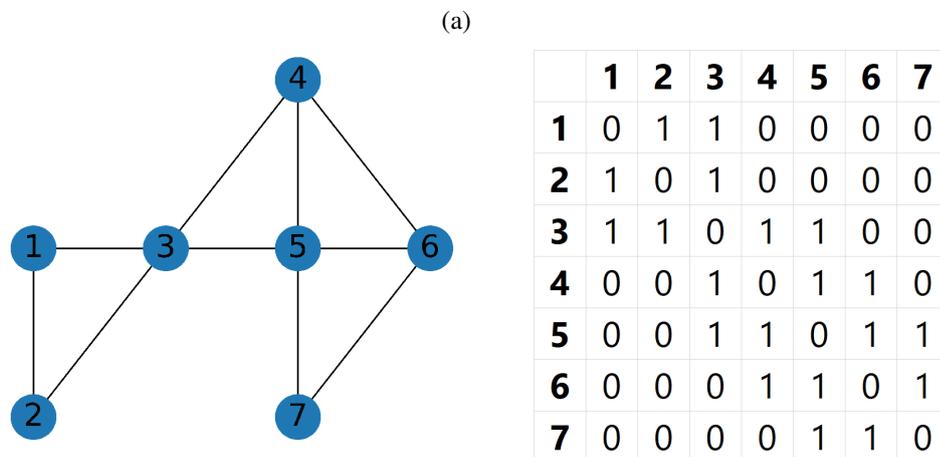
## 2 PROCESSAMENTO DE SINAIS SOBRE GRAFOS

Neste capítulo, são apresentadas as definições básicas utilizadas no processamento de sinais sobre grafos; discute-se como estas definições podem ser aplicadas para representar sinais sobre o domínio dos vértices, considerando operadores como o Laplaciano e o seu papel no estabelecimento da transformada de Fourier aplicada a sinais sobre grafos.

### 2.1 CONCEITOS BÁSICOS

Pode-se definir um grafo como um conjunto de vértices conectados por arestas. O grafo pode ser representado por  $(\mathcal{V}, \mathcal{E})$ , em que  $\mathcal{V}$  é um conjunto finito de  $N$  vértices e  $\mathcal{E}$  é um subconjunto de  $\mathcal{V}^2$  que representa as arestas. A matriz de adjacência  $A$  é utilizada para indicar se um par de vértices está conectado ou não por uma aresta. Mais precisamente, a matriz de adjacência  $A$  é construída da seguinte forma: se  $A_{i,j} \neq 0$ , então o vértice  $v_j \in \mathcal{V}$  influencia (ou seja, é *adjacente*) o vértice  $v_i \in \mathcal{V}$  e se diz que há uma aresta de  $v_j$  para  $v_i$  com peso  $A_{i,j}$ . Para o escopo desta dissertação, é suficiente considerar apenas os casos em que  $A_{i,j} \in \mathbb{R}$ . Por sua própria definição,  $A$  é uma matriz quadrada de dimensões  $N \times N$  (FEOFILOFF; KOHAYAKAWA; WAKABAYASHI, 2011; THULASIRAMAN; SWAMY, 1992).

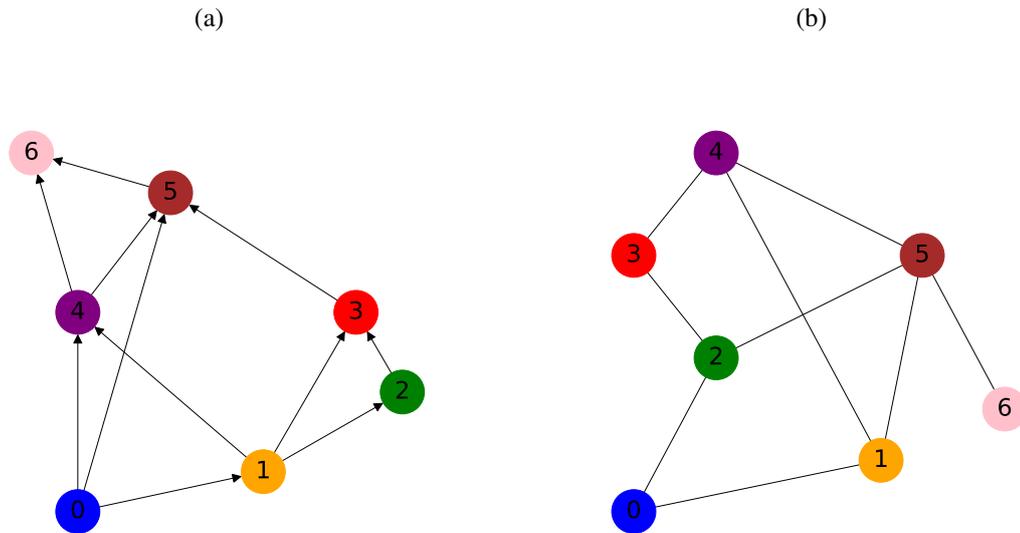
Figura 3 – Exemplo de matriz de adjacência e grafo



Fonte: O Autor (2022).

Um grafo é dito não direcionado quando sua matriz de adjacência é simétrica. Neste caso, é como se as arestas pudessem ser percorridas nos dois sentidos ou como se os vértices conectados por uma aresta se influenciassem mutuamente e com a mesma ponderação. Em um grafo direcionado, a matriz de adjacência não é simétrica e as arestas só podem ser percorridas no sentido em que estão indicadas. Também é possível que um nó tenha conexão com ele mesmo, isto é chamado de *loop*. A Figura 4a mostra um exemplo de grafo direcionado. A Figura 4b um exemplo de grafo não direcionado (RIBEIRO; LIMA, 2018).

Figura 4 – Exemplo da representação gráfica de grafos (a) direcionados e (b) não direcionados.

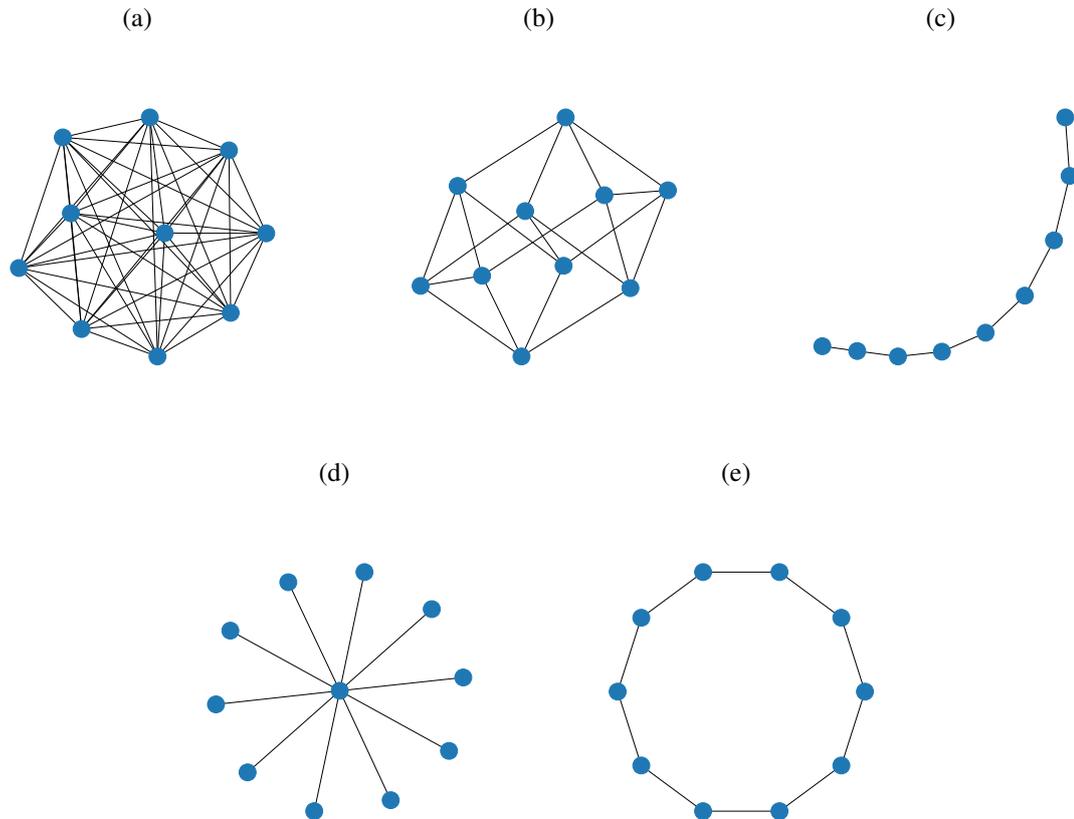


Fonte: O Autor (2022).

Ao longo deste trabalho, a palavra grafo será utilizada como sinônimo de grafo não direcionado. Caso o grafo ao qual se está referindo seja direcionado, isso será dito no texto. No tocante à ponderação de uma aresta, quanto maior o valor do respectivo peso, maior é o grau de conectividade ou de influência de um vértice sobre o outro. A intensidade da conexão entre dois nós é de grande importância para o processamento de sinais sobre grafos, uma vez que isso pode representar uma informação importante para o modelo que se está empregando. No caso de um modelo de grafo para representar uma rede de sensores que se distribuem em determinada região, a ponderação das arestas pode ser um número com proporcionalidade inversa à distância (geográfica ou geodésica) entre cada par de sensores, por exemplo. Um grafo que represente a conectividade entre subestações de um sistema de distribuição de energia pode empregar como fator de ponderação das arestas a condutância das linhas que utilizadas para conexão dessas estruturas e assim por diante (RIBEIRO; LIMA, 2018).

Um grafo está conectado se existe pelo menos um caminho que ligue quaisquer vértices. Um grafo é dito completo quando todos os vértices têm uma aresta entre si, conforme mostrado na Figura 5a. A Figura 5b mostra um grafo regular, que é quando todos os nós têm o mesmo grau. Em um grafo em linha todos os vértices estão conectados a dois outros vértices em série, exceto o primeiro e o último vértice, como mostrado na Figura 5c. Para um grafo em estrela, todos os vértices estão conectados a um vértice central. Um exemplo de grafo em estrela é mostrado na Figura 5d. Um grafo circular tem todos os vértices conectados de forma cíclica. Este tipo de grafo pode ser utilizado para a representação de um sinal cíclico, semelhante ao processamento digital de sinais clássicos. Esta característica será explorada mais adiante. Um exemplo deste tipo de grafo pode ser visto na Figura 5e.

Figura 5 – Exemplos de Grafos



Fonte: O Autor (2022).

O processo de construção do grafo é um dos passos mais importantes para o processamento de sinais sobre grafos (GSP). Este não é um processo tão trivial e diferentemente do processamento de sinais clássico, a topologia não é dada junto com o problema. É necessário construir os vértices, as arestas e seus respectivos pesos. Este processo vai depender de quais dados estão disponíveis no momento em que o grafo está sendo montado. Para isso, é preciso analisar quais as características mais relevantes por meio das quais os vértices se relacionam. Por exemplo, o peso das arestas pode ser definido a partir da distância euclidiana ou da distância geodésica entre os vértices. Neste processo, é interessante definir um limiar a partir do qual a interação entre estes dois vértices é tão pequena que pode ser considerada nula. Isto facilita a visualização do grafo, pois limita a quantidade de arestas, além de tornar a matriz de adjacência mais esparsa, facilitando a construção da matriz e diminuindo o custo computacional para realizar operações (RIBEIRO; LIMA, 2018).

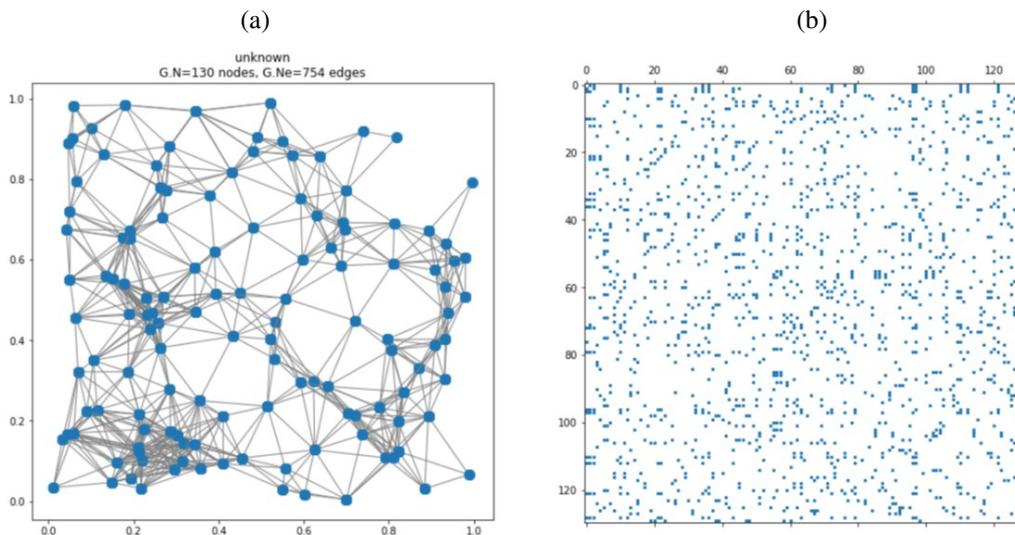
Na Figura 6a, está a representação de um grafo formado a partir de uma rede de sensores. Cada ponto representa a localização geográfica de um sensor. Na Figura 6b está a matriz de adjacência. Existem diversas funções que podem ser utilizadas para definir o peso das arestas. Como o exemplo trata de sensores de temperatura, é possível inferir algumas informações, como

a variação ser linear e inversamente proporcional a distância entre os sensores. Assim, para definir o peso das arestas do grafo mostrado na Figura 6a, pode-se utilizar a equação

$$\mathcal{E}_{ij} = e^{-\frac{1}{2}d_{ij}^2}, \quad (1)$$

se  $d_{ij} \leq \mathcal{T}$ , caso contrário  $\mathcal{E}_{ij} = 0$  ( $d_{ij}$  corresponde à distância entre os sensores associados aos nós  $i$  e  $j$ , e  $\tau$  representa um limiar).

Figura 6 – Exemplo de (a) representação de uma rede de sensores por meio de um grafo e (b) representação gráfica de sua matriz de adjacência.



Fonte: O Autor (2022).

## 2.2 SINAIS SOBRE GRAFOS

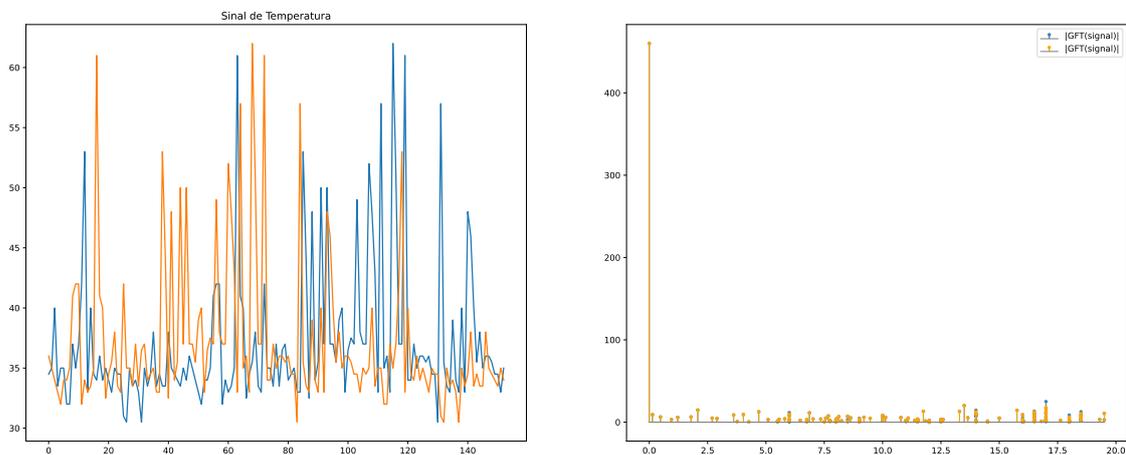
No processamento digital de sinais clássico, os sinais são amostrados sequencialmente, seja de forma temporal ou espacial. Caso o intervalo entre as amostras seja sempre igual, pode-se afirmar que, dado um sinal  $x(n)$ , ele será precedido de  $x(n - 1)$  e sucedido por  $x(n + 1)$ . A representação gráfica deste sinal pode ser feita no plano cartesiano, onde o eixo das ordenadas representa a intensidade do sinal e o eixo das abscissas a distância do ponto inicial onde foi feita a amostragem. De modo geral, é possível organizar qualquer grafo desta forma. A definição de quais nós serão associados a  $x(n)$ ,  $x(n - 1)$  ou  $x(n + 1)$  pode ser feita de forma arbitrária ou seguindo regras pré-definidas. De fato, dados coletados de forma sequencial em intervalos de tempo iguais podem ser considerados como um grafo direcionado em linha ou em anel, se assumirmos alguma condição de periodicidade (STANKOVIC; DAKOVIC; SEJDIC, 2019).

Os sinais aplicados sobre um grafo são obtidos associando-se um valor a cada um dos nós. Estes valores podem ser reais ou complexos. No caso do grafo visto na Figura 6, pode-se considerar que a leitura dos sensores de temperatura forma um sinal sobre o grafo. Os

sinais aplicados sobre um grafo podem ser analisados utilizando-se técnicas análogas às do processamento clássico, como a transformada da Fourier. Estas técnicas de processamento serão explicadas com detalhes mais adiante. Entretanto, ao olhar para a topologia do grafo mostrado na Figura 6, não fica claro como deve ser feito o ordenamento do sinal.

No processamento de sinais clássico, o ordenamento do vetor  $x = [x(0), x(1), \dots, x(n - 1)]^T$  é feito de forma bastante intuitiva, como dito anteriormente. Porém, em um grafo com topologia distinta isto é feito de forma arbitrária. Normalmente, o resultado do processamento dos dados não depende muito da forma como o sinal foi organizado. Na Figura 7 pode-se verificar que o mesmo sinal aplicado sobre o mesmo grafo, mas com ordenamentos diferentes resultam em transformadas de Fourier sobre o grafo bastante similares.

Figura 7 – Comparação da transformada de Fourier sobre o grafo de dois sinais com ordenamentos diferentes



Fonte: O Autor (2022).

Desta forma, também é possível representar um sinal aplicado sobre um grafo qualquer da mesma forma que um sinal clássico e também utilizar versões generalizadas das técnicas aplicadas ao processamento de sinais clássico. Qualquer processamento linear em um dado vértice é considerado como uma combinação linear do valor deste vértice e dos demais vértices em sua vizinhança.

### 2.2.1 Matrizes de Grau, Adjacência e Laplaciana

A matriz de adjacência é um dos principais pilares no processamento de sinais sobre grafos. Ela é construída a partir da topologia do grafo e tem propriedades importantes que serão discutidas mais adiante. A matriz de adjacência contém a relação entre os vértices de um grafo. Como parte da topologia do grafo, verifica-se que os vértices estão conectados entre si por meio de arestas às quais são atribuídos pesos. O tamanho da matriz de adjacência depende exclusivamente da quantidade de vértices presentes no grafo (RIBEIRO; LIMA, 2018).

O grau de um vértice é dado pelo somatório dos pesos das arestas que chegam ou saem dele. A matriz de grau de um grafo é uma matriz diagonal formada pelo grau de todos os vértices do grafo. De forma geral temos que a matriz de grau é dada por

$$\mathbf{D} = \begin{bmatrix} D_0 & 0 & 0 & 0 \\ 0 & D_1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & D_n \end{bmatrix}, \quad (2)$$

em que a diagonal principal da matriz é preenchida pelo grau dos vértices (FONINI, 2019).

Considerando a definição da matriz de grau do grafo, a respectiva matriz Laplaciana é dada por

$$\mathbf{L} = \mathbf{D} - \mathbf{A}, \quad (3)$$

onde  $\mathbf{D}$  é a matriz de grau e  $\mathbf{A}$  a matriz de adjacência. Aqui vale ressaltar que a matriz Laplaciana não considera *loops*, já que seus valores seriam adicionados tanto a matriz de grau quanto a matriz de adjacência, o que não alteraria o resultado da diferença. Apesar disso, normalmente a matriz Laplaciana é utilizada em aplicações de grafos não direcionados. As matrizes  $\mathbf{L}$ ,  $\mathbf{D}$  e  $\mathbf{A}$  são simétricas para grafos não direcionados (RIBEIRO; LIMA, 2018).

Considerando o exemplo de grafo mostrado na Figura 4b, sua matriz de adjacência pode ser representada por

$$\begin{bmatrix} 0 & 0,7 & 0,3 & 0 & 0 & 0 & 0 \\ 0,7 & 0 & 0 & 0 & 0,4 & 0,1 & 0 \\ 0,3 & 0 & 0 & 0,2 & 0 & 0,6 & 0 \\ 0 & 0 & 0,2 & 0 & 0,5 & 0 & 0 \\ 0 & 0,4 & 0 & 0,5 & 0 & 0,5 & 0 \\ 0 & 0,1 & 0,6 & 0 & 0,5 & 0 & 0,3 \\ 0 & 0 & 0 & 0 & 0 & 0,3 & 0 \end{bmatrix}. \quad (4)$$

É possível inferir que a matriz de grau é dada por

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1,2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1,1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1,4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1,5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0,3 \end{bmatrix} \quad (5)$$

e sua matriz Laplaciana por

$$\begin{bmatrix} 1 & -0,7 & -0,3 & 0 & 0 & 0 & 0 \\ -0,7 & 1,2 & 0 & 0 & -0,4 & -0,1 & 0 \\ -0,3 & 0 & 1,1 & -0,2 & 0 & -0,6 & 0 \\ 0 & 0 & -0,2 & 0,7 & -0,5 & 0 & 0 \\ 0 & -0,4 & 0 & -0,5 & 1,4 & -0,5 & 0 \\ 0 & -0,1 & -0,6 & 0 & -0,5 & 1,5 & -0,3 \\ 0 & 0 & 0 & 0 & 0 & -0,3 & 0,3 \end{bmatrix}. \quad (6)$$

### 2.2.2 Deslocamento de Sinais sobre Grafos

Para definir o deslocamento de um sinal sobre um grafo, pode-se considerar um sinal  $\mathbf{x}$ , aplicado sobre um grafo  $\mathcal{G}$ , em que  $x(n)$  corresponde a amostra do sinal no vértice  $n$ . Assim, pode-se dizer que o deslocamento de um sinal sobre um grafo é o “movimento” da amostra a partir do vértice  $n$  ao longo do grafo (STANKOVIC; DAKOVIC; SEJDIC, 2019). Este movimento deve ser feito em todo o grafo.

O papel de operador de deslocamento no grafo pode ser desempenhado por qualquer matriz, que ao ser multiplicada pelo sinal no grafo  $\mathbf{x}$  produz outro sinal cujas amostras resultam de uma combinação linear das amostras do sinal original. A escolha de uma matriz específica para o referido papel normalmente envolve a possibilidade de haver alguma interpretação física significativa ou algum paralelo com o processamento digital de sinais clássico (LIMA; DE OLIVEIRA NETO, 2022).

Dentre as matrizes utilizadas, é possível destacar a matriz de adjacência, escolhida como bloco de construção pela abordagem GSP proveniente do chamado processamento algébrico de sinais (PUSCHEL; MOURA, 2008); e a matriz Laplaciana, em que os conceitos básicos são desenvolvidos a partir da teoria espectral de grafos (SHUMAN et al., 2013). Embora estas duas sejam as mais estudadas e aplicadas em ferramentas de GSP, outras matrizes vem sendo propostas na literatura (STANKOVIĆ et al., 2020).

Neste trabalho será utilizada a matriz de adjacência como operador de deslocamento sobre o grafo. Deste modo, o sinal deslocado pela unidade representado como  $\mathbf{x}_1$  é obtido utilizando-se a matriz de adjacência por

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}. \quad (7)$$

Isto é, a matriz de adjacência age como um filtro de deslocamento unitário para um dado sinal. Da mesma forma, para se deslocar um sinal duas vezes, pode-se fazer

$$\mathbf{x}_2 = \mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}^2\mathbf{x}. \quad (8)$$

Estendendo, pode-se dizer que o deslocamento do sinal por um fator  $m$  é dado por

$$\mathbf{x}_m = \mathbf{A}(\mathbf{A}\mathbf{x}_{m-1}) = \mathbf{A}^m\mathbf{x}. \quad (9)$$

### 2.3 TRANSFORMADA DE FOURIER SOBRE GRAFOS

No processamento digital de sinais clássico existem vários tipos de sinais que diferem um do outro ao longo do tempo mas que possam possuir características de frequência em comum. A análise e processamento destes sinais é, na maioria das vezes, feito no domínio da frequência. Na prática o domínio da frequência é uma outra forma de redesenhar um sinal onde o eixo da abscissa não mais representa o tempo e sim a frequência. A conversão de um sinal de tempo contínuo para o domínio da frequência é feita utilizando-se a transformada de Fourier, dada por (OPPENHEIM; WILLSKY; NAWAB, 1997)

$$\mathcal{F}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt. \quad (10)$$

No processamento digital, os sinais são computados e armazenados de forma discreta, de modo que transformada de Fourier correspondente é dada por (OPPENHEIM; SCHAFER, 2013)

$$\mathcal{X}(\omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-i\omega n}. \quad (11)$$

Esta é amplamente utilizada no processamento digital de sinais, pela possibilidade de fornecer diversas informações para a análise e a resolução de problemas, auxiliando, por exemplo, no projeto de filtros (SANDRYHAILA; MOURA, 2013b).

A análise espectral do processamento digital de sinais clássico pode ser estendida para o processamento de sinais sobre grafos. A transformada de Fourier de um grafo pode ser obtida a partir da matriz de adjacência. Considerando-se um sinal  $\mathbf{x}$ , a transformada de Fourier deste sinal sobre o grafo é dada por (SANDRYHAILA; MOURA, 2013a)

$$\mathbf{X} = \mathbf{U}^{-1}\mathbf{x}, \quad (12)$$

em que  $\mathbf{U}$  é uma matriz com os autovetores da matriz de adjacência do grafo em suas colunas. Escrevendo-se  $\mathbf{X}$  como um vetor  $X(k)$  com  $k = 0, 1, \dots, N - 1$ , tem-se que os elementos do vetor  $\mathbf{X}$  podem ser escritos como

$$X(k) = \sum_{n=0}^{N-1} x(n)u_k(n). \quad (13)$$

Assim,  $X(k)$  é uma projeção do sinal considerado sobre o  $k$ -ésimo autovetor. Pode-se entender a transformada de Fourier como uma decomposição do sinal sobre os autovetores (DERI; MOURA, 2017). Como a matriz de adjacência de um grafo não direcionado é simétrica, seus autovetores formam uma base ortogonal.

Analogamente, a transformada de Fourier inversa é obtida com

$$\mathbf{x} = \mathbf{U}\mathbf{X}. \quad (14)$$

Ao aplicar estes conceitos a um grafo em anel direcionado, pode-se verificar que a auto-decomposição da matriz de Adjacência coincide com a transformada de Fourier discreta (SANDRYHAILA; MOURA, 2014).

A decomposição espectral de um sinal sobre um grafo pode ser feita na matriz Laplaciana da mesma forma que na matriz de adjacência, utilizando a auto-decomposição dessa matriz, isto é  $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$  (SHUMAN et al., 2013). Logo, a transformada de Fourier de um sinal  $\mathbf{x}$  baseada na decomposição da matriz Laplaciana pode ser dada por

$$\mathbf{X} = \mathbf{U}^{-1}\mathbf{x}, \quad (15)$$

em que  $\mathbf{U}$  é a matriz com os autovetores da matriz Laplaciana em suas colunas. Da mesma forma, a transformada de Fourier inversa é dada por

$$\mathbf{x} = \mathbf{U}\mathbf{X}. \quad (16)$$

Apesar de estes métodos de obtenção das GFTs de um grafo já estarem consolidados na literatura atual, outras definições alternativas ainda são objetos de estudo, como visto em (SARDELLITTI; BARBAROSSA; LORENZO, 2017), em que os autores apresentam uma GFT cuja base para se adequar melhor à representação de sinais sobre grafos direcionados.

A transformada de Fourier sobre um grafo, identificada pelo acrônimo GFT (do inglês *Graph Fourier Transform*), implica na transformação do sinal de um domínio de vértices para o domínio da frequência (SHUMAN; RICAUD; VANDERGHEYNST, 2016). Este conceito é bem intuitivo quando se fala de um sinal gerado no domínio do tempo. Porém, como foi visto anteriormente, os sinais representados sobre um grafo podem se referir a domínios não só no tempo, como também leitura de sensores, potência elétrica em subestações, etc.

A justificativa matemática relacionada a interpretação do domínio da frequência para sinais em grafos já se encontram consolidadas na literatura. Dentro deste contexto, os autovalores do grafo são análogos as frequências do processamento de sinais clássico. Da mesma forma, os autovetores do grafo são análogos as componentes de frequência (RIBEIRO; LIMA, 2018).

Para um sinal clássico, considerando-se o próprio conceito de frequência, espera-se que sinais que variem mais rapidamente tenham componentes de alta frequência com maior magnitude. Devido a isto, quanto maior a variação de valores associado a um autovetor, este e o seu respectivo autovalor estão relacionados a frequências mais altas. Portanto, de modo geral, é necessário ordenar a base de autovetores (colunas da matriz  $\mathbf{U}$ ) para que obedeça esta característica. Independente da base de autovetores e autovalores utilizada, a ordenação pode ser feita utilizando a variação total  $TV$  (do inglês, *total variation*), definida como o somatório da magnitude das diferenças entre duas amostras consecutivas (SANDRYHAILA; MOURA, 2014)

$$TV((s)) = \sum_{n=0}^{N-1} |s_n - s_{(n-1 \bmod N)}|, \quad (17)$$

em que  $s$  é uma série temporal periódica de período  $N$ .

No caso de grafos, as amostras consecutivas são definidas pelo operador de deslocamento no grafo, podendo-se utilizar a matriz de adjacência. Desta forma tem-se que a variação total para grafos,  $TV_G$ , é definida por:

$$TV_G(s) = \|s - A^{\text{norm}}s\|_1, \quad (18)$$

em que  $\|\cdot\|_1$  denota a norma  $l_1$  e  $A^{\text{norm}}$  é a matriz de adjacência normalizada (SANDRYHAILA; MOURA, 2014).

De fato, altas componentes de frequência em um grafo intuitivamente indicam que o sinal aplicado sobre ele varia de forma mais abrupta entre os nós adjacentes. Analogamente, componentes de frequência menores indicam que o sinal varia suavemente ao longo do grafo.

## 2.4 FILTRAGEM

Um filtro em um grafo pode ser implementado tanto no domínio dos vértices quanto no domínio da frequência. Considerando função de transferência de um grafo como  $G(\Lambda)$ , no domínio espectral a implementação de um filtro pode ser feita diretamente. A GFT do sinal dada por  $\mathbf{X} = \mathbf{U}^{-1}\mathbf{x}$  pode ser diretamente multiplicada por  $G(\Lambda)$  para obter a transformada do sinal de saída  $\mathbf{Y} = G(\Lambda)\mathbf{X}$ . Por fim, o sinal de saída será dado pela transformada de Fourier inversa de  $\mathbf{Y}$ . Logo,  $\mathbf{y} = \mathbf{U}\mathbf{Y}$ . Este procedimento, por envolver a multiplicação de muitas matrizes, pode ter um custo computacional bastante alto para grafos com muitos vértices e matrizes de adjacência muito grandes. Neste caso, seria mais vantajoso a implementação dos filtros no domínio dos vértices (STANKOVIC; DAKOVIC; SEJDIC, 2019).

Um sistema em um grafo pode ser implementado como a combinação linear de um sinal e de suas versões deslocadas. Assim, podemos escrever o sinal de saída da seguinte forma

$$y = h_0\mathbf{A}^0\mathbf{x} + h_1\mathbf{A}^1\mathbf{x} + \dots + h_{M-1}\mathbf{A}^{M-1}\mathbf{x} = \sum_{m=0}^{M-1} h_m\mathbf{A}^m\mathbf{x}, \quad (19)$$

sendo  $\mathbf{A}^0 = \mathbf{I}$ , a matriz identidade.

Para implementar o mesmo filtro descrito anteriormente utilizando (19), é necessário achar os coeficientes  $h_0, h_1, \dots, h_{M-1}$  de modo que a transformada de Fourier de  $H(\Lambda)$  seja aproximadamente igual a  $G(\Lambda)$ . Conforme mostrado em (STANKOVIC; DAKOVIC; SEJDIC, 2019), a função de transferência no domínio dos vértices é dada por

$$H(\lambda_k) = \frac{Y(k)}{X(k)} = (h_0 + h_0\lambda_k + \dots + h_{M-1}\lambda_k^{M-1}). \quad (20)$$

Como espera-se que  $H(\Lambda_k)$  e  $G(\Lambda_k)$  sejam iguais para  $k = 0, 1, \dots, N - 1$ , tem-se o seguinte sistema de equações

$$\begin{cases} h_0 + h_1\lambda_0^1 + \dots + h_{M-1}\lambda_0^{M-1} = G(\lambda_0) \\ h_0 + h_1\lambda_1^1 + \dots + h_{M-1}\lambda_1^{M-1} = G(\lambda_1) \\ \vdots \\ h_0 + h_1\lambda_{N-1}^1 + \dots + h_{M-1}\lambda_{N-1}^{M-1} = G(\lambda_{N-1}), \end{cases} \quad (21)$$

o qual pode ser representado matricialmente por

$$\mathbf{V}_\lambda h = g, \quad (22)$$

em que  $\mathbf{V}_\lambda$  é a matriz de Vandermonde dos autovalores  $\lambda_k$ ,  $h$  é o vetor dos coeficientes procurados transposto e  $g$  é a diagonal da matriz  $G(\lambda_{N-1})$ .

## 2.5 OUTRAS OPERAÇÕES REALIZADAS SOBRE GRAFOS

Esta seção irá mostrar outras operações utilizadas em GSP, porém, como estão fora do escopo geral deste trabalho, seus conceitos não serão aprofundados.

### 2.5.1 Deslocamentos no domínio da frequência

O deslocamento de um sinal sobre um grafo no domínio da frequência pode ser definido da mesma forma como é feito no domínio dos vértices (STANKOVIC; DAKOVIC; SEJDIC, 2019). Considerando que o sinal sobre um grafo  $z(n)$  é o produto de dois sinais  $x(n)y(n)$ , então, aplicando (13), sua GFT é dada por

$$GFT[z(n)] = GFT[x(n)y(n)] = \sum_{n=0}^{N-1} x(n)y(n)u_k(n) = \sum_{n=0}^{N-1} \sum_{i=0}^{N-1} X(i)u_i(n)y(n)u_k(n), \quad (23)$$

ficando

$$GFT[x(n)y(n)] = \sum_{i=0}^{N-1} X(i)Y_i(k), \quad (24)$$

de forma que

$$Y_i(k) = \sum_{n=0}^{N-1} y(n)u_i(n)u_k(n) \quad (25)$$

pode ser considerado como um deslocamento de  $Y(k)$  por  $i$ .

### 2.5.2 Convolução de Sinais sobre Grafos

Definindo-se a convolução de dois sinais de tempo contínuo como

$$y(n) = x(n) * h(n), \quad (26)$$

a convolução destes sinais no domínio da frequência é dada por

$$Y(k) = X(k)H(k). \quad (27)$$

Desta forma, o resultado da operação de convolução entre os sinais  $x(n)$  e  $h(n)$  é a transformada de Fourier sobre o grafo inversa, também chamada de IGFT (do inglês *Inverse Graph Fourier Transform*), de  $Y(k)$ , dada por

$$y(n) = x(n) * h(n) = \sum_{k=0}^{N-1} Y(k)u_k(n) = \sum_{k=0}^{N-1} X(k)H(k)u_k(n). \quad (28)$$

### 3 APLICAÇÕES DE GSP EM REDES DE SENSORES

Este capítulo descreve algumas técnicas de processamento de sinais sobre grafos aplicadas em redes de sensores. Inicialmente, são apresentadas duas técnicas de identificação de anomalias em redes de sensores. A segunda técnica, além de identificar que há uma anomalia, também mostra como encontrar qual o sensor que está com leituras irregulares. Na última parte é mostrada uma técnica para determinar os locais mais suscetíveis a concentrar poluentes no ar, baseando-se numa modelagem feita de uma cidade em que os edifícios são considerados como nós e a influência de um edifício sobre outro como os vértices.

#### 3.1 DETECÇÃO DE MAL FUNCIONAMENTO EM REDES DE SENSORES

Atualmente a utilização de redes de sensores para monitoramento das condições climáticas é bastante comum. Sensores de ar, temperatura e pressão, por exemplo, são largamente utilizados para realizar estas tarefas (MAO et al., 2019). Porém a manutenção desta rede de sensores demanda bastante tempo de deslocamento de um operador até o local de instalação para que sejam verificadas as condições físicas e o funcionamento dos sensores.

Algumas técnicas de processamento de sinais sobre grafos (GSP) podem ser empregadas para detectar e localizar anomalias no sinal, em especial no domínio da frequência (SANDRYHAILA; MOURA, 2014). Em alguns casos estas anomalias podem ser interpretadas como mal funcionamento dos sensores, entretanto, esta não é uma regra. A partir da topologia de instalação dos sensores e do sinal gerado, é possível construir um grafo e aplicar técnicas de GSP para filtragem dos sinais. Normalmente, componentes de baixa frequência estão associados a pequenas oscilações do sinal, enquanto componentes de alta frequência estão associadas a oscilações mais abruptas.

Grandezas físicas como pressão atmosférica, temperatura, etc. têm pequenas variações com o distância. Desta forma é seguro afirmar que sensores geograficamente próximos tendem a ter leituras semelhantes. Desta forma, medidas muito diferentes em sensores próximos podem indicar o mal funcionamento de um desses sensores. Este conceito já foi aplicado em trabalhos como (DRAYER; ROUTTENBERG, 2018).

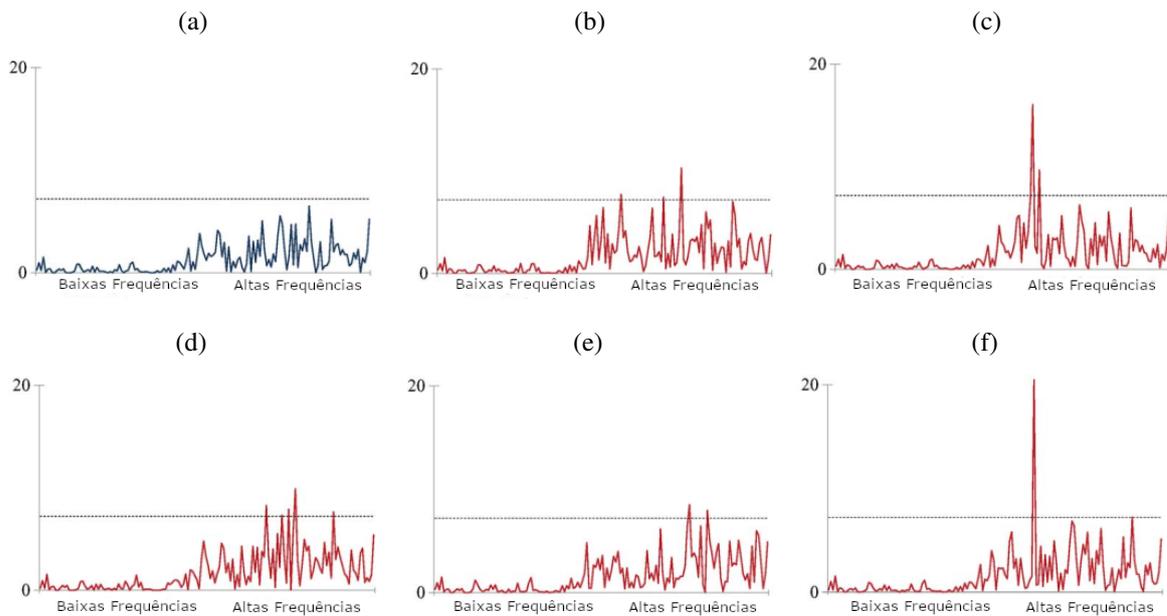
É comum utilizar filtros para processar sinais sobre grafos. Igualmente ao processamento de sinais clássicos, os filtros podem ser classificados como passa-baixa, passa-alta e passa-faixa. Segundo a técnica proposta em (SANDRYHAILA; MOURA, 2014), é possível detectar o mal-funcionamento de sensores extraíndo-se os componentes de alta frequência com um filtro passa-alta. Após isso, deve-se submeter o resultado a uma limiarização. Se algum dos coeficientes da transformada de Fourier do sinal for maior do que este limiar, isso significa que há um mal-funcionamento em um dos sensores. O autor define este limiar como o máximo valor absoluto dos coeficientes da transformada de Fourier das últimas três medições do sinal, assumindo que a

mudança no valor da amostra associada a cada vértice com o tempo é considerada.

Para verificar esta tese, é feito um experimento com os dados de 150 estações de monitoramento climático espalhadas nos Estados Unidos. Foram utilizados os dados do ano de 2003, resultando num total de 54.750 medições. Com uma alteração de 20 graus Fahrenheit em uma das medições, o sinal é processado através de um filtro passa-alta, sendo os dados dos últimos três dias utilizados para definir o limiar a partir do qual é detectado o mal-funcionamento. Este teste foi realizado para cada uma das medições, resultando num total de 54.750 testes. Estas alterações foram detectadas em 89% dos casos (SANDRYHAILA; MOURA, 2014).

A Figura 8 mostra os gráficos de magnitude dos coeficientes espectrais do sinal real (8a) e dos sinais corrompidos na cidade de Colorado Springs (8b), Tampa (8c), Atlanta (8d), Reno (8e) e Portland (8f). É possível verificar que uma alteração no sinal real leva a um aumento das magnitudes dos coeficientes espectrais na região de alta frequência.

Figura 8 – Magnitude dos Coeficientes Espectrais.



Fonte: Adaptado de Sandryhaila; Moura (2014)

Apesar da boa precisão em detectar estas pequenas alterações no sinal, não é fornecida uma forma de localizar onde as alterações ocorreram. Com esta técnica, também não é possível determinar se há uma ou mais alterações no sinal.

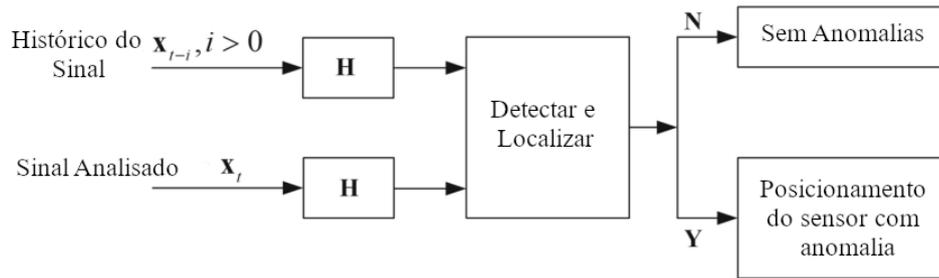
### 3.2 DETECÇÃO E LOCALIZAÇÃO DE ANOMALIAS EM REDES DE SENSORES

A maioria das técnicas de detecção de anomalias utilizadas em GSP normalmente identifica apenas que há uma anomalia no sinal. A detecção de qual o nó exato em que está a

anomalia é mais difícil uma vez que a transformada de Fourier do grafo é uma operação que considera todos os nós associados ao grafo.

O algoritmo proposto em (MOU et al., 2020) verifica não só se há uma anomalia no sinal, mas propõe uma técnica para identificar onde a anomalia está ocorrendo. Para isto, é utilizado um filtro de dois canais, conforme o esquema mostrado na Figura 9. Na entrada do filtro, o sinal que se quer analisar e o seu histórico são separados, assumindo-se que não há anomalias no histórico (MOU et al., 2020).

Figura 9 – Filtro de dois canais.



Fonte: Adaptado de Mou et al. (2020).

De modo geral, espera-se que haja poucas oscilações nos sinais gerados por redes de sensores. Assim as componentes de alta frequência devem ter baixa magnitude. Caso algum dos sensores apresente um defeito e haja uma alteração, a leitura apresentada por este sensor provavelmente irá diferir dos sensores vizinhos. Isto naturalmente irá levar a um aumento na magnitude das componentes de alta frequência do sinal.

Ao se comparar as transformadas de Fourier do histórico do sinal com a do sinal que se quer analisar, este aumento da magnitude das componentes de alta frequência pode indicar que há uma anomalia no sinal. Entretanto, como dito anteriormente, apenas com a análise da transformada de Fourier dos sinais não é possível determinar em que nó está localizada a anomalia.

Na Figura 9, os dados do sinal de entrada são representados por  $x$ , o filtro aplicado por  $H$  e o sinal filtrado por  $y = Hx$ . O filtro  $H$  é um passa-alta, utilizado para extrair as componentes de alta frequência do sinal.

Para realizar a detecção e localização da anomalia no sinal, o esquema proposto em (MOU et al., 2020) realiza a filtragem do sinal suspeito de conter uma anomalia  $x_t$  e do histórico do sinal  $X_{t-i}$ , em que  $i = 1, 2, \dots, K$ . Os sinais filtrados  $y_t = Hx_t$  e  $y_{t-i} = Hx_{t-i}$ , em que  $i = 1, 2, \dots, K$ , são então utilizados no estágio de detecção e localização de anomalias. Para isto são calculados dois termos, a amplitude residual máxima

$$\theta_t = \|y_t - \frac{1}{K} \sum_{i=1}^K y_{t-i}\|_{\infty} \quad (29)$$

e a amplitude residual média

$$\theta_K = \frac{2}{K(K-1)} \sum_{\substack{i,j=1,\dots,K \\ i \neq j}} \|y_{t-i} - y_{t-j}\|_\infty. \quad (30)$$

Caso haja mal-funcionamento de algum sensor, é de se esperar que haja um aumento nas oscilações de alta frequência, uma vez que a leitura de um determinado nó irá diferir dos nós adjacentes. Desta forma, caso a amplitude residual máxima do sinal analisado seja muito maior do que a amplitude residual média do histórico do sinal, é provável que haja uma anomalia neste sinal.

É importante ressaltar que, dentro do próprio sinal, é natural que haja variações ao longo do tempo e que, em determinado ponto, a amplitude residual máxima de um sinal pode ser maior do que a amplitude residual média, ainda que não haja anomalias no sinal. Portanto, para se comparar os valores de  $\theta_t$  e  $\theta_K$ , é necessário multiplicar  $\theta_K$  por um fator  $\gamma$ . Isto serve para evitar a sinalização de anomalias de forma incorreta em um sinal.

O último passo do esquema proposto, para localizar onde está a anomalia, é o de localizar em que nó foi detectada a anomalia. Isto é determinado como

$$k_t^* = \operatorname{argmax} |z_t(k)|, \quad (31)$$

em que  $k_t^*$  é o índice do nó onde a anomalia foi detectada e

$$z_t = y_t - \frac{1}{K} \sum_{i=1}^K y_{t-i}. \quad (32)$$

Conforme destacado pelo autor em (MOU et al., 2020), as vezes é possível detectar uma anomalia na leitura dos sensores calculando-se a diferença entre a amplitude residual máxima do sinal sem aplicar o filtro passa alta. Isto é, aplicando-se a fórmula para o calculo de amplitude residual máxima e média, respectivamente (29) e (30), diretamente no sinal  $x_t$ , isto é

$$\theta_t = \|x_t - \frac{1}{K} \sum_{i=1}^K x_{t-i}\|_\infty, \quad (33)$$

$$\theta_K = \frac{2}{K(K-1)} \sum_{\substack{i,j=1,\dots,K \\ i \neq j}} \|x_{t-i} - x_{t-j}\|_\infty. \quad (34)$$

Entretanto, a localização do nó que apresenta a anomalia se torna difícil, uma vez que o sinal contém muitas componentes de baixa frequência, o que pode acabar escondendo a localização da anomalia.

Também é importante ressaltar que, a depender da topologia do problema, como no caso de sensores de umidade e temperatura do ar, variações climáticas abruptas podem levar a discrepância entre as medições, levando a detecção de falsas anomalias. Entretanto, estes fenômenos são passageiros e tendem a se normalizar assim que as condições normais são reestabelecidas.

### 3.3 MODELAGEM DE POLUIÇÃO DO AR EM AMBIENTES URBANOS

Atualmente, um dos grandes problemas em grandes cidades é a poluição do ar atmosférico e seus efeitos sobre a saúde humana. Especialmente em países com economias emergentes, a poluição do ar tem aumentado de forma constante como resultado de atividades industriais e maior utilização de veículos que utilizam energia de combustíveis fósseis. Estima-se que mais de 8 milhões de pessoas morram por ano em decorrência dos efeitos da poluição do ar (AZAM; RIAHI-ZANJANI; BALALI-MOOD, 2016).

O monitoramento da qualidade do ar tem se tornado cada vez mais necessário em grandes centros. Como isto demanda a instalação de redes de sensores, definir quais os melhores locais para estas instalações pode se tornar uma tarefa complexa que demanda um estudo longo de uma grande área. Conforme mostrado em (JAIN; MOURA; KONTOKOSTA, 2014), é possível prever a influência da poluição emitida pelo consumo de combustíveis poluentes pelos edifícios em uma grande cidade, detectando *hotspots* de poluição, onde a concentração de matéria particulada e poluentes será maior. Monitorando estas áreas é possível obter um bom retrato da qualidade do ar em uma grande cidade.

Pode-se considerar um ambiente urbano como um grafo. Para isto, os edifícios podem ser considerados como os vértices. Cada edifício tem uma localização geográfica única e uma informação de consumo de combustível poluente associado. Além disso, eles se relacionam com o ambiente ao seu redor. Numa cidade com  $N$  edifícios, pode-se representar esta informação por  $G = (V, W)$ , em que  $V = v_0, v_1, \dots, v_{N-1}$  são os vértices do grafo  $G$  e  $W$  a matriz de adjacência ponderada. A matriz de adjacência corresponde a forma como cada edifício, ou vértice, se relaciona com os demais, sendo comum a utilização da distância euclidiana (1). Entretanto, esta relação também pode ser definida com base nas propriedades de cada edifício, como o tamanho ou consumo de energia.

O sinal aplicado sobre este grafo pode ser definido como um vetor  $X \in \mathbb{R}^N$ , sendo  $X = (x_0, x_1, \dots, x_{N-1})$ . O sinal  $X$  pode ser formado pelo consumo de energia, água, gás para aquecimento, dados físicos do edifício, etc.

Para adaptar este modelo ao da cidade de Nova York, as arestas entre os edifícios foram definidas em (JAIN; MOURA; KONTOKOSTA, 2014) utilizando o modelo de dispersão Gaussiano

$$C(x, y, z, H) = \frac{Q}{2\pi \cdot u \cdot \sigma_y \cdot \sigma_z} \cdot \exp\left[-\frac{y^2}{2\sigma_y^2}\right] \cdot \left[\exp\left[-\frac{(z-H)^2}{2\sigma_z^2}\right] + \exp\left[-\frac{(z+H)^2}{2\sigma_z^2}\right]\right], \quad (35)$$

onde  $C$  é a concentração média de poluentes (em  $g/m^3$ ),  $x$  é a distância da fonte de poluição na direção do vento (em  $m$ ),  $y$  é a distância da fonte de poluição perpendicular a direção do vento (em  $m$ ),  $z$  é a distância vertical em relação ao solo (em  $m$ ),  $H$  é a distância vertical, em relação ao solo, em que os poluentes estão sendo emitidos (em  $m$ ),  $u$  é a velocidade do vento (em  $m/s$ ),  $Q$  é a quantidade de poluente emitida pela fonte de poluição (em  $g/s$ ),  $\sigma_y$  é o parâmetro

de dispersão urbano na direção horizontal e  $\sigma_z$  é o parâmetro de dispersão urbano na direção vertical (JAIN; MOURA; KONTOKOSTA, 2014).

O objetivo de utilizar este tipo de distribuição é poder analisar como a poluição do ar em uma grande cidade é distribuída de forma mais simplificada, verificando como determinadas áreas são influenciadas pelos seus arredores. Também é importante ressaltar que o objetivo deste modelo não é quantificar a concentração dos poluentes nestas áreas, mas sim verificar quais locais estão suscetíveis a maiores concentrações de poluentes. Com este tipo de informação, a instalação de sensores para verificação da qualidade do ar pode ser mais direcionada para estas áreas. A fim de simplificar a modelagem, a direção do vento é considerada como constante e sempre numa mesma direção (do oeste para o leste), as fontes de emissão e exposição estão todas no nível do solo, não há dispersão no eixo  $z$  e  $Q$  é um valor estático de todas as emissões, ao invés de uma taxa dada em  $g/s$  (JAIN; MOURA; KONTOKOSTA, 2014).

Aplicando estas suposições em (35), tem-se

$$C = \frac{Q}{\sigma_y} \exp \left[ -\frac{y^2}{2\sigma_y^2} \right], \quad (36)$$

em que

$$\sigma_y = 0.22x(1 + 0.0004x)^{-\frac{1}{2}} \quad (37)$$

para a cidade de Nova York (JAIN; MOURA; KONTOKOSTA, 2014).

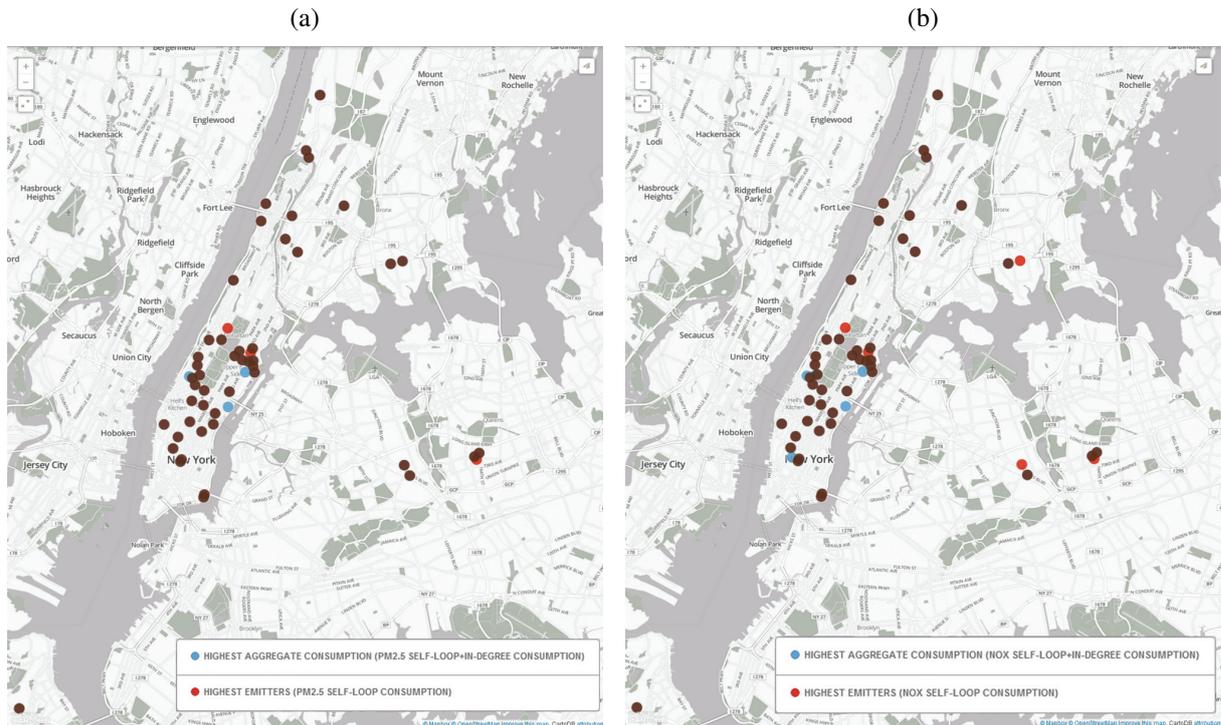
Os elementos da matriz de adjacência são então dados por

$$A_{ij} = \frac{1}{\sigma_y} \exp \left( -\frac{[y_{ij}]^2}{2\sigma_y^2} \right) \quad (38)$$

se  $y_{ij}$  for maior do que  $\gamma$  e  $A_{ij} = 0$  em outros casos, sendo  $y_{ij}$  a distância física Euclidiana entre os edifícios  $i$  e  $j$ . O limiar  $\gamma$  é igual a  $2.54e - 10$ , obtido considerando  $x$  e  $y$  iguais a 1.500m em (35) e (37). Também é designado um valor para a diagonal principal da matriz de adjacência, a fim de representar o impacto da poluição emitida pelo próprio edifício sobre ele mesmo. O sinal considerado para o experimento é um vetor com o consumo total de três tipos de óleo de aquecimento. Como cada combustível contribui de forma diferente para a poluição, os valores foram ponderados antes de serem processados sobre o grafo (JAIN; MOURA; KONTOKOSTA, 2014).

A fim de comparar os resultados da verificação de locais mais suscetíveis a poluição emitida pelos edifícios, dois métodos são empregados. O convencional utiliza os dados de emissão de poluente dos próprios edifícios para criar uma classificação destes locais. O segundo método utiliza o grau dos vértices do grafo criado. Esta informação representa o quanto que cada edifício é influenciado pela sua própria emissão de poluentes e pela emissão dos edifícios adjacentes. A Figura 10a mostra a estimativa dos 50 edifícios mais expostos a poluentes do tipo PM2.5 (matéria particulada de até 2.5mm de diâmetro); e a Figura 10b mostra a estimativa dos

Figura 10 – Estimativa de Exposição a Poluição de edifícios na cidade de Nova York.



Fonte: Jain; Moura; Kontokosta (2014)

50 edifícios mais expostos a poluentes do tipo NOx (óxidos de nitrogênio) (JAIN; MOURA; KONTOKOSTA, 2014).

Na Figura 10, os maiores emissores de poluentes estão marcados com a cor vermelha, enquanto os locais estimados como tendo maiores concentrações de poluentes devido as próprias emissões e seus arredores estão marcados na cor azul. Locais em que os maiores emissores de poluentes e estimados como mais afetados coincidem, estão marcados com vermelho escuro. Como é possível observar na Figura 10, pode-se verificar que os locais suscetíveis a maiores concentrações de poluentes nem sempre estão ligados aos edifícios com maiores concentrações de emissão. Isto se dá pois em alguns locais pode ocorrer que a influência somada de vários emissores menores acarrete em uma concentração maior de poluentes naquela área. Em contrapartida, locais com uma grande fonte emissora de poluentes cercada por fontes menores leve a uma dispersão maior destes poluentes, ocasionando uma menor concentração nestes locais (JAIN; MOURA; KONTOKOSTA, 2014).

## 4 ESTUDO DE CASO

Para aplicação dos conceitos desenvolvidos ao longo do capítulo anterior, foi realizado um estudo de caso utilizando as leituras de temperatura de quadros de distribuição de energia elétrica de um grande hospital da região metropolitana do Recife/PE.

Na manutenção de quadros elétricos, um dos problemas mais recorrentes é a formação de pontos quentes que, caso não sejam identificados e tratados, podem levar a uma parada parcial ou total do funcionamento do quadro. Desta forma é interessante aplicar os conceitos de detecção de anomalias para identificar se a mudança de temperatura de um quadro elétrico se deve diretamente ao aumento de carga ou se há a formação de um ponto quente.

Normalmente, a análise da temperatura dos quadros é feita por um operador que realiza a coleta dos sinais e leva em conta fatores como o qual o setor do hospital que o quadro está alimentando, as potências e os tipos de equipamentos que estão ligados ao quadro, o histórico de manutenção do quadro, tipo de ambiente em que ele está instalado, entre outros. Todo este processo demanda tempo e experiência do operador. Assim, foram aplicadas algumas técnicas de processamento de sinais sobre grafos para que, a partir dos dados de temperatura coletados dos quadros, fosse possível identificar anomalias no sinal e apontar possíveis defeitos.

Outra premissa importante é que o procedimento para coleta das amostras de temperatura é padronizado, sendo feito apenas com o hospital em pleno funcionamento. Com os quadros a plena carga, os pontos quentes em quadros elétricos são mais facilmente identificáveis. Também vale destacar que as amostras não são coletadas todas na mesma data. Apesar disso, o sinal de temperatura não é comprometido, pois ele tende a variar lentamente ao longo dos dias, uma vez que o surgimento de pontos quentes não é um processo abrupto e leva tempo até demonstrar seus efeitos. Desta forma, desde que as amostras sejam coletadas em horários semelhantes com diferença de poucos dias elas tendem a retratar a realidade. Os horários para a realização das inspeções foram definidos com base na memória de massa de consumo de energia do hospital, onde foram verificados os picos de consumos ao longo do dia e da semana.

### 4.1 CARACTERÍSTICAS DO SINAL E MONTAGEM DO GRAFO

Os dados utilizados foram coletados entre novembro de 2017 e outubro de 2021. Para definir os quadros utilizados nesta aplicação, foi realizado um pré-filtro onde levou-se em conta o histórico do sinal de temperatura destes. No período de tempo definido, alguns quadros foram desativados, substituídos por quadros novos, ou ainda foram instalados apenas posteriormente. Desta forma, o sinal de temperatura destes quadros estava incompleto. Portanto, não foram considerados para a realização dos testes.

Com isto, sobraram 153 quadros que são utilizados como vértices do grafo. Entre os anos de 2017 e 2021 foram realizadas 17 medições de temperatura em todos estes quadros, totalizando

2601 amostras.

Na literatura atual existem diversas formas de modelar este tipo de grafo (MEI; MOURA, 2016; SARDELLITTI; BARBAROSSA; LORENZO, 2016). Para a definição dos vértices foi proposta a utilização dos critérios a seguir.

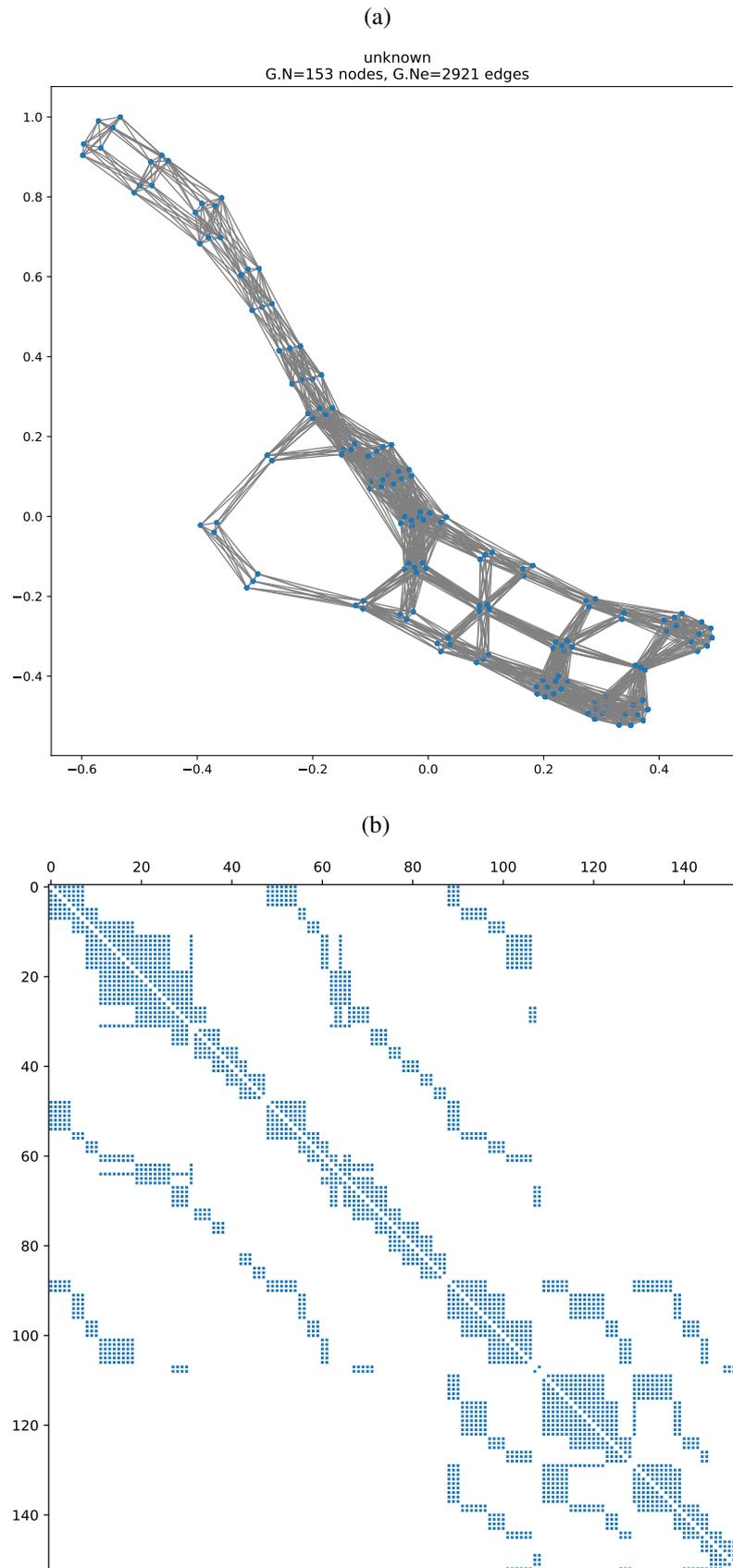
- Quadros que alimentam o mesmo setor têm arestas entre si com valor unitário, pois, apesar de alimentarem equipamentos diferentes, caso o setor esteja em funcionamento o consumo de energia elétrica será até certo ponto proporcionalmente semelhante;
- Quadros que alimentam setores diferentes, mas que estejam no mesmo andar cujos blocos são adjacentes também têm arestas entre si com valor unitário, pois, de modo geral, os setores adjacentes funcionam em horários semelhantes;
- Setores que têm algum grau de dependência, como por exemplo bloco cirúrgico e UTI cirúrgica, têm arestas entre si com valor unitário;
- Quadros que estão em pavimentos diretamente superiores ou diretamente inferiores têm arestas entre si de valor igual a 0,5, pois estão sendo alimentados por ramais próximos no barramento de distribuição.

O grafo gerado é não direcionado. Um dos motivos para essa escolha é que as técnicas mostradas ao longo do Capítulo 3 se aplicam a grafos não direcionados. O grafo poderia ser gerado de forma direcionada, utilizando, por exemplo, as conexões elétricas existentes entre os quadros, onde as arestas partiriam de um quadro que é alimentado para o alimentador. Pode parecer contra intuitivo as arestas partirem no sentido contrário ao da alimentação elétrica, porém o foco deste grafo é a detecção de superaquecimento. Neste caso, quadros que demandam muita energia elétrica de um outro quadro acaba por elevar a temperatura do segundo. Consequentemente, o efeito térmico é contrário ao sentido da alimentação elétrica. Entretanto essa abordagem não foi possível, pois o hospital em que os quadros foram analisados é um prédio antigo com mais de 50 anos de construção e não conta mais com os projetos de alimentação e distribuição elétrica originais.

A fim de facilitar a visualização do grafo, foram geradas coordenadas em duas dimensões. Foi utilizado o algoritmo *Spring* (*seed* = 40) para gerar as coordenadas. Este algoritmo é fornecido pela própria biblioteca do PyGSP, mostrada na Seção 1.2. A Figura 11a mostra a representação do grafo utilizando o algoritmo *Spring* e a Figura 11b a matriz de adjacência. o grafo gerado possui 153 nós e 2921 arestas e é do tipo não direcionado.

Antes de iniciar as análises do sinal, um ponto a se destacar é que a temperatura média dos quadros varia ao longo do ano. Sendo verificadas temperaturas mais elevadas nos meses de verão e menores nos meses de inverno. Como todos os serviços do hospital funcionam normalmente durante o ano inteiro, a hipótese mais provável para esta sazonalidade é a utilização

Figura 11 – Grafo e Matriz de Adjacência.



Fonte: O Autor (2022).

de aparelhos de ar-condicionado. Os aparelhos de ar-condicionado são mais utilizados durante os meses mais quentes, levando a um aumento da carga. Com o aumento da carga as temperaturas médias internas dos quadros também se elevam. Além de a temperatura ambiente nos corredores, onde fica a maioria dos quadros, também ser mais elevada.

Na Figura 12 é possível ver a variação da temperatura média dos quadros ao longo do tempo. Nos meses de verão as temperaturas são mais elevadas, enquanto que nos meses de inverno elas são menores. No gráfico mostrado na Figura 12, há dois pontos que chamam a atenção. A média de temperatura dos meses de julho de 2020 e julho de 2021 está bem abaixo das demais. Devido a pandemia de Covid-19, parte dos serviços ambulatoriais do hospital foram desativados por algum tempo. Isto coincidiu com a diminuição da temperatura média dos quadros. Em relação aos dados coletados no ano de 2022, eles serão utilizados mais a frente para verificar se as conclusões tiradas a partir do processamento do sinal se mostraram coerentes.

Figura 12 – Temperatura média ao longo do ano.



Fonte: O Autor (2022).

Embora os dados tenham sido coletados ao longo de cinco anos, sua quantidade pode ser considerada pequena. Portanto o processamento feito nesta análise tem precisão limitada, embora os resultados possam ser considerados como ponto de partida para trabalhos futuros. A frequência de coleta de dados não é muito alta resumidamente por dois motivos: a velocidade com que a temperatura dos quadros varia não é muito alta, pois a formação de pontos quentes ocorre normalmente de forma lenta e gradual; e como os dados são coletados de forma manual, utilizando-se câmera termográfica a execução diária ou mesmo mensal da atividade se torna inviável em termos práticos.

A maioria dos algoritmos mostrados ao longo do Capítulo 3 funciona analisando tanto o sinal atual quanto seu histórico. No caso do sinal de temperatura geral dos quadros elétricos, há anomalias em todo o histórico de medições, pois após as termografias realizadas houve a

recomendação de manutenção preditiva pelo operador em ao menos um quadro elétrico em todos os casos.

A recomendação de manutenção preditiva era feita pelo operador sempre que algum quadro apresentava um desvio de temperatura ou padrão termográfico incomum. Para esta análise, são levados em conta a temperatura geral do quadro (se está dentro dos limites estabelecidos em tabela de referência), o ambiente em que o quadro se encontra (se é uma sala técnica climatizada, um corredor com pouca ventilação etc.), o tipo de carga que ele alimenta (se são cargas pequenas, como de iluminação, ou cargas maiores, como condicionadores de ar) e o estado de conservação do quadro (alguns quadros são originais da construção do prédio e têm mais de 50 anos, estando perto do fim de sua vida útil).

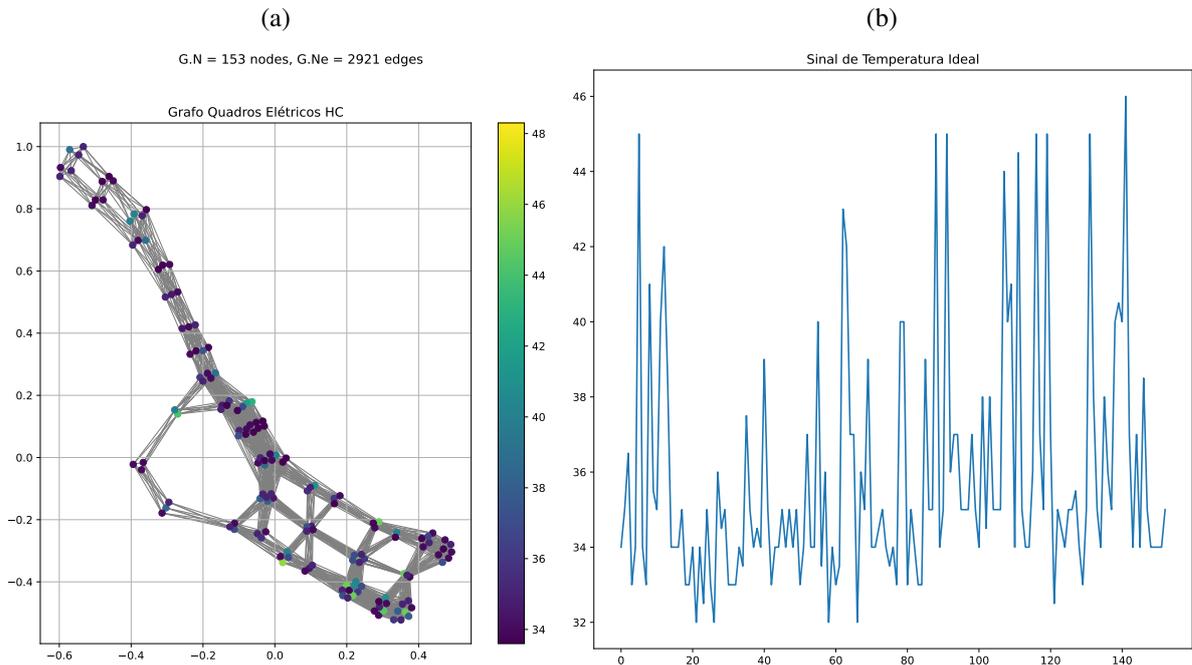
Os algoritmos que são utilizados a seguir dependem da obtenção de um histórico de sinal sem anomalias. Como os sinais reais não possuem esta característica, foi gerado artificialmente um sinal padrão com os valores de temperatura esperados para cada quadro. Este sinal foi gerado tendo como base a série histórica de temperatura dos quadros, o tipo de carga que estes quadros alimentam, o ambiente em que eles se encontram e seu estado de conservação. Nesta análise, também foi considerado que as leituras mais recentes têm maior impacto no valor esperado uma vez que o plano de manutenção preventiva e preditiva foi iniciado junto com as primeiras inspeções termográficas realizadas, fazendo com que o surgimento de novas anomalias diminuísse e se estabilizasse com o passar do tempo.

A Figura 13a mostra o sinal de temperatura plotado sobre a topologia citada anteriormente. A Figura 13b mostra o sinal de temperatura ideal criado utilizando-se os critérios citados anteriormente. Este sinal variou entre 32°C e 46°C.

Para avaliar se os algoritmos utilizados foram capazes de acertar ou não as predições dos quadros com anomalias, foram utilizadas as ordens de serviço de manutenção preditiva geradas na época em que cada termografia foi realizada. Observa-se que no início mais ordens de serviço eram geradas uma vez que ainda não havia plano de manutenção preventiva implantado no prédio, bem como a experiência do profissional responsável por desempenhar esta atividade ainda era pouca. Assim, há de se considerar que pode haver a indicação de falsos positivos por parte do operador.

Uma característica particular deste sinal também é que a indicação de uma anomalia não necessariamente significa que uma leitura foi feita de forma errada. Neste caso é mais provável que grandes variações indiquem o surgimento de um ponto quente ou outro tipo de problema de manutenção, do que indicar uma leitura errada. Portanto, na análise que apresentada nas seções seguintes, as anomalias encontradas são consideradas como um problema de manutenção do quadros, e não como um erro de leitura do sinal.

Figura 13 – Grafo com sinal ideal.



Fonte: O Autor (2022).

## 4.2 ANÁLISE DA GFT

Algumas vezes, ao aplicar a transformada de Fourier a um sinal e analisar a GFT isoladamente, é possível determinar se há anomalia no sinal. Ao proceder com a análise da transformada de Fourier dos sinais e desconsiderando a componente de frequência zero, foi verificado que alguns sinais possuíam algumas particularidades em sua GFT.

A Figura 14a mostra a GFT do sinal de janeiro de 2018 apresentou uma componente bastante destacada em relação às outras. Ao analisar o sinal original, é possível verificar que há uma leitura indicando uma medição de 102°C. Em relação à leitura anterior à de novembro de 2017, a variação de temperatura neste quadro foi de 45°C. Esta variação, por si só já, é um forte indício de que há uma anomalia neste nó e, provavelmente, esta é uma das razões para que a GFT do sinal tenha alguns componentes tão destacados em relação a outros.

A GFT do sinal de outubro de 2019 está mostrado na Figura 14b e também teve uma de suas componentes muito destacada em relação às demais. Neste caso, ao analisar o próprio sinal e compará-lo com os anteriores, também foi visto que há uma grande variação de temperatura em um dos componentes, da ordem de 34°C. Isto também ocorre com a GFT do sinal de julho de 2019, mostrado na Figura 14c em que houve uma variação de 31°C em uma das leituras.

As demais GFTs analisadas não mostraram grandes variações em suas componentes. Com isso fica evidente que, para o sinal que está sendo analisado, não é possível inferir muitas informações apenas observando a GFT. Os nós destacados nas Figuras 14a, 14b e 14c representam

os quadros que foram considerados como anomalias pelo operador na época em que a termografia foi realizada.

### 4.3 DETECÇÃO E LOCALIZAÇÃO DE ANOMALIAS

A técnica aplicada para detectar as anomalias no sinal é descrita na seção 3.2. Conforme mostrado, um filtro passa-alta de dois canais é aplicado ao sinal de entrada. Utilizando-se 29 e 30 é possível identificar se há alguma anomalia no sinal, caso o valor do resíduo máximo obtido seja maior que o valor do resíduo médio normalizado por um fator  $\gamma$ .

Ao longo desta seção, será considerado como um **acerto** os casos em que o algoritmo e o operador fizerem a mesma indicação de que um nó é anômalo. O **percentual de acerto** é a relação entre o número de acertos do algoritmo e o número de nós considerados como anomalia pelo operador na época em que o sinal foi coletado. Um **falso positivo** é considerado quando o algoritmo indica como anomalia um nó que não foi considerado pelo operador desta forma. O **percentual de falsos positivos** é a relação entre a quantidade de falsos positivos indicados pelo algoritmo e o total de anomalias indicadas pelo algoritmo. Um **verdadeiro positivo** é quando o algoritmo e o operador concordam, seja uma indicação de anomalia ou uma não-anomalia. E o **percentual de verdadeiro positivo** é o percentual total de verdadeiros positivos em relação ao total de quadros.

Foram feitos vários testes variando-se o valor de  $\gamma$  a fim de encontrar o fator ideal que deveria ser utilizado. Para isso, foram gerados 100 valores de  $\gamma$  uniformemente distribuídos entre 0 e 2. A Figura 15 mostra a relação entre os percentuais de acertos e falsos positivos e os valores de  $\gamma$ .

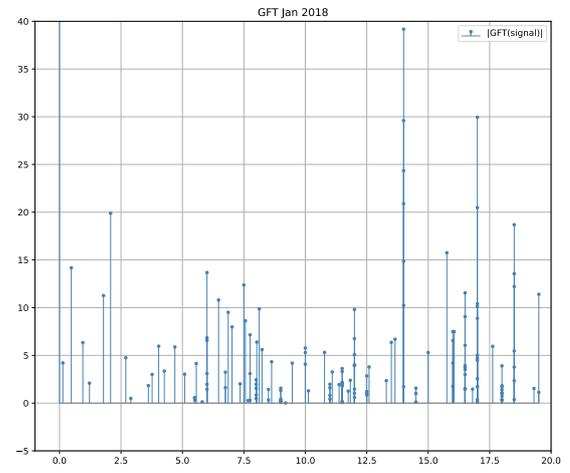
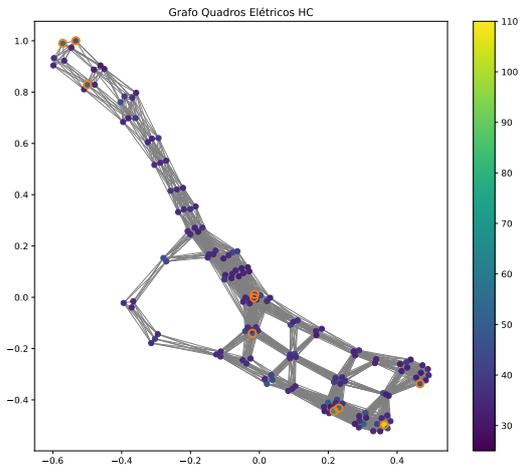
Ao analisar o gráfico mostrado na Figura 15 é possível ver que tanto o percentual de acertos quanto o percentual de falsos positivos diminuem a medida que o valor de  $\gamma$  aumenta. Isto se dá pois valores menores de  $\gamma$  fazem com que o limiar para indicar um quadro como anomalia diminua, fazendo com que mais quadros sejam indicados como anomalia. Como uma quantidade maior de quadros é indicada como anomalia pelo algoritmo, é de se esperar que mais acertos ocorram. Porém, isto leva a indicação de variações de temperatura comuns nos quadros como anomalias, uma vez que o limiar de indicação diminuiu. Observando os dois gráficos, é possível ver que a quantidade de falsos positivos se estabiliza na faixa dos 5% para os valores de  $\gamma \geq 1,4$ . Já o percentual de acertos continua decaindo de forma linear à medida que o valor de  $\gamma$  aumenta.

Considerando todos os valores de  $\gamma$  no intervalo de 0 a 2, a Figura 16 mostra o percentual de acerto médio para cada mês. Ao analisar o gráfico, é possível verificar que a média de acertos é maior nos sinais a partir do ano de 2020. Um dos fatores que podem explicar esta diferença é que os dados iniciais não contavam com nenhum histórico de manutenção, o que pode ter levado o operador a indicar quadros como anomalias. Com o passar do tempo, foi criado um histórico e,

Figura 14 – Alguns sinais sobre grafos e respectivas GFTs.

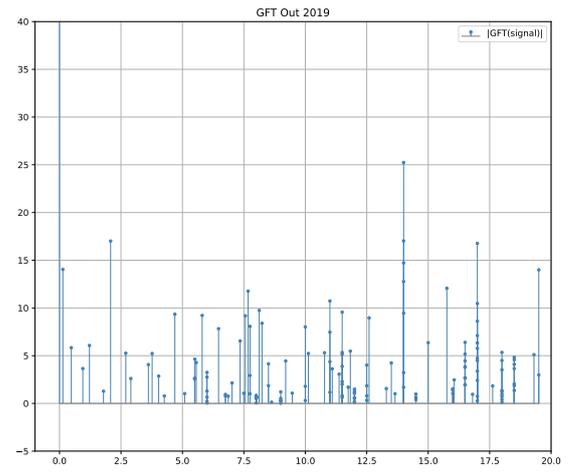
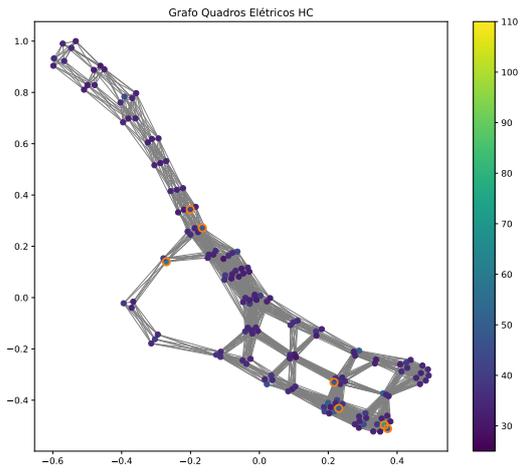
(a)

Grafo Quadros



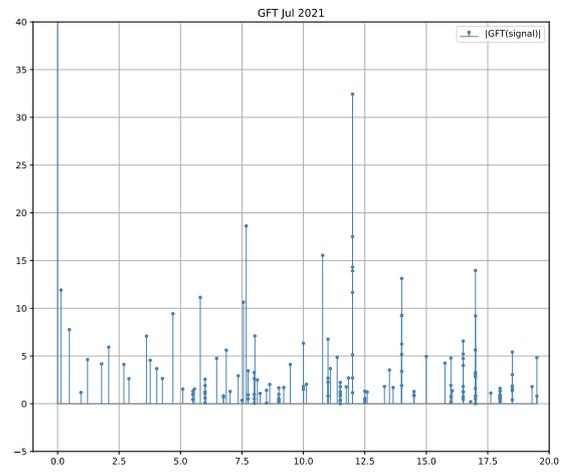
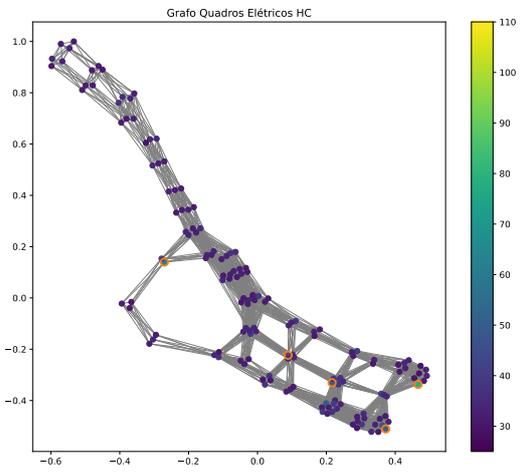
(b)

Grafo Quadros



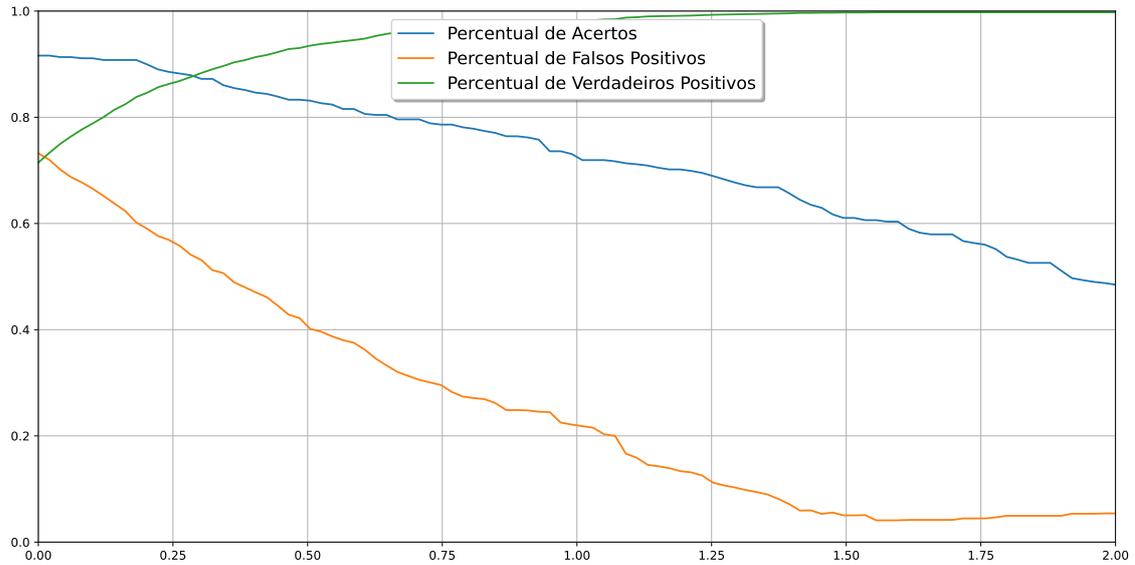
(c)

Grafo Quadros



Fonte: O Autor (2022).

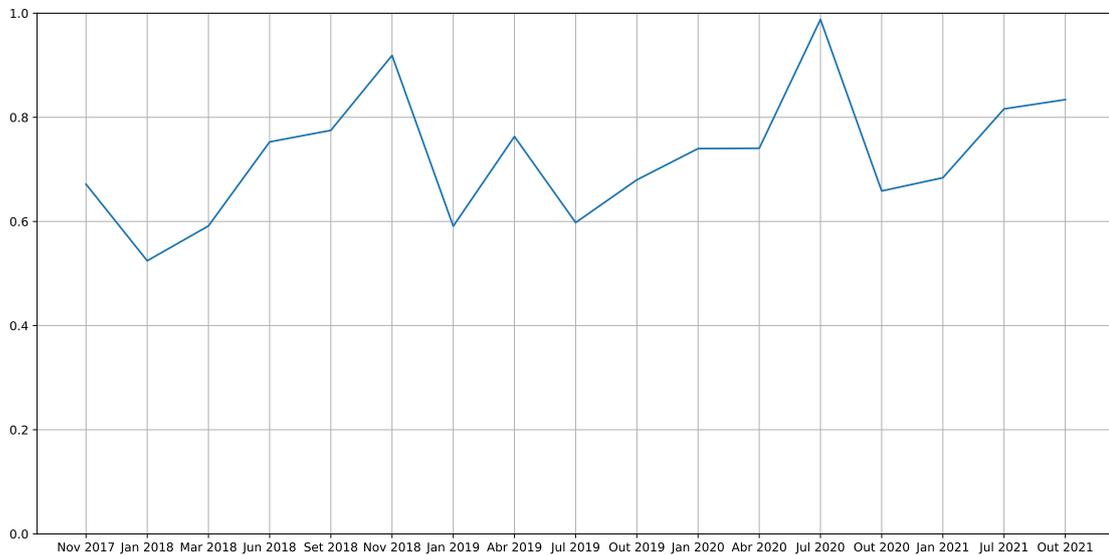
Figura 15 – Percentual de acerto de nós anomalias e falsos positivos x variação do valor de gama.



Fonte: O Autor (2022).

além disso, uma maior experiência do operador levou a critérios mais claros para indicação de quadros como anomalias, tornando o processo menos suscetível a erros.

Figura 16 – Percentual de acerto médio por sinal.



Fonte: O Autor (2022).

Para fins de análise de desempenho do algoritmo apresentado, o valor de  $\gamma$  foi escolhido tomando como referência o máximo percentual de acerto para um percentual de falsos positivos menor do que 8%. O valor de  $\gamma$  que satisfaz estas condições foi  $\gamma = 1,41$ . Todas as análises feitas a seguir tomaram como base este valor. Para este valor de gama, o percentual de acertos foi de aproximadamente 72%. O percentual de verdadeiros positivos para este valor de gama

é de aproximadamente 99,5%. Isto é, considerando a identificação de todas as amostras como anomalias ou não-anomalias.

Na Figura 16 é possível observar que o percentual de verdadeiros positivos sobe rapidamente a medida que o valor de gama aumenta. Neste contexto, isso se dá pela baixa quantidade de anomalias que existe no sinal no sinal em geral. Das 2601 amostras de temperatura, apenas 267 foram consideradas como anomalias (aproximadamente 10,3% das amostras).

Outro ponto a se destacar é que a quantidade de quadros indicados como falsos positivos se mostrou baixa em números absolutos, com média de 0,8 falsos positivos por sinal. Em apenas 7 dos 17 sinais analisados houve a indicação de falsos positivos. Em termos gerais, cerca de 7,44% dos quadros indicados como anomalias foram considerados como falsos positivos durante a análise do sinal. Isto contribui bastante para elevar o percentual de verdadeiros positivos.

Outra característica que pode ser observada foi que em alguns casos a variação de temperatura dos quadro indicados como anomalia pelo algoritmo em relação ao valor esperado de leitura era menor do que a variação de outros quadros que não foram indicados como anomalia. Isso mostra que apenas a variação da temperatura num determinado ponto não é determinante para que um quadro seja considerado como anomalia. É esperado que as arestas entre os quadros tenha influência nesta determinação.

## 5 CONSIDERAÇÕES FINAIS

Esta dissertação apresentou os fundamentos do processamento de sinais sobre grafos, tendo como foco o uso de ferramentas de GSP no cenário de redes de sensores e um estudo de caso no qual foram analisados dados provindos de medições periódicas de temperatura em quadros elétricos de um hospital. Os pontos abordados são sumarizados a seguir:

- Foram descritos os conceitos fundamentais de GSP, desde os conceitos básicos, nomenclaturas utilizadas, tipos de grafos, sinais aplicados sobre grafos, deslocamento destes sinais, técnicas de filtragem e filtros utilizados, e algumas aplicações de GSP em redes de sensores.
- Dentre as técnicas abordadas, está a detecção de anomalias em redes de sensores através da aplicação de um filtro passa-alta no sinal sobre o grafo, utilizando a transformada de Fourier em grafos (SANDRYHAILA; MOURA, 2014). Apesar da boa precisão em detectar nós com anomalias em suas medições, este tipo de processamento não contempla formas eficientes de detectar em qual nó, ou nós, está a anomalia, o que pode ser necessário em algumas aplicações.
- Outra técnica de processamento vista foi o processamento do sinal em um filtro de dois canais (MOU et al., 2020), no qual o histórico deste mesmo sinal passa por um dos canais com um filtro passa alta e o sinal que se deseja analisar é passado no outro canal. Então há a aplicação de um algoritmo para detectar e localizar os nós com valores suspeitos de serem anômalos, através da amplitude residual máxima e média da filtragem.
- Outra aplicação importante descrita é a utilização da própria modelagem do grafo para prever locais com maiores concentrações de poluentes no ar baseado no consumo de óleo de aquecimento em edifícios de grandes cidades (JAIN; MOURA; KONTOKOSTA, 2014). No exemplo abordado, é possível verificar que não apenas a própria emissão pode ser utilizada para determinar locais com níveis maiores ou menores de poluentes, mas também a influencia dos próprios arredores deve ser considerada.
- No estudo de caso apresentado, a técnica mostrada por (MOU et al., 2020) foi adaptada para detectar a formação de pontos quentes em quadros elétricos de um hospital. Esta abordagem utilizou as mesmas premissas de processamento dos sinais mostradas pelos autores. Entretanto, a interpretação dos resultados obtidos foi adaptada para o cenário em que foi aplicado. Como alterações na temperatura de quadros elétricos podem ocorrer de forma relativamente rápida, sem necessariamente indicar mal funcionamento dos sensores de temperatura, neste estudo de caso as anomalias detectadas foram consideradas como a formação de pontos quentes no quadro elétrico.

- Para realizar o estudo de caso, foi montado um grafo a partir de dados de temperatura coletados pelo próprio autor desta dissertação. A topologia do grafo foi montada a partir de critérios definidos com base na experiência e conhecimento de caso do autor. Os resultados obtidos com o processamento dos dados foi comparado as dados reais de manutenção dos quadros, o que mostrou que o algoritmo foi capaz de identificar com um baixo número de falsos positivos em quais quadros havia a formação de pontos quentes.
- Apesar da pequena quantidade de dados analisados, cerca de 2.601 medições, os resultados são considerados satisfatórios por abrirem perspectiva para trabalhos futuros. Além disto, os resultados obtidos corroboram com o fato que técnicas de GSP podem ser aplicadas para detectar diferentes tipos de fenômenos, mesmo que inicialmente estes não sejam identificados como um tipo de sinal clássico. Além de serem capazes de extrair informações que não estão disponíveis ao analisar apenas os dados brutos.

## 5.1 TRABALHOS FUTUROS

A partir do estudo realizado neste trabalho, é possível elencar os possíveis desdobramentos:

- Devido a baixa quantidade de dados analisados, novos estudos com uma maior base de dados podem ser feitos para confirmar os resultados obtidos neste trabalho, uma vez que a quantidade de amostras reduzidas pode levar a vícios ocultos no processamento que seriam mais facilmente identificados com um maior volume.
- A partir deste trabalho, é possível realizar a instalação de uma rede de sensores nestes quadros que possam coletar informações com maior frequência e monitorá-los. Neste sentido, a adaptação das técnicas apresentadas em (JAIN; MOURA; KONTOKOSTA, 2014) para o cenário abordado no estudo de caso pode indicar os melhores pontos para serem instalados os sensores, de forma a otimizar a coleta de dados ao mesmo tempo que reduza os custos de instalação.
- Com um maior volume de dados, outras formas de aproximação do problema mostrado também pode ser feita, como a aplicação de algoritmos para detecção de bordas e análise de entropia como as mostradas em (NIGRIS et al., 2017). Novos estudos também podem verificar se é possível identificar a formação de pontos quentes nos próprios circuitos que fazem a interligação dos quadros, uma vez que instalações elétricas antigas são mais suscetíveis a falhas de isolamento.
- Outros parâmetros dos quadros elétricos também podem ser analisados para a comparação com os resultados do processamento baseado na temperatura, como as medidas das correntes e das tensões. Como a formação de pontos quentes aumenta a resistência elétrica

dos elementos, o processamento de grandezas elétricas pode trazer alguma correlação entre a formação dos pontos quentes e alterações nestas grandezas.

## REFERÊNCIAS

- AKELLA, P. N. et al. Creation and maintenance of social relationship network graphs. **Google Patents**, US Patent 7,539,697, maio 26 2009. Citado na página 11.
- ANIRUDH, R.; THIAGARAJAN, J. J. **Bootstrapping Graph Convolutional Neural Networks for Autism Spectrum Disorder Classification**. arXiv, 2017. Disponível em: <<https://arxiv.org/abs/1704.07487>>. Citado na página 10.
- AZAM, A. G.; RIAHI-ZANJANI, B.; BALALI-MOOD, M. Effects of air pollution on human health and practical measures for prevention in Iran. **Journal of Research in Medical Sciences**, v. 21, 09 2016. Citado na página 33.
- BLONDEL, V. D. et al. Fast unfolding of communities in large networks. **Journal of Statistical Mechanics: Theory and Experiment**, v. 2008, n. 10, p. P10008, oct 2008. Disponível em: <<https://dx.doi.org/10.1088/1742-5468/2008/10/P10008>>. Citado na página 12.
- CHEN, S. **Data Science with Graphs: A Signal Processing Perspective**. Tese (Doutorado) — Carnegie Mellon University, 2016. Url<http://repository.cmu.edu/dissertations/724>. Citado na página 12.
- CHEUNG, M. et al. Graph signal processing and deep learning: Convolution, pooling, and topology. **IEEE Signal Processing Magazine**, Institute of Electrical and Electronics Engineers (IEEE), v. 37, n. 6, p. 139–149, nov 2020. Disponível em: <<https://doi.org/10.1109%2Fmosp.2020.3014594>>. Citado na página 13.
- DEFFERRARD, M. et al. **PyGSP: Graph Signal Processing in Python**. Disponível em: <<https://github.com/epfl-lts2/pygsp/>>. Citado na página 14.
- DERI, J. A.; MOURA, J. M. F. Spectral projector-based graph Fourier transforms. **IEEE Journal of Selected Topics in Signal Processing**, v. 11, n. 6, p. 785–795, 2017. Citado na página 24.
- DONG, X. et al. Graph signal processing for machine learning: A review and new perspectives. **CoRR**, abs/2007.16061, 2020. Disponível em: <<https://arxiv.org/abs/2007.16061>>. Citado na página 13.
- DRAYER, E.; ROUTTENBERG, T. Detection of false data injection attacks in power systems with graph Fourier transform. In: **2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)**. [S.l.: s.n.], 2018. p. 890–894. Citado na página 29.
- EGILMEZ, H. E.; ORTEGA, A. Spectral anomaly detection using graph-based filtering for wireless sensor networks. In: IEEE. **2014 IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)**. [S.l.], 2014. p. 1085–1089. Citado na página 12.
- FEOFILOFF, P.; KOHAYAKAWA, Y.; WAKABAYASHI, Y. Uma introdução sucinta à teoria dos grafos. 2011. Url<https://www.ime.usp.br/pf/teoriadosgrafos/>. Citado na página 17.
- FONINI, P. A. M. **A Didatic Introduction to Graph Signal Processing Techniques and Applications**. Dissertação (Mestrado) — Universidade Federal do Rio de Janeiro, Rio de Janeiro, Mar 2019. Citado na página 22.

- GOLDSBERRY, L. et al. Brain signal analytics from graph signal processing perspective. **IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**, p. 851–855, 2017. Citado na página 11.
- GRIPON, V.; ORTEGA, A.; GIRAULT, B. An inside look at deep neural networks using graph signal processing. In: IEEE. **2018 Inf. Theory and Applications Workshop (ITA)**. [S.l.], 2018. p. 1–9. Citado na página 13.
- HAGBERG, A. A.; SCHULT, D. A.; SWART, P. J. Exploring network structure, dynamics, and function using networkx. In: VAROQUAUX, G.; VAUGHT, T.; MILLMAN, J. (Ed.). **Proceedings of the 7th Python in Science Conference**. Pasadena, CA USA: [s.n.], 2008. p. 11 – 15. Disponível em: <[http://conference.scipy.org/proceedings/SciPy2008/paper\\_2/](http://conference.scipy.org/proceedings/SciPy2008/paper_2/)>. Citado na página 14.
- HOSSEINALIPOUR, S. et al. Infection analysis on irregular networks through graph signal processing. **IEEE Transactions on Network Science and Engineering**, v. 7, n. 3, p. 1939–1952, 2020. Citado na página 12.
- HUANG, W. et al. A graph signal processing perspective on functional brain imaging. **Proceedings of the IEEE**, v. 106, n. 5, p. 868–885, 2018. Citado na página 10.
- HUANG, W. et al. Graph frequency analysis of brain signals. **IEEE J. Sel. Topics Signal Process.**, IEEE, v. 10, n. 7, p. 1189–1203, 2016. Citado na página 11.
- JABLONSKI, I. Graph signal processing in applications to sensor networks, smart grids, and smart cities. **IEEE Sensors Journal**, v. 17, n. 23, p. 7659–7666, 2017. Citado 2 vezes nas páginas 10 e 12.
- JAIN, R. K.; MOURA, J. M.; KONTOKOSTA, C. E. Big data + big cities: Graph signals of urban air pollution [exploratory sp]. **IEEE Signal Processing Magazine**, v. 31, n. 5, p. 130–136, 2014. Citado 6 vezes nas páginas 12, 33, 34, 35, 46 e 47.
- LIMA, J. B.; DE OLIVEIRA NETO, J. R. A graph signal processing approach to Fourier-like number-theoretic transforms. **Digital Signal Processing**, v. 131, p. 103761, 2022. ISSN 1051-2004. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1051200422003785>>. Citado na página 23.
- MAO, F. et al. Low-cost environmental sensor networks: Recent advances and future directions. **Frontiers in Earth Science**, v. 7, p. 221, 09 2019. Citado na página 29.
- MASOUMI, M.; REZAEI, M.; HAMZA, A. B. Shape analysis of carpal bones using spectral graph wavelets. In: **Vertex-Frequency Analysis of Graph Signals**. [S.l.]: Springer, 2019. p. 419–436. Citado na página 10.
- MEI, J.; MOURA, J. M. F. Signal processing on graphs: Causal modeling of unstructured data. **IEEE Trans. Signal Process.**, IEEE, 2016. Citado na página 37.
- MOHAN, D. M. et al. Wavelets on graphs with application to transportation networks. In: IEEE. **17th International IEEE Conference on Intelligent Transportation Systems (ITSC)**. [S.l.], 2014. p. 1707–1712. Citado na página 12.
- MOU, M. et al. Malfunction detection and localization algorithm for wireless sensor network. **Circuits, Systems, and Signal Processing**, n. 40, p. 501–509, Jul 2020. Citado 3 vezes nas páginas 31, 32 e 46.

NIGRIS, S. de et al. Fractional graph-based semi-supervised learning. In: . [S.l.: s.n.], 2017. p. 356–360. Citado na página 47.

OPPENHEIM, A. V.; SCHAFER, R. W. **Discrete-time Signal Processing**. [S.l.]: Pearson, 2013. ISBN 9781292025728. Citado na página 24.

OPPENHEIM, A. V.; WILLSKY, A. S.; NAWAB, S. H. **Signals and Systems**. Prentice Hall, 1997. (Prentice-Hall signal processing series). ISBN 9780138147570. Disponível em: <<https://books.google.com.br/books?id=LwQqAQAAMAAJ>>. Citado na página 24.

ORTEGA, A. et al. Graph signal processing: Overview, challenges, and applications. **Proceedings of the IEEE**, v. 106, n. 5, p. 808–828, 2018. Citado 3 vezes nas páginas 10, 11 e 12.

PEROZZI, B.; AL-RFOU, R.; SKIENA, S. DeepWalk. In: **Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining**. ACM, 2014. Disponível em: <<https://doi.org/10.1145%2F2623330.2623732>>. Citado na página 10.

PIRAYRE, A. et al. Brane cut: biologically-related a priori network enhancement with graph cuts for gene regulatory network inference. **BMC Bioinformatics**, BioMed Central, v. 16, n. 1, p. 368, 2015. Citado na página 11.

PUSCHEL, M.; MOURA, J. M. F. Algebraic signal processing theory: Foundation and 1-D time. **IEEE Transactions on Signal Processing**, v. 56, n. 8, p. 3572–3585, 2008. Citado na página 23.

RABIEI, H. et al. Estimating the complexity of the cerebral cortex folding with a local shape spectral analysis. In: **Vertex-Frequency Analysis of Graph Signals**. [S.l.]: Springer, 2019. p. 437–458. Citado na página 11.

RAMAKRISHNA, R.; SCAGLIONE, A. Grid-graph signal processing (grid-GSP): A graph signal processing framework for the power grid. **IEEE Transactions on Signal Processing**, Institute of Electrical and Electronics Engineers (IEEE), v. 69, p. 2725–2739, 2021. Disponível em: <<https://doi.org/10.1109%2Ftsp.2021.3075145>>. Citado na página 11.

RIBEIRO, G.; LIMA, J. Graph signal processing in a nutshell. **Journal of Communication and Information Systems**, v. 33, n. 1, Jul. 2018. Disponível em: <<https://jcis.sbrt.org.br/jcis/article/view/563>>. Citado 6 vezes nas páginas 17, 18, 19, 21, 22 e 25.

SAKIYAMA, A. et al. Efficient sensor position selection using graph signal sampling theory. In: **IEEE. 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. [S.l.], 2016. p. 6225–6229. Citado na página 12.

SANDRYHAILA, A.; MOURA, J. M. Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure. **IEEE Signal Processing Magazine**, v. 31, n. 5, p. 80–90, 2014. Citado na página 10.

SANDRYHAILA, A.; MOURA, J. M. F. Discrete signal processing on graphs. **IEEE Trans. Signal Process.**, IEEE, v. 61, n. 7, p. 1644–1656, 2013. Citado na página 24.

SANDRYHAILA, A.; MOURA, J. M. F. Discrete signal processing on graphs: Graph Fourier transform. In: **2013 IEEE International Conference on Acoustics, Speech and Signal Processing**. [S.l.: s.n.], 2013. p. 6167–6170. Citado na página 24.

SANDRYHAILA, A.; MOURA, J. M. F. Discrete signal processing on graphs: Frequency analysis. **IEEE Transactions on Signal Processing**, v. 62, n. 12, p. 3042–3054, 2014. Citado 6 vezes nas páginas 12, 25, 26, 29, 30 e 46.

SARDELLITTI, S.; BARBAROSSA, S.; LORENZO, P. D. Graph topology inference based on transform learning. In: **2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)**. IEEE, 2016. Disponível em: <<https://doi.org/10.1109/globalsip.2016.7905863>>. Citado na página 37.

SARDELLITTI, S.; BARBAROSSA, S.; LORENZO, P. D. On the graph Fourier transform for directed graphs. **IEEE J. Sel. Topics Signal Process.**, IEEE, 2017. Citado na página 25.

SHUMAN, D. I. et al. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. **IEEE Signal Process. Mag.**, IEEE, v. 30, n. 3, p. 83–98, 2013. Citado 3 vezes nas páginas 11, 23 e 25.

SHUMAN, D. I.; RICAUD, B.; VANDERGHEYNST, P. Vertex-frequency analysis on graphs. **Applied and Computational Harmonic Analysis**, Elsevier, v. 40, n. 2, p. 260–291, 2016. Citado na página 25.

SPELTA, M.; MARTINS, W. Online temperature estimation using graph signals. In: . [S.l.: s.n.], 2018. Citado na página 12.

STANKOVIC, L.; DAKOVIC, M.; SEJDIC, E. Introduction to graph signal processing. In: \_\_\_\_\_. [S.l.: s.n.], 2019. p. 3–108. ISBN 978-3-030-03573-0. Citado 5 vezes nas páginas 11, 20, 23, 26 e 27.

STANKOVIĆ, L. et al. Vertex-frequency graph signal processing: A comprehensive review. **Digital Signal Processing**, v. 107, p. 102802, 2020. ISSN 1051-2004. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1051200420301470>>. Citado na página 23.

THULASIRAMAN, K.; SWAMY, M. N. S. **Graphs: Theory and Algorithms**. [S.l.]: Wiley Online Library, 1992. 97–125 p. Citado na página 17.

ZHANG, J.; MOURA, J. M. F. Diffusion in social networks as sis epidemics: Beyond full mixing and complete graphs. **IEEE Journal of Selected Topics in Signal Processing**, v. 8, n. 4, p. 537–551, 2014. Citado na página 10.

## APÊNDICE A – CÓDIGOS UTILIZADOS PELO AUTOR EM PYTHON

### DEFINIÇÃO DE FUNÇÕES, CONSTANTES E CARREGAMENTO DE DADOS EXTERNOS

```
##### FUNÇÕES #####
def comparar_listas(lista, gabarito):
    acerto = 0
    falso_pos = 0
    for i in range(len(lista)):
        if lista[i] in gabarito:
            acerto += 1
        else:
            falso_pos += 1
    per_falso_pos = falso_pos / len(lista)
    perc_acerto = acerto / len(gabarito)
    return perc_acerto, acerto, falso_pos, per_falso_pos

##### CONSTANTES #####

baseFiles = 'Arquivos_PDF/'      # caminho dos arquivos pdf
colorbar = 1      # colocar colorbar nos grafos
vertex_size = 40      # tamanho dos vértices
salvar_graficos = 1      # = 1 > plota grafo e salva
taps = 8      # grau do polinômio utilizado
ruído = 0.2      # porcentagem de ruído adicionado
nPassos = 100      # num de testes para achar ordem frac
lim_a = 6      # Valor máximo da ordem fracionária procuradas
Faixa = 13      # Valor Empírico
Faixa2 = 4.51      # Meta das frequências serão filtradas
format_string = '{:.4f}'      # formato de impressao de num

##### CARREGAR DADOS EXTERNOS #####

AdjQ = np.loadtxt('AdjQ.txt', delimiter = ',')
sTempQ = np.loadtxt('sTempQ.txt', delimiter = ',')
coordenadas = np.loadtxt('CoordQ2D.txt', delimiter = ',')
coordenadas3D = np.loadtxt('CoordQ.txt', delimiter = ',')
sTempIdeal = np.loadtxt('sTempIdeal.txt', delimiter = ',')
x_ax = np.loadtxt('tagMeses.txt', delimiter = ',')
```

```
AdjQ2 = np.loadtxt('AdjQ2.txt', delimiter = ', ')
sTempQ2 = np.loadtxt('sTempQ2.txt', delimiter = ', ')
```

## GERAR GRAFOS E FILTROS

```
##### INICIALIZAR GRAFO #####
```

```
print("Inicializando grafos e plotando os resultados")
G_Lap = graphs.Graph(W=AdjQ, coords = coordenadas)
G_Lap.set_coordinates(kind='spring', seed = 40)
```

```
##### DEFINIR FILTRO PASSA ALTA 1 #####
```

```
print("Gerando filtro passa alta 1 do tipo  $x/(1+x)$ ")
high_pass_1 = filters.Filter(G_Lap, lambda x: x/(1+x))
print("Inicializando testes com ruído em apenas um dos nós.")
```

```
##### GABARITO DE MANUTENÇÃO DOS QUADROS #####
```

```
# ESTE É O GABARITO COM OS QUADROS QUE FORAM SINALIZADOS
# COMO COM PONTO QUENTE PELO OPERADOR NA ÉPOCA.
# É UMA LISTA DE LISTAS.
```

```
with open('gabarito.txt') as f:
    lines = f.readlines()
gabarito = []
for i in range(len(lines)):
    str = lines[i]
    arr = str.split(',')
    integer_map = map(int, arr)
    integer_list = list(integer_map)
    gabarito.append(integer_list)
```

```
# A VARIÁVEL DADOS_G VAI SALVAR OS VALORES DE GAMA,
# MÉDIA PERCENTUAL DE ACERTOS E TOTAL DE FALSOS
# POSITIVOS A VARIÁVEL DADOS_B SERÁ UTILIZADA PARA
# GUARDAR OS DADOS BRUTOS DA PREVISÃO DO ALGORITMO
```

```
dados_g = []
dados_b = []
```

```
dados_pg = []
dados_pb = []
```

## APLICAR ALGORITMO DE DETECÇÃO E LOCALIZAÇÃO DE ANOMALIAS

```
# ITERAR ALGORITMO PARA DESCOBRIR O MELHOR VALOR DE GAMA

val = np.linspace(0, 2, 100)

for gama in val:
    print('Analisando sinal para gama = {}'.format(gama))
    nodes_highlight = []
    dados_temp = []
    dados_ptemp = []
    for s in range(sTempQ.shape[1]):
        cont = 1
        print('Analisando sinal de {}'.format(x_ax[s]))
        sTempQN = [0]*(sTempQ.shape[0])
        for k in range(0,sTempQ.shape[0]):
            sTempQN[k] = sTempQ[:,s][k]
        sTempQN = np.array(sTempQN)

        sTempQI = [0]*(sTempIdeal.shape[0])
        for k in range(0,sTempIdeal.shape[0]):
            sTempQI[k] = sTempIdeal[k]
        sTempQI = np.array(sTempQI)

    # FILTRAR AMBOS OS SINAIS NO DOMÍNIO DOS NÓS

    tSignalFilt = high_pass_1.filter(sTempQN)
    tSignalFilt_1 = high_pass_1.filter(sTempQI)
    tSignalFilt_2 = high_pass_1.filter(sTempQI)
    tSignalFilt_3 = high_pass_1.filter(sTempQI)

    # CALCULAR AMPLITUDE RESIDUAL MÁXIMA

    k = 3 # quantidade de sinais no histórico
    yt = tSignalFilt # sinal suspeito filtrado
```

```

yt_1 = tSignalFilt_1      # sinal esperado
yt_2 = tSignalFilt_1
yt_3 = tSignalFilt_1

yt_h = np.array([yt_1, yt_2, yt_3])
tsum = np.add(0, yt_h.sum(axis=0))
ar_max_t = yt - (1/k * tsum)
ar_max = LA.norm(ar_max_t, np.inf)

# CALCULAR AMPLITUDE RESIDUAL MÉDIA

tpsum = LA.norm(yt_1, np.inf)
ar_med = (2/(k*(k-1)))*tpsum
if ar_max > gama*ar_med:
    zt = yt - (1/k * tsum)
    nodes_highlight_temp = np.argwhere
    (zt > ar_med*gama)
    nodes_highlight.append
    (nodes_highlight_temp.flatten().tolist())

for m in range(len(nodes_highlight)):
    temp = comparar_listas(nodes_highlight[m], gabarito[m])
    dados_temp.append(temp)
dados_temp = np.array(dados_temp)
m_temp = np.mean(dados_temp, axis=0)
s_temp = np.sum(dados_temp, axis=0)
dados_g.append([gama, m_temp[0], s_temp[2], m_temp[3]])
dados_b.append([gama, dados_temp])

# FAZENDO ANÁLISE APENAS NOS ÚLTIMOS DOIS ANOS DO SINAL

for p in range(len(nodes_highlight)-7, len(nodes_highlight)):
    temp = comparar_listas(nodes_highlight[p], gabarito[p])
    dados_ptemp.append(temp)
dados_ptemp = np.array(dados_ptemp)
m_temp = np.mean(dados_ptemp, axis=0)
s_temp = np.sum(dados_ptemp, axis=0)
dados_pg.append([gama, m_temp[0], s_temp[2], m_temp[3]])
dados_pb.append([gama, dados_ptemp])

```

```

dados_g = np.array(dados_g)
dados_pg = np.array(dados_pg)
# A VARIÁVEL DADOS CONTÉM, NESTA ORDEM:
# GAMA, PERC ACERTO, TOTAL DE FALSOS POSITIVOS,
# PERC FALSOS POSITIVOS

```

## PLOTAGEM DOS RESULTADOS OBTIDOS

```
##### PLOTAR GRÁFICO DE PRECISÃO X FALSOS POSITIVOS #####
```

```

t = dados_g[:,0]
data1 = dados_g[:,1]
data2 = dados_g[:,3]
fig, ax1 = plt.subplots(figsize=(16,8))
ax1.plot(t, data1, label='Percentual de Acertos')
ax1.plot(t, data2, label='Percentual de Falsos Positivos')
legend = ax1.legend(loc='upper center', shadow=True, fontsize
='x-large')
plt.grid()
plt.xlim([0,2])
plt.ylim([0,1])
if salvar_graficos:
    plt.savefig('gama_test.pdf', dpi=300,
    bbox_inches='tight')

```

```
##### PLOTAR GRÁFICO PRECISÃO POR MÊS (TODOS OS MESES) #####
```

```

fig, ax2 = plt.subplots(figsize=(16,8))
t = x_ax
dados_parc = []
for n in range(len(dados_b)):
    datat = np.array(dados_b[n][1])
    data3 = datat[:,0]
    ax2.plot(t, data3)
    dados_parc.append(data3)
plt.grid()

```

```

plt.ylim([0,1])
if salvar_graficos:
    plt.savefig('teste_mes_geral.pdf', dpi=300, bbox_inches
                ='tight')

##### PLOTAR GRÁFICO DE MÉDIA DE PRECISÃO POR MÊS #####

dados_parc = np.array(dados_parc)
fig, ax3 = plt.subplots(figsize=(16,8))
t = x_ax
data4 = np.mean(dados_parc, axis=0)
ax3.plot(t, data4)
plt.grid()
plt.ylim([0,1])
if salvar_graficos:
    plt.savefig('teste_mes_media.pdf', dpi=300,
                bbox_inches='tight')

#### PLOTAR GRÁFICO DE PRECISÃO POR MÊS DOS ÚLTIMOS
DOIS ANOS ####

t = dados_pg[:,0]
data5 = dados_pg[:,1]
data6 = dados_pg[:,3]
fig, ax4 = plt.subplots(figsize=(16,8))
ax4.plot(t, data5, label='Percentual de Acertos')
ax4.plot(t, data6, label='Percentual de Falsos Positivos')
legend = ax4.legend(loc='upper center', shadow=True, fontsize
                    ='x-large')
plt.grid()
plt.xlim([0,2])
plt.ylim([0,1])
if salvar_graficos:
    plt.savefig('gama_test_2_years.pdf', dpi=300,
                bbox_inches='tight')

plt.close('all')

fig, ax8 = plt.subplots(1, figsize = (11,9), dpi=300)

```

```

_ = ax8.spy(G_Lap.W, markersize=1)
if salvar_graficos:
    plt.savefig('matriz_de_adjacencia.pdf', dpi=300,
                bbox_inches='tight')
fig, ax9 = plt.subplots(1, figsize = (11,9), dpi=300)
G_Lap.plot(show_edges = True, vertex_size = 30)
if salvar_graficos:
    plt.savefig('grafo_sem_sinal.pdf', dpi=300,
                bbox_inches='tight')

##### PLOTAR GRAFO COM SINAL IDEAL #####

fig, ax5 = plt.subplots(1, figsize = (11,9), dpi=300)
plotting.plot_signal(G_Lap, sTempIdeal, ax=ax5, plot_name
='Grafo Quadros Elétricos HC', vertex_size=35)
plt.suptitle('G.N = {} nodes, G.Ne = {} edges'.format(G_Lap.N,
G_Lap.Ne))
plt.grid()
if salvar_graficos:
    plt.savefig('grafo_sinal_ideal.pdf', dpi=300,
                bbox_inches='tight')

fig, ax6 = plt.subplots()
ax6.plot(sTempIdeal, label = 'Sinal de Temperatura Ideal')
ax6.set_title('Sinal de Temperatura Ideal')
if salvar_graficos:
    plt.savefig('sinal_temperatura_ideal.pdf', dpi=300,
                bbox_inches='tight')

##### PLOTAR SAZONALIDADE #####
med_saz = np.mean(sTempQ, axis=0)
fig, ax7 = plt.subplots(1, figsize = (23,9), dpi=300)
plt.plot(x_ax, med_saz)
plt.title('Média de Temperatura ao Longo do Tempo')
if salvar_graficos:
    plt.savefig('sazonalidade_temperatura_media.pdf',
                dpi=300, bbox_inches='tight')

```

```

##### IMPRIMIR PLOTAGEM #####
for i in range(sTempQ.shape[1]):
    nd_list = nodes_highlight[i]
    fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(23,9),
    dpi=300)
    fig.suptitle('Grafo + GFT')
    plotting.plot_signal(G_Lap, sTempQ[:,i], ax=ax1,
    vertex_size=35, limits=[25,110], highlight=nd_list,
    plot_name ='Grafo Quadros Elétricos HC')
    plt.suptitle('G.N = {} nodes, G.Ne = {} edges'.format
    (G_Lap.N, G_Lap.Ne))
    plt.suptitle('Grafo Quadros')
    plt.grid()

    sTemperatura = sTempQ[:,i]
    G_Lap.compute_fourier_basis()
    tSignalQ = G_Lap.gft(sTemperatura)

    (markers, stemlines, baseline) = ax2.stem(G_Lap.e.real,
    np.abs(tSignalQ), label=' |GFT(signal)| ')
    plt.setp(markers, marker='o', markersize=2, markeredgecolor
    ="steelblue", markeredgewidth=2)
    plt.setp(stemlines, linestyle="-", color="steelblue",
    linewidth=0.3 )
    plt.setp(baseline, linestyle="-", color="grey", linewidth=1)
    plt.title('GFT ' + x_ax[i])
    plt.xlim([-1, 20])
    plt.ylim([-5, 40])
    ax2.legend(loc='upper right')
    if salvar_graficos:
        plt.savefig(baseFiles + 'grafo+gft_{}.pdf'
        .format(x_ax[i]), dpi=300, bbox_inches='tight')

```