

UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO

BACHARELADO EM CIÊNCIA DE DADOS

TRABALHO DE CONCLUSÃO DE CURSO (TCC)

CINTIA IZUMI SHINODA

DANIEL REGINO MONÇÃO

FERNANDO MIGUEL ESCRIBANO MARTINEZ

PEDRO JOSÉ DE ASSIS BARBEITO CAMARA

ROGÉRIO GONÇALVES DA SILVA

THIAGO AKIRA FERREIRA

WILLY PAULINO DE OLIVEIRA GOMES

**Vulnerabilidade e resiliência da rede de transporte público de São Paulo: uma análise
topológica baseada em teoria dos grafos**

São Paulo – SP

2026

UNIVERSIDADE VIRTUAL DO ESTADO DE SÃO PAULO

CINTIA IZUMI SHINODA

DANIEL REGINO MONÇÃO

FERNANDO MIGUEL ESCRIBANO MARTINEZ

PEDRO JOSÉ DE ASSIS BARBEITO CAMARA

ROGÉRIO GONÇALVES DA SILVA

THIAGO AKIRA FERREIRA

WILLY PAULINO DE OLIVEIRA GOMES

Vulnerabilidade e resiliência da rede de transporte público de São Paulo: uma análise topológica baseada em teoria dos grafos

Trabalho de Conclusão de Curso apresentado ao curso de Ciência de Dados da Universidade Virtual do Estado de São Paulo (UNIVESP), como requisito parcial para obtenção do título de Bacharel em Ciência de Dados.

Orientador: Me. Anderson Nogueira Cotrim.

São Paulo – SP

2026

CINTIA IZUMI SHINODA

DANIEL REGINO MONÇÃO

FERNANDO MIGUEL ESCRIBANO MARTINEZ

PEDRO JOSÉ DE ASSIS BARBEITO CAMARA

ROGÉRIO GONÇALVES DA SILVA

THIAGO AKIRA FERREIRA

WILLY PAULINO DE OLIVEIRA GOMES

Vulnerabilidade e resiliência da rede de transporte público de São Paulo: uma análise topológica baseada em teoria dos grafos

Trabalho de Conclusão de Curso apresentado ao curso de Ciência de Dados da Universidade Virtual do Estado de São Paulo (UNIVESP), como requisito parcial para obtenção do título de Bacharel em Ciência de Dados.

Aprovado em: _____

Conceito: _____

BANCA EXAMINADORA

Me. Anderson Nogueira Cotrim
Orientador

Profa. Dra. Elia Yathie Matsumoto
Banca

RESUMO

A rede de transporte público da Região Metropolitana de São Paulo transporta cotidianamente milhões de passageiros e articula a mobilidade de aproximadamente 22 milhões de habitantes. Este trabalho aplica a Teoria dos Grafos à análise topológica do subsistema operado pela SPTrans, com o objetivo de identificar paradas e estações topologicamente críticas e avaliar a resiliência da rede sob diferentes regimes de perturbação. A partir do *feed* GTFS coletado em março de 2026, o sistema foi modelado como grafo simples, não-direcionado e não-ponderado em representação *L-space*, com integração intermodal das sub-redes de Metrô e CPTM por meio de arestas de transferência com raio de 400 metros. A análise de vulnerabilidade combinou três métricas de centralidade: *degree*, *betweenness* e *closeness*, e identificou a Estação Brás como nó articulador da rede, concentrando aproximadamente 38% dos caminhos mínimos com apenas sete vizinhos diretos. A análise de resiliência simulou cinco cenários de remoção sequencial: quatro ataques direcionados e falha aleatória com 30 repetições, e confirmou empiricamente o paradigma de Albert, Jeong e Barabási (2000): a rede apresenta elevada tolerância a falhas distribuídas, com remoção aleatória de 20% dos nós preservando 62% do componente principal, e fragilidade aguda a ataques concentrados em nós de alto grau, com remoção dos nós por *degree centrality* reduzindo o componente principal à metade. A análise da integração intermodal identificou ainda um limite institucional da SPTrans: 48 estações da CPTM em ramais que se estendem a municípios da Região Metropolitana permanecem desconectadas da malha de ônibus, refletindo a fronteira de atribuição operacional. Os achados sustentam recomendações práticas para a priorização de planos de contingência e para a articulação metropolitana entre operadoras.

Palavras-chave: Teoria dos Grafos; *General Transit Feed Specification* (GTFS); Transporte Público; Vulnerabilidade de Redes; Resiliência de Redes.

LISTA DE ILUSTRAÇÕES

Figura 1 — Modelo Entidade-Relacionamento dos arquivos utilizados na construção do grafo	23
Figura 2 — Distribuição Espacial das Paradas e Estações por Modal	30
Figura 3 — Recortes regionais (3 km) das arestas de transferência intermodal em cinco regiões estratégicas	35
Figura 4 — Evolução do tamanho do componente principal sob diferentes cenários de remoção	41
Figura 5 — Evolução do parâmetro κ sob diferentes cenários de remoção	42

LISTA DE TABELAS

Tabela 1 — Dimensões dos arquivos GTFS: número de registros e atributos	28
Tabela 2 — Composição dos componentes conexos do grafo antes da integração intermodal	31
Tabela 3 — Comparação dos raios de transferência intermodal testados.....	32
Tabela 4 — Distribuição das estações não-integradas por linha	33
Tabela 5 — Características do grafo antes e após a integração intermodal.....	34
Tabela 6 — Top 10 paradas por <i>degree centrality</i>	36
Tabela 7 — Top 10 paradas por <i>betweenness centrality</i>	37
Tabela 8 — Top 10 paradas por <i>closeness centrality</i>	38
Tabela 9 — Sobreposição entre os <i>rankings</i> de top 10 das três métricas	39
Tabela 10 — Frações críticas de remoção por cenário	43

SUMÁRIO

1 INTRODUÇÃO	9
1.1 MOBILIDADE URBANA EM SÃO PAULO	9
1.2 PROBLEMA DE PESQUISA	9
1.3 OBJETIVOS	10
1.3.1 Objetivo Geral	10
1.3.2 Objetivos Específicos	10
1.4 JUSTIFICATIVA	11
1.5 ESTRUTURA DO TRABALHO	11
2 FUNDAMENTAÇÃO TEÓRICA	13
2.1 MOBILIDADE URBANA E SISTEMAS DE TRANSPORTE PÚBLICO	13
2.1.1 Sistemas de Transporte Público em Regiões Metropolitanas	13
2.1.2 Dados Abertos e o Padrão GTFS	14
2.2 TEORIA DOS GRAFOS E REDES COMPLEXAS	14
2.2.1 Conceitos Básicos	14
2.2.2 O Paradigma das Redes Complexas	15
2.3 REPRESENTAÇÕES DE REDES DE TRANSPORTE EM GRAFOS	15
2.3.1 <i>L-Space</i>	15
2.3.2 <i>P-Space</i> e <i>C-Space</i>	16
2.4 MÉTRICAS DE CENTRALIDADE	16
2.4.1 <i>Degree Centrality</i>	17
2.4.2 <i>Betweenness Centrality</i>	17
2.4.3 <i>Closeness Centrality</i>	18
2.4.4 Complementaridade entre Métricas	18
2.5 VULNERABILIDADE E RESILIÊNCIA ESTRUTURAL DE REDES	19
2.5.1 Definições	19
2.5.2 O Paradigma De Albert, Jeong e Barabási	19
2.5.3 Métricas de Impacto e o Parâmetro de Molloy-Reed (κ)	19
3 METODOLOGIA	21
3.1 DADOS GTFS DA SPTRANS	21
3.1.1 Coleta dos Dados	21

3.1.2 Análise Exploratória dos Dados	22
3.2 CONSTRUÇÃO DO GRAFO.....	23
3.2.1 Modelagem em <i>L-Space</i>	23
3.2.2 Integração Intermodal das Sub-Redes de Metrô e CPTM.....	24
3.3 CÁLCULO DAS MÉTRICAS DE CENTRALIDADE.....	25
3.4 SIMULAÇÕES DE FALHA	25
3.5 FERRAMENTAS COMPUTACIONAIS.....	27
4 RESULTADOS E DISCUSSÃO	28
4.1 ANÁLISE EXPLORATÓRIA E CARACTERÍSTICAS DA REDE	28
4.1.1 Caracterização do <i>Feed</i> GTFS.....	28
4.1.2 Caracterização do Grafo Antes da Integração Intermodal.....	30
4.1.3 Caracterização do Grafo Após a Integração Intermodal.....	32
4.2 ANÁLISE DE VULNERABILIDADE: PARADAS E ESTAÇÕES CRÍTICAS	36
4.2.1 <i>Ranking</i> por <i>Degree Centrality</i>	36
4.2.2 <i>Ranking</i> por <i>Betweenness Centrality</i>	37
4.2.3 <i>Ranking</i> por <i>Closeness Centrality</i>	38
4.2.4 Sobreposição entre os <i>Rankings</i>	39
4.3 ANÁLISE DE RESILIÊNCIA: SIMULAÇÃO DE FALHAS.....	40
4.3.1 Fragmentação do Componente Principal.....	40
4.3.2 Evolução do Parâmetro κ de Molloy-Reed	42
4.3.3 Síntese: Frações Críticas por Cenário.....	43
4.4 IMPLICAÇÕES PARA PLANEJAMENTO URBANO	44
4.4.1 Priorização de Paradas e Estações para Planos de Contingência	44
4.4.2 Robustez do Sistema diante de Falhas Distribuídas	44
4.4.3 Limites Institucionais da Rede da SPTrans e Integração Metropolitana.....	45
5 CONCLUSÃO.....	46
5.1 SÍNTESE DOS ACHADOS	46
5.2 CONTRIBUIÇÕES	47
5.3 LIMITAÇÕES	47
5.4 TRABALHOS FUTUROS	48
REFERÊNCIAS.....	50

APÊNDICE A - MODELO ENTIDADE-RELACIONAMENTO DO GTFS DA SPTRANS.....	52
APÊNDICE B - CÓDIGO-FONTE E NOTEBOOKS	53

1 INTRODUÇÃO

Este capítulo introduz o problema de pesquisa e situa o trabalho em seu contexto urbano e acadêmico. A Seção 1.1 caracteriza a mobilidade urbana em São Paulo e a relevância do sistema de transporte público metropolitano; a Seção 1.2 formula a pergunta de pesquisa e as perguntas específicas derivadas dela; a Seção 1.3 apresenta os objetivos geral e específicos; a Seção 1.4 articula a relevância social e acadêmica da pesquisa; e a Seção 1.5 descreve a organização do trabalho.

1.1 MOBILIDADE URBANA EM SÃO PAULO

A mobilidade urbana constitui um dos maiores desafios das grandes metrópoles contemporâneas, com impactos diretos na qualidade de vida, no acesso a oportunidades de emprego, educação e saúde.

Em São Paulo, o sistema de transporte público operado pela São Paulo Transporte S/A (SPTrans) atende o município e trechos adjacentes da Região Metropolitana, transportando em média 7,1 milhões de passageiros por dia útil, através de uma rede composta por mais de 1.300 linhas de ônibus, 6 linhas de metrô e 7 linhas de trem metropolitano, totalizando mais de 22 mil paradas e estações.

1.2 PROBLEMA DE PESQUISA

A escala e a complexidade do sistema descritos na seção anterior tornam não-trivial a tarefa de identificar quais elementos da rede são realmente determinantes para sua operação. Em sistemas com essas dimensões e complexidade, a importância de uma parada depende de propriedades estruturais da rede como um todo, não sendo redutível à intuição operacional ou à sua localização geográfica.

A literatura recente em análise topológica de redes de transporte público tem avançado significativamente nessa direção, com aplicações em sistemas como o metrô de Chongqing (SONG *et al.*, 2024), a rede multimodal de Madri (MOURONTE-LÓPEZ, 2021) e 42 redes metropolitanas analisadas comparativamente por Massobrio e Cats (2024). Esses estudos convergem em achados relevantes: a remoção de nós com alto *betweenness* compromete a integridade da rede de forma desproporcionalmente maior do que falhas aleatórias, e a importância topológica de uma estação nem sempre corresponde à sua centralidade aparente.

Apesar desses avanços, persistem duas lacunas: em primeiro lugar, conforme apontam Massobrio e Cats (2024), pouco se sabe sobre os aspectos topológicos do processo de recuperação após falhas. Em segundo lugar, estudos sobre a rede de transporte público de São Paulo que integrem a análise de vulnerabilidade e resiliência topológica utilizando dados GTFS atualizados permanecem escassos. A ausência é particularmente relevante dada a escala do sistema e seu contexto latino-americano, sub-representado na literatura internacional.

Diante desse cenário, o presente trabalho formula a seguinte questão central: **como a estrutura topológica da rede de transporte público da SPTrans condiciona sua vulnerabilidade e sua resiliência diante de diferentes regimes de perturbação?**

Decorrem dessa questão três perguntas específicas: (i) qual a estrutura topológica da rede e como ela se distribui em componentes conexos? (ii) quais nós concentram as métricas de centralidade que indicam papel articulador na rede? e (iii) como o tamanho do maior componente conexo evolui sob ataques direcionados em comparação com falhas aleatórias?

1.3 OBJETIVOS

1.3.1 OBJETIVO GERAL

Analisar a estrutura da rede de transporte público do município de São Paulo por meio da Teoria dos Grafos aplicada a dados da São Paulo Transporte S/A (SPTrans) no formato *General Transit Feed Specification* (GTFS), identificando paradas/estações críticas para a resiliência do sistema.

1.3.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos são listados a seguir:

- Modelar a rede de transporte público do município de São Paulo como um grafo, utilizando dados GTFS estáticos, onde paradas/estações são nós, e conexões entre paradas consecutivas são arestas;
- Calcular e comparar métricas de centralidade para identificar as paradas/estações mais críticas da rede;
- Identificar pontos de vulnerabilidade na rede;
- Simular cenários de falha na rede para avaliar a resiliência do sistema;

1.4 JUSTIFICATIVA

A relevância deste trabalho situa-se em duas dimensões complementares: social e acadêmica.

Do ponto de vista social, a identificação de paradas e estações topologicamente críticas oferece subsídios concretos para o planejamento e a operação do sistema de transporte público de São Paulo. Em uma rede que transporta em média 7,1 milhões de passageiros por dia útil (SPTrans, 2026), a interrupção de pontos estruturalmente articuladores pode comprometer significativamente a mobilidade de parcela expressiva da população. O conhecimento sistemático desses pontos permite, entre outras possibilidades, priorizar investimentos em redundância, em corredores estratégicos, antecipar planos de contingência para falhas em estações-chave e fundamentar decisões sobre expansões da rede que reforcem sua robustez estrutural.

Do ponto de vista acadêmico, o trabalho contribui em três frentes. Primeiro, amplia a base empírica da literatura ao aplicar métodos consolidados de análise topológica a uma rede de grande escala em contexto latino-americano, sub-representado nos estudos comparativos disponíveis (MASSOBRIO; CATS, 2024). Segundo, integra em um mesmo desenho metodológico as dimensões de vulnerabilidade e resiliência, que são frequentemente tratadas isoladamente, sobre uma rede multimodal real. Terceiro, demonstra a viabilidade de conduzir análises rigorosas de redes urbanas a partir de dados abertos em formato GTFS e de ferramentas computacionais de uso livre, contribuindo para a reprodutibilidade e a democratização desse tipo de estudo no Brasil.

1.5 ESTRUTURA DO TRABALHO

O trabalho está organizado em cinco capítulos articulados em torno de um percurso analítico que parte da abstração teórica em direção à interpretação prática dos achados. O Capítulo 2 apresenta os fundamentos teóricos que sustentam a análise, com particular atenção à articulação entre Teoria dos Grafos, métricas de centralidade e os conceitos de vulnerabilidade e resiliência estrutural, que raramente é discutida de forma integrada em trabalhos sobre o transporte público brasileiro.

O Capítulo 3 operacionaliza esses fundamentos em procedimentos concretos de análise sobre os dados GTFS da SPTrans, detalhando a modelagem da rede em representação *L-space*, a integração intermodal das sub-redes do Metrô e CPTM por meio de arestas de transferência, o cálculo das métricas de centralidade e o protocolo das simulações de falha. O Capítulo 4

apresenta os resultados das análises organizados em quatro frentes que dialogam diretamente com a literatura discutida no Capítulo 2: caracterização da rede, identificação de paradas e estações críticas, simulação de cenários de remoção e implicações para o planejamento.

O Capítulo 5 sintetiza as contribuições do trabalho em diálogo com as perguntas de pesquisa formuladas na Seção 1.2, discute as limitações inerentes ao recorte adotado e indica possibilidades de pesquisas futuras. As Referências reúnem as obras citadas, e os Apêndices disponibilizam materiais complementares: o esquema completo do *feed* GTFS (Apêndice A) e a descrição dos Jupyter Notebooks desenvolvidos (Apêndice B).

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os fundamentos conceituais que sustentam a análise topológica desenvolvida no trabalho. A organização segue do contexto urbano e operacional para o instrumental analítico: a Seção 2.1 situa o transporte público como objeto de pesquisa em sistemas urbanos; a Seção 2.2 introduz os conceitos da Teoria dos Grafos e do estudo de redes complexas; a Seção 2.3 discute as representações alternativas de redes de transporte como grafos; a Seção 2.4 formaliza as métricas de centralidade adotadas; e a Seção 2.5 articula os conceitos de vulnerabilidade e resiliência estrutural mobilizados nas análises.

2.1 MOBILIDADE URBANA E SISTEMAS DE TRANSPORTE PÚBLICO

A mobilidade urbana é uma dimensão central do funcionamento das cidades contemporâneas, condicionando o acesso da população a oportunidades de trabalho, educação, saúde e serviços. Em metrópoles densas, sistemas eficientes de transporte público reduzem externalidades associadas ao uso intensivo do automóvel particular e contribuem para a equidade no acesso ao território (UN-HABITAT, 2018). O Objetivo de Desenvolvimento Sustentável 11.2 da Organização das Nações Unidas estabelece, nesse sentido, o acesso a sistemas de transporte seguros, acessíveis e sustentáveis como meta global, com indicadores específicos de cobertura caminhável a paradas (UN-HABITAT, 2018).

2.1.1 SISTEMAS DE TRANSPORTE PÚBLICO EM REGIÕES METROPOLITANAS

Sistemas de transporte público metropolitanos são tipicamente multimodais, articulando modos com diferentes capacidades, velocidades e regimes operacionais. Linhas ferroviárias e metroviárias oferecem alta capacidade em corredores troncais; linhas de ônibus operam em malha capilar sobre a rede viária; modos complementares conectam as extremidades do sistema às origens e destinos dos passageiros. A operação integrada desses modos depende de transferências em estações de baldeação, nas quais o usuário articula trechos da viagem em diferentes meios.

A integração intermodal não é apenas operacional, manifesta-se também na estrutura tarifária, com sistemas de bilhete único, e na cartografia funcional do sistema, que apresenta modais como uma rede unificada ao usuário. Do ponto de vista institucional, contudo, a operação de cada modal é frequentemente atribuída a concessionárias distintas, com fronteiras territoriais nem sempre coincidentes. No caso da região metropolitana de São Paulo, as linhas

de ônibus municipais são operadas pela São Paulo Transporte S.A. (SPTrans), enquanto o Metrô e a CPTM operam em sistemas próprios, e o transporte intermunicipal por ônibus está sob a Empresa Metropolitana de Transportes Urbanos (EMTU).

2.1.2 DADOS ABERTOS E O PADRÃO GTFS

A disponibilização de dados abertos pelos operadores de transporte público tem viabilizado análises sistemáticas e reprodutíveis de sistemas em diferentes cidades. O padrão *General Transit Feed Specification* (GTFS) estabelece uma especificação aberta para representação de dados de transporte público, mantida pela MobilityData (2026), e tornou-se referência global. Em sua versão estática, o GTFS descreve a estrutura operacional do sistema por meio de arquivos textuais que registram linhas, viagens, paradas, horários, geometria dos trajetos e regras tarifárias. Em sua versão *Realtime*, fornece informação dinâmica sobre a execução do serviço (chegadas, atrasos, posição dos veículos).

A adoção do padrão GTFS por operadores por todo o mundo é o que permite hoje conduzir análises comparáveis entre sistemas e replicar metodologias em diferentes contextos urbanos. O *feed* estático da SPTrans, utilizado neste trabalho (Subseção 3.1.1), é representativo desse padrão.

2.2 TEORIA DOS GRAFOS E REDES COMPLEXAS

A Teoria dos Grafos é o ramo da matemática que estuda estruturas abstratas formadas por um conjunto de nós (ou vértices) e um conjunto de arestas que conectam pares de nós. Originada no século XVIII com o problema das pontes de Königsberg, formulado por Leonhard Euler, a teoria se consolidou ao longo do século XX como instrumento central para a modelagem de sistemas em que a estrutura de relações entre componentes importa mais do que os atributos individuais dos componentes.

2.2.1 CONCEITOS BÁSICOS

Um grafo G é definido por um par $G = (V, E)$, em que V é o conjunto dos nós e E é o conjunto de arestas. Em um grafo não-direcionado, cada aresta é um par não-ordenado $\{u, v\}$, indicando que a relação entre u e v é simétrica. Em um grafo não-ponderado, todas as arestas têm o mesmo peso, em geral implicitamente 1. O grau de um nó v , denotado por $k(v)$, é o número de arestas incidentes em v .

Dois nós u e v são ditos conexos se existe um caminho que os conecta. Um componente conexo é um subconjunto maximal de nós em que todos os pares são conexos entre si. Um grafo é dito conexo se possui um único componente conexo. Em grafos não-conexos, o componente principal é o componente com o maior número de nós.

2.2.2 O PARADIGMA DAS REDES COMPLEXAS

A partir do final da década de 1990, o estudo de grandes redes empíricas (biológicas, sociais, tecnológicas, urbanas) revelou que muitas dessas estruturas apresentam regularidades estatísticas que as distinguem de grafos puramente aleatórios. Distribuições de graus com cauda longa, alto coeficiente de *clustering*, distâncias médias pequenas mesmo em redes grandes e organização em torno de poucos nós altamente conectados (*hubs*) são propriedades recorrentes (ALBERT; JEONG; BARABÁSI, 2000). O campo que se ocupa dessas regularidades é hoje conhecido como estudo das redes complexas, e fornece o referencial dentro do qual redes de transporte público são analisadas neste trabalho.

Redes de transporte público apresentam características estruturais que combinam elementos de redes regulares (a malha viária impõe geometria) e de redes complexas (a distribuição de viagens concentra-se em poucos terminais e corredores). A consequência prática é que análises topológicas dessas redes precisam mobilizar instrumentos do estudo de redes complexas, em particular, métricas de centralidade e medidas de robustez, adaptados ao contexto urbano.

2.3 REPRESENTAÇÕES DE REDES DE TRANSPORTE EM GRAFOS

A representação de uma rede de transporte público como grafo não é unívoca, uma vez que a definição do que constitui um nó e do que constitui uma aresta admite alternativas, cada uma com implicações para o tipo de análise possível. Três representações são correntes na literatura (PALK *et al.*, 2026; MOURONTE-LÓPEZ, 2021): *L-space*, *P-space* e *C-space*.

2.3.1 *L-SPACE*

Na representação *L-space*, ou *space-of-infrastructure*, paradas são nós e duas paradas são conectadas por uma aresta se forem consecutivas em ao menos uma viagem do sistema. Trata-se da representação mais próxima da estrutura física da rede: as arestas correspondem a trechos efetivamente percorridos entre paradas sequenciais. A *L-space* é a representação

predominante na literatura de análise topológica de transporte público, conforme apontam Palk *et al.* (2026) em revisão sistemática, e adotada neste trabalho (Subseção 3.2.1).

2.3.2 *P-SPACE E C-SPACE*

Na representação *P-space*, ou *space-of-service*, paradas continuam a ser nós, mas duas paradas são conectadas por uma aresta se pertencerem à mesma linha, independentemente de serem consecutivas. Cada linha torna-se um sub-grafo completo, refletindo a ideia de que duas paradas servidas pela mesma linha podem ser alcançadas sem transferência. Essa representação produz grafos consideravelmente mais densos que a *L-space* e é apropriada para análises centradas no número de transferências entre origem e destino.

Já na representação *C-space*, ou espaço de rotas, são as próprias linhas que constituem os nós, ou seja, duas linhas são conectadas por uma aresta se compartilham ao menos uma parada. A *C-space* descarta a granularidade das paradas individuais e expõe a estrutura de articulação do sistema no nível das rotas, sendo particularmente útil para o estudo de transferências e da topologia operacional do serviço.

A literatura também emprega representações bipartidas, em que paradas e linhas constituem conjuntos disjuntos de nós conectados por arestas que registram pertencimento (*B-space*), úteis quando se deseja preservar simultaneamente as duas dimensões.

A escolha entre essas representações condiciona o significado das métricas e a interpretação dos achados. A *L-space* foi adotada neste trabalho por ser a representação que melhor reflete a conectividade da rede, uma propriedade central para a análise de vulnerabilidade e resiliência sob remoção de nós, conduzida nas Seções 4.2 e 4.3.

2.4 MÉTRICAS DE CENTRALIDADE

Métricas de centralidade são funções que atribuem a cada nó de um grafo um valor numérico que expressa sua importância topológica relativa. Diferentes definições de centralidade capturam diferentes noções de importância (no presente trabalho: quantidade de conexões diretas, posição em caminhos mínimos e proximidade média aos demais nós).

A literatura recomenda o uso combinado de múltiplas métricas para análises de vulnerabilidade em redes de transporte (PALK *et al.*, 2026; SONG *et al.*, 2024). Esta seção formaliza as três métricas adotadas neste trabalho.

2.4.1 DEGREE CENTRALITY

A *degree centrality* de um nó v é definida como o número de arestas incidentes em v , normalizado pelo número máximo possível de conexões em um grafo com N nós (FREEMAN, 1979):

$$C_D(v) = \frac{k(v)}{N-1}$$

Onde:

$C_D(v)$ = *degree centrality* do nó v

$k(v)$ = grau de v (número de arestas incidentes em v)

N = número total de nós no grafo

Em redes de transporte público, nós com alta *degree centrality* correspondem a *hubs* operacionais (paradas ou estações em que múltiplas linhas convergem).

2.4.2 BETWEENNESS CENTRALITY

A *betweenness centrality* de um nó v mede a frequência com que v aparece nos caminhos mínimos entre todos os pares de nós da rede (FREEMAN, 1977):

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Onde:

$C_B(v)$ = *betweenness centrality* do nó v

σ_{st} = número total de caminhos mínimos entre os nós s e t

$\sigma_{st}(v)$ = número desses caminhos que passam por v

A soma é feita sobre todos os pares de nós s, t distintos de v

Para fins de comparabilidade entre redes de tamanhos distintos, a métrica é frequentemente normalizada pelo fator $\frac{2}{[(N-1)(N-2)]}$ em grafos não direcionados:

$$C_B^{norm}(v) = \frac{2}{(N-1)(N-2)} \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Nós com alta *betweenness* funcionam como pontes obrigatórias entre regiões da rede, o que significa que sua remoção tende a desconectar conjuntos significativos de pares origem-destino. O cálculo exato da *betweenness* tem complexidade $O(VE)$ pelo algoritmo de Brandes

(2001), o que torna a métrica computacionalmente custosa em redes de grande escala. Aproximações por amostragem são frequentemente adotadas em análises empíricas (MASSOBRIO; CATS, 2024), conforme detalhado na Seção 3.3.

2.4.3 CLOSENESS CENTRALITY

A *closeness centrality* mede a proximidade média de um nó aos demais nós da rede, definida como o inverso da soma das distâncias mínimas de v a todos os outros nós (FREEMAN, 1979):

$$C_c(v) = \frac{N - 1}{\sum_{u \neq v} d(v, u)}$$

Onde:

$C_c(v)$ = *closeness centrality* do nó v

N = número de nós

$d(v, u)$ = distância do caminho mínimo (em número de arestas) entre v e u

A soma é feita sobre todos os nós u distintos de v

Valores altos de *closeness* indicam nós a partir dos quais o restante da rede é alcançável com poucos passos em média. Em redes de transporte público metropolitanas, a *closeness* tende a destacar nós centrais geograficamente, dada a relação entre topologia e geografia nesse tipo de sistema.

2.4.4 COMPLEMENTARIDADE ENTRE MÉTRICAS

As três métricas capturam dimensões distintas de criticidade. *Degree centrality* mede a magnitude local de conectividade; *betweenness centrality* mede a posição estrutural em caminhos mínimos; *closeness centrality* mede a proximidade global média. Estudos empíricos em redes de transporte público têm mostrado que os *rankings* produzidos por essas três métricas exibem sobreposição reduzida, o que sustenta a recomendação de uso combinado (PALK *et al.*, 2026; SONG *et al.*, 2024). A análise apresentada na Subseção 4.2.4 confirma esse padrão para a rede da SPTrans.

2.5 VULNERABILIDADE E RESILIÊNCIA ESTRUTURAL DE REDES

No estudo topológico de redes, vulnerabilidade e resiliência são conceitos articulados que descrevem, respectivamente, a propensão da rede a sofrer degradação sob perturbação e sua capacidade de manter integridade estrutural. Esta seção define operacionalmente os dois conceitos (Subseção 2.5.1), apresenta o paradigma teórico que orienta o trabalho (Subseção 2.5.2) e formaliza as métricas de impacto adotadas (Subseção 2.5.3).

2.5.1 DEFINIÇÕES

No estudo topológico de redes, vulnerabilidade refere-se à propensão da rede a sofrer degradação significativa em sua estrutura quando perturbada, enquanto resiliência refere-se à capacidade de manter integridade estrutural diante de perturbações (MASSOBRIO; CATS, 2024). Embora os dois conceitos sejam intrinsecamente articulados, eles são operacionalizados de formas distintas: a vulnerabilidade é tipicamente avaliada pela identificação de nós ou arestas cuja remoção produziria maior impacto na rede, enquanto resiliência é avaliada pela observação do comportamento da rede sob sequências de remoção.

2.5.2 O PARADIGMA DE ALBERT, JEONG E BARABÁSI

O estudo seminal de Albert, Jeong e Barabási (2000) estabeleceu que redes com distribuição de graus de cauda longa exibem um comportamento dual diante de perturbações, ou seja, alta tolerância a falhas aleatórias e alta fragilidade a ataques direcionados a nós de alto grau. A intuição é direta: na falha aleatória, a probabilidade de remover um nó qualquer é proporcional à sua frequência na rede, e como a maioria dos nós tem grau baixo, falhas aleatórias raramente removem *hubs*; em ataques direcionados, ao contrário, os *hubs* são removidos primeiro, e sua remoção desmonta rapidamente a estrutura de conectividade global.

Esse paradigma orientou desde então um amplo conjunto de estudos sobre robustez de redes empíricas, incluindo redes de transporte público (HOLME *et al.*, 2002; SONG *et al.*, 2024) e é o referencial teórico das simulações de remoção sequencial conduzidas na Seção 3.4 e analisadas na Seção 4.3.

2.5.3 MÉTRICAS DE IMPACTO E O PARÂMETRO DE MOLLOY-REED (κ)

A avaliação quantitativa do impacto de remoções na estrutura da rede recorre a métricas que capturam diferentes aspectos da fragmentação. A mais usual é o tamanho do maior

componente conexo (*largest connected component*, LCC), expresso em geral como fração do tamanho original: a redução do LCC sinaliza fragmentação observada da rede.

Um critério teórico complementar é fornecido pelo parâmetro κ de Molloy e Reed (1995), definido a partir dos dois primeiros momentos da distribuição de graus:

$$\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$$

Onde:

κ = parâmetro de Molloy-Reed

$\langle k \rangle$ = grau médio da rede (primeiro momento da distribuição de graus)

$\langle k^2 \rangle$ = valor esperado do quadrado de graus (segundo momento da distribuição de graus)

Molloy e Reed demonstraram que, em grafos aleatórios não correlacionados, a condição $\kappa > 2$ é necessária para a existência de um componente gigante de tamanho macroscópico. O valor $\kappa = 2$, portanto, funciona como limiar teórico e, à medida que κ se aproxima de 2 por valores superiores, a rede aproxima-se da fragmentação iminente; $\kappa < 2$ sinaliza que o componente gigante não pode ser sustentado pela distribuição de graus vigente.

A combinação entre o LCC, métrica de fragmentação observada, e κ , indicador de fragmentação iminente, fornece a base para a análise de resiliência conduzida na Seção 4.3.

3 METODOLOGIA

A presente pesquisa classifica-se como experimental, uma vez que envolve a realização de simulações computacionais de cenários de falha e ataques direcionados à rede de transporte público. A abordagem adotada possui natureza quantitativa, fundamentada na aplicação de métricas numéricas de centralidade, conectividade e indicadores de resiliência de redes complexas, permitindo avaliar o comportamento estrutural do sistema diante de diferentes níveis de perturbação.

Não há participação de seres humanos na pesquisa. O objeto de estudo consiste exclusivamente em dados estruturados que representam a rede de transporte público de São Paulo, incluindo estações, conexões e relações topológicas entre os elementos da infraestrutura analisada.

3.1 DADOS GTFS DA SPTRANS

Esta seção descreve a base empírica do trabalho. A Subseção 3.1.1 detalha a coleta de dados (origem, data e período de validade) e a Subseção 3.1.2 apresenta procedimentos da análise exploratória conduzida.

3.1.1 COLETA DOS DADOS

Os dados consistem nos arquivos GTFS estáticos disponibilizados publicamente pela São Paulo Transporte S/A (SPTrans) em seu portal de desenvolvedores, coletados em 10 de março de 2026. O conjunto, composto por dez arquivos no formato `.txt`, descreve a estrutura operacional do sistema, incluindo linhas, viagens, paradas, horários, frequências, geometria dos trajetos e regras tarifárias, em conformidade com a especificação GTFS mantida pela MobilityData (2026).

Conforme declarado no arquivo `calendar.txt`, o *feed* possui período de validade entre 1º de outubro de 2023 e 1º de setembro de 2026, configurando-se como uma representação estática do sistema no momento da coleta.

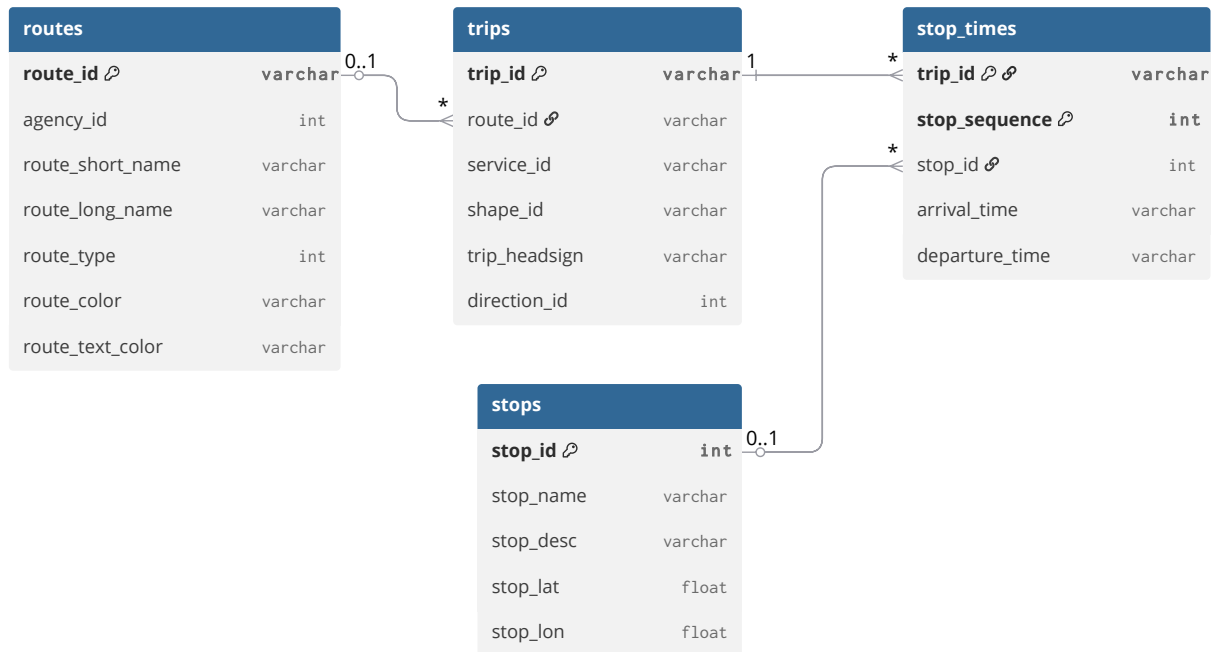
Como dados complementares, foram utilizados os `shapefiles (.shp)` de limites municipais do Estado de São Paulo, disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2024), para a contextualização espacial das paradas no contexto da Região Metropolitana de São Paulo.

3.1.2 ANÁLISE EXPLORATÓRIA DOS DADOS

A análise exploratória teve por objetivo compreender a estrutura, dimensão e qualidade do *feed* e identificar os arquivos pertinentes à modelagem da rede como grafo. Foi conduzida verificação da consistência dos arquivos, identificação de valores ausentes, análise descritiva dos atributos e inspeção espacial da distribuição das paradas, esta última por meio de *spatial join* entre as coordenadas e os limites municipais do IBGE.

Da exploração resultou a seleção de quatro arquivos: `routes.txt` (linhas), `trips.txt` (viagens), `stops.txt` (paradas e estações) e `stop_times.txt` (sequência ordenada de paradas por viagem). Em conjunto, esses arquivos contêm todas as informações necessárias à construção do grafo em representação *L-space*. Os demais arquivos foram desconsiderados por não contribuírem para a modelagem topológica adotada: `agency.txt` contém apenas metadados da operadora; `calendar.txt` e `frequencies.txt` descrevem dimensões temporais não incorporadas na análise estática; `shapes.txt` armazena a geometria dos trajetos, útil para visualização cartográfica, mas não para a definição de adjacências entre paradas; e `fare_attributes.txt` e `fare_rules.txt` referem-se à estrutura tarifária, alheia ao escopo do estudo. A Figura 1 apresenta o modelo relacional dos quatro arquivos efetivamente utilizados¹.

¹ A versão completa do diagrama, com os dez arquivos do *feed* GTFS da SPTrans, encontra-se no Apêndice A e está disponível em <https://raw.githubusercontent.com/cintia-shinoda/public-transit-graph-theory/main/images/GTFS-SPTrans/MER-GTFS-SPTrans.svg>.

Figura 1 — Modelo Entidade-Relacionamento dos arquivos utilizados na construção do grafo

Fonte: Elaborado pelos autores (2026) com base em MobilityData (2026).

3.2 CONSTRUÇÃO DO GRAFO

A construção do grafo foi realizada em dois estágios: a modelagem em representação *L-space* a partir do arquivo `stop_times.txt` (Subseção 3.2.1) e a integração intermodal das sub-redes do Metrô e CPTM ao componente principal por meio de arestas de transferência (Subseção 3.2.2).

3.2.1 MODELAGEM EM *L-SPACE*

A rede foi modelada em representação *L-space*: paradas são nós e duas paradas são conectadas por uma aresta se forem consecutivas em ao menos uma viagem do sistema. Trata-se da representação mais adotada na literatura de análise topológica de transporte público (PALK *et al.*, 2026), tendo sido empregada também por Mouronte-López (2021) e Song *et al.* (2024) em estudos com objetivos análogos.

A escolha por um grafo não-direcionado decorre da natureza operacional do sistema, em que a quase totalidade das linhas opera em ambos os sentidos. A escolha por um grafo não-ponderado se alinha com Palk *et al.* (2026) e Mouronte-López (2021), isolando os efeitos da estrutura topológica. Informações de nível de serviço, ainda que disponíveis no *feed* via `frequencies.txt`, foram deliberadamente abstraídas.

A construção foi realizada a partir do arquivo `stop_times.txt`, que registra a sequência ordenada de paradas para cada viagem. Para cada `trip_id`, foram extraídas as paradas em ordem de `stop_sequence` e adicionadas arestas entre paradas adjacentes na sequência. *Self-loops* eventualmente identificados (arestas em que origem e destino possuíam o mesmo identificador) foram removidos. A caracterização do grafo resultante e a análise de seus componentes conexos são apresentadas na Seção 4.1.

3.2.2 INTEGRAÇÃO INTERMODAL DAS SUB-REDES DE METRÔ E CPTM

A análise topológica do grafo bruto seria limitada a uma representação parcial do sistema, uma vez que o componente principal corresponde exclusivamente à rede de ônibus, enquanto as estações de Metrô e CPTM permanecem como sub-redes isoladas.

Essa fragmentação não reflete a realidade operacional do sistema, na qual passageiros articulam diariamente os três modais por meio de baldeações em estações de integração. Essa característica precisa ser representada no grafo para que as análises subsequentes de centralidade, vulnerabilidade e resiliência sejam significativas. Caso a integração não fosse efetuada, a análise reportaria as paradas críticas da rede de ônibus somente e não da rede de transporte público.

Para representar a integração no grafo, foram criadas arestas de transferência conectando estações de Metrô e CPTM a paradas de ônibus situadas em sua vizinhança imediata. A operação consiste em, para cada estação do Metrô ou CPTM, identificar todas as paradas de ônibus localizadas dentro de um raio de 400 metros e adicionar uma aresta entre os nós correspondentes. Essa implementação deu-se por meio de uma estrutura de árvores espaciais (`scipy.spatial.cKDTree`), construída sobre as coordenadas geográficas das paradas de ônibus, permitindo consultas eficientes de todos os pontos dentro do raio especificado para cada estação.

A escolha do raio de 400 metros segue prática consolidada em planejamento de transporte público que adota essa distância como referência de cobertura caminhável a paradas de ônibus (EL-GENEIDY *et al.*, 2014), equivalente a aproximadamente cinco minutos de caminhada. Trata-se de valor mais conservador que o padrão internacional de 500 metros adotado pelo indicador SDG 11.2.1 (UN-HABITAT, 2018), o que reforça o critério aplicado neste estudo. Para validar empiricamente essa escolha no contexto da rede da SPTrans, foi conduzida análise comparativa com raios de 200 e 600 metros, cujos resultados são apresentados na Subseção 4.1.3.

Após a integração, apenas um componente do grafo permaneceu isolado: a travessia hidroviária na Represa Billings, naturalmente segregada da malha de ônibus dada sua localização aquática. O trecho de paradas de ônibus na Rua Prof. Francisco Pinheiro, isolado no grafo bruto, passou a integrar o componente principal por meio das arestas de transferência intermodal adicionadas, indicando que a proximidade entre essas paradas e estações de Metrô ou CPTM permitiu sua reconexão à rede.

3.3 CÁLCULO DAS MÉTRICAS DE CENTRALIDADE

A análise de vulnerabilidade da rede apoia-se em três métricas de centralidade, calculadas sobre o componente principal integrado: *degree centrality*, *betweenness centrality* e *closeness centrality*. As três métricas, formalizadas no Capítulo 2, capturam dimensões distintas e complementares do papel topológico de um nó na rede: *degree* quantifica o número de conexões diretas; *betweenness* mede a frequência com que um nó aparece nos caminhos mínimos entre outros pares; e *closeness* avalia a proximidade média do nó em relação aos demais.

A combinação das três métricas é recomendada por Palk *et al.* (2026) como abordagem mais robusta do que a dependência de um único indicador, dado que cada métrica pode identificar nós críticos por razões diferentes, achado corroborado por Song *et al.* (2024) na análise da rede de metrô de Chongqing.

Em razão da escala do componente principal (mais de 22 mil nós), o cálculo exato de *betweenness centrality* envolve custo computacional elevado. Adotou-se, portanto, o método de aproximação por amostragem implementado no NetworkX (parâmetro κ), que estima a métrica a partir de um subconjunto de nós-fonte amostrados aleatoriamente. Essa abordagem é amplamente utilizada em estudos com redes de grande escala (MASSOBRIO; CATS, 2024) e produz *rankings* altamente correlacionados com o cálculo exato. Para garantir a reprodutibilidade dos resultados, foi fixada uma semente aleatória (*random seed*) na amostragem.

Degree centrality e *closeness centrality* foram calculadas de forma exata, uma vez que apresentam complexidade computacional substancialmente menor.

3.4 SIMULAÇÕES DE FALHA

A avaliação da resiliência da rede foi conduzida por meio de simulações de remoção sequencial de nós, em dois regimes de perturbação derivados do paradigma de Albert, Jeong e

Barabási (2000): ataque direcionado e falha aleatória. O ataque direcionado remove nós em ordem decrescente de uma métrica de centralidade, simulando perturbações concentradas em pontos topologicamente críticos. A falha aleatória remove nós em ordem aleatória, simulando perturbações difusas como panes generalizadas. Em conjunto, os dois regimes permitem caracterizar o padrão de tolerância e fragilidade da rede.

Quatro variantes de ataque direcionado foram simuladas, uma para cada uma das três métricas de centralidade calculadas na Seção 3.3 (*degree*, *betweenness* e *closeness*) e uma adicional baseada em um *ranking* composto. O *ranking* composto é a média aritmética das três centralidades, cada uma previamente normalizada para o intervalo $[0,1]$ por divisão pelo valor máximo da respectiva métrica. Seu objetivo é testar se a combinação das três dimensões de criticidade produz um ataque mais ou menos eficaz que cada métrica isoladamente.

Em todas as variantes, a ordem de remoção foi definida a partir do *ranking* calculado sobre o grafo original, sem recálculo das centralidades a cada remoção. Trata-se de uma escolha conservadora. A remoção sequencial com recálculo, usual em estudos com horizonte de simulação mais longo (HOLME *et al.*, 2002), tende a produzir ataques mais eficientes, mas a custo computacional substancialmente maior. A interpretação dos resultados leva em conta essa decisão.

A falha aleatória foi simulada em 30 repetições independentes, com sementes aleatórias fixadas entre 0 e 29 para garantir reprodutibilidade. As curvas reportadas correspondem à média entre repetições, com o desvio-padrão exibido como banda de incerteza.

Em ambos os regimes, foi removida uma fração de nós entre 0% e 20% do componente principal integrado. O passo de remoção foi refinado no intervalo inicial, em que ocorre a maior parte da fragmentação topológica esperada para redes de cauda longa: 0,1% entre 0% e 5% e 0,5% entre 5% e 20%, totalizando aproximadamente 80 pontos por curva. O limite superior de 20% foi adotado por ser suficiente para capturar a transição de fase em todos os cenários simulados.

A cada ponto da curva, foram calculadas duas métricas de impacto. A primeira é o tamanho relativo do maior componente conexo (LCC), expresso como percentual do componente principal original, que quantifica o grau de fragmentação da rede. A segunda é o parâmetro κ de Molloy e Reed (1995), definido como $\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$, onde k é o grau médio e $\langle k^2 \rangle$ é o segundo momento da distribuição de graus. O valor $\kappa = 2$ corresponde ao limiar teórico para a existência de um componente gigante em grafos aleatórios não correlacionados. Valores acima de 2 indicam que a rede mantém integridade estrutural, já valores abaixo de 2 sinalizam a

iminência ou consumação da fragmentação. As curvas e as frações críticas para cada cenário são apresentadas na Seção 4.3.

3.5 FERRAMENTAS COMPUTACIONAIS

A construção e a análise do grafo foram realizadas em linguagem Python 3.11. Para a manipulação dos dados do *feed* GTFS, foram utilizadas as bibliotecas Pandas e NumPy. A modelagem do grafo, o cálculo das métricas de centralidade (Seção 3.3) e as simulações de resiliência (Seção 3.4) foram implementadas com a biblioteca NetworkX, padrão para análise de grafos em Python. A integração intermodal das sub-redes de Metrô e CPTM (Subseção 3.2.2) utilizou a estrutura de árvore espacial `cKDTree` da biblioteca SciPy para consultas eficientes de vizinhança por raio. Para as operações geoespaciais, foi utilizada a biblioteca GeoPandas. As visualizações estáticas foram produzidas com Matplotlib.

A reprodutibilidade dos resultados foi garantida pela fixação de *seeds* aleatórias nos procedimentos estocásticos: `seed=42` na aproximação da *betweenness centrality* (Seção 3.3) e `seed=0` a `seed=29` nas 30 repetições da simulação de falha aleatória (Seção 3.4). O código-fonte completo encontra-se em quatro Jupyter Notebooks sequenciais, está disponível em <https://github.com/cintia-shinoda/public-transit-graph-theory>. A descrição detalhada dos notebooks é apresentada no Apêndice B.

4 RESULTADOS E DISCUSSÃO

Este capítulo apresenta e discute os resultados das análises realizadas sobre o componente principal integrado da rede da SPTrans. A organização segue o percurso analítico do trabalho. A Seção 4.1 reúne a caracterização exploratória do *feed* GTFS, do grafo bruto e dos efeitos da integração intermodal sobre a estrutura da rede. A Seção 4.2 apresenta a análise de vulnerabilidade, identificando paradas e estações topologicamente críticas a partir das três métricas de centralidade. A Seção 4.3 apresenta a análise de resiliência, com simulações de remoção de nós sob ataque direcionado e falha aleatória. A Seção 4.4 articula as implicações dos achados para o planejamento da rede.

4.1 ANÁLISE EXPLORATÓRIA E CARACTERÍSTICAS DA REDE

A análise exploratória dos dados GTFS coletados em 10 de março de 2026 e a caracterização do grafo construído nas etapas metodológicas (Capítulo 3) são aqui apresentadas em três subseções: a caracterização do *feed* (4.1.1), a análise do grafo antes da integração intermodal (4.1.2) e os efeitos da integração das sub-redes do Metrô e CPTM (4.1.3).

4.1.1 CARACTERIZAÇÃO DO *FEED* GTFS

O *feed* GTFS da SPTrans é composto por dez arquivos, cuja dimensão é sintetizada na Tabela 1. Destacam-se a expressiva volumetria de *shapes.txt* (1.134.502 registros), que descreve a geometria detalhada dos trajetos, e de *stop_times.txt* (98.434 registros), que sustenta a construção do grafo descrita na Seção 3.2. Os arquivos utilizados na modelagem da rede como grafo (*routes.txt*, *trips.txt*, *stops.txt* e *stop_times.txt*) totalizam 123.939 registros distribuídos em 23 atributos.

Tabela 1 — Dimensões dos arquivos GTFS: número de registros e atributos

Arquivo	Número de registros	Número de atributos
<i>agency</i>	1	5
<i>calendar</i>	6	10
<i>fare_attributes</i>	6	6
<i>fare_rules</i>	5.379	5
<i>frequencies</i>	40.146	4

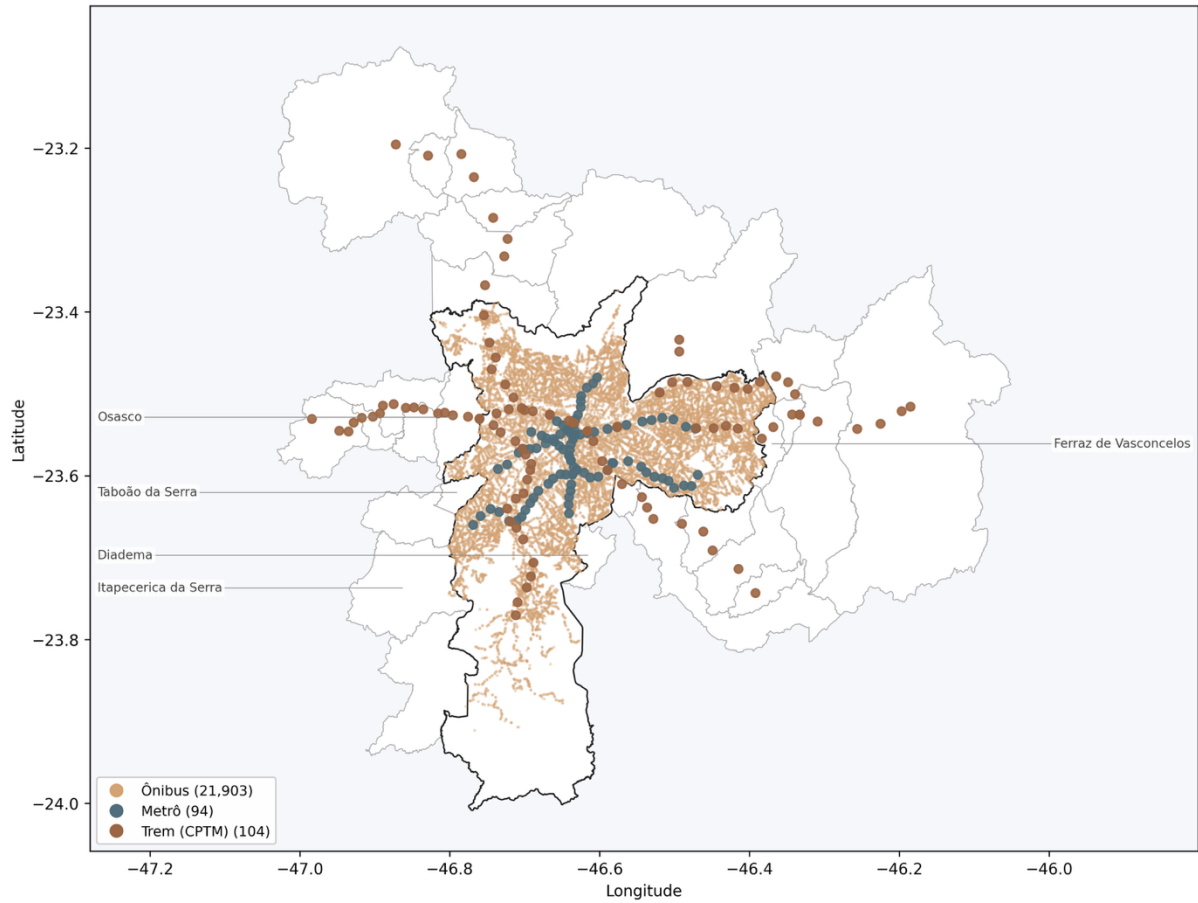
routes	1.348	7
shapes	1.134.502	5
stop_times	98.434	5
stops	22.103	5
trips	2.254	6

Fonte: Elaborado pelos autores (2026), com base em SPTrans (2026).

A consolidação das paradas e estações com pelo menos uma viagem associada resultou em 22.101 nós com modal identificado, distribuídos em três modais: 21.903 paradas de ônibus (99,1%), 94 estações do Metrô (0,4%) e 104 estações da CPTM (0,5%). As duas paradas listadas em `stops.txt`, mas ausentes em `stop_times.txt`, foram desconsideradas, por não integrarem qualquer viagem do sistema.

A Figura 2 apresenta a distribuição espacial das paradas e estações por modal sobre os limites municipais. Observa-se forte concentração no município de São Paulo, com as paradas de ônibus formando uma malha densa que cobre o território da capital, enquanto as estações de Metrô se concentram no eixo central e as estações da CPTM se estendem por municípios vizinhos ao longo dos eixos das linhas sobre trilhos.

Figura 2 — Distribuição Espacial das Paradas e Estações por Modal



Fonte: Elaborado pelos autores (2026), com base no IBGE (2024) e na SPTrans (2026).

O *spatial join* entre as coordenadas das paradas e os limites municipais identificou a presença de paradas em 29 municípios distintos. A maior parte concentra-se no município de São Paulo (21.965 paradas, 99,4%), enquanto os demais municípios apresentam presença marginal.

Apenas cinco municípios (Osasco, Itapeverica da Serra, Diadema, Taboão da Serra e Ferraz de Vasconcelos) apresentam pelo menos cinco paradas, e foram identificados nominalmente na Figura 2. Os demais municípios da Região Metropolitana, atendidos por outras operadoras (notadamente EMTU para o transporte intermunicipal), não compõem o sistema sob responsabilidade da SPTrans e, portanto, não integram o escopo da rede analisada.

4.1.2 CARACTERIZAÇÃO DO GRAFO ANTES DA INTEGRAÇÃO INTERMODAL

A modelagem em representação *L-space* a partir do arquivo `stop_times.txt` identificou inicialmente 29.556 conexões únicas entre paradas. Durante a construção, foram detectados e removidos 16 *self-loops* (arestas em que origem e destino possuíam o mesmo

identificador). O grafo resultante contém 22.101 nós e 29.540 arestas, com grau médio de 2,67 e densidade de 0,000121, indicando uma rede esparsa (característica esperada para sistemas de transporte público em malha viária).

A análise dos componentes conexos revelou 14 componentes desconectados. O componente principal corresponde à rede de ônibus e contém 21.890 nós (99% do total). Os 13 componentes menores correspondem majoritariamente às linhas do Metrô e da CPTM, que aparecem como sub-redes isoladas por não compartilharem paradas consecutivas com as linhas de ônibus no *feed* GTFS. A Tabela 2 detalha a composição dos componentes.

Tabela 2 — Composição dos componentes conexos do grafo antes da integração intermodal

#	Quantidade de nós	Modal / Linha
1	21.890	Rede de ônibus
2	42	CPTM - ramo Itapevi (Santa Rita, Eng. Cardoso)
3	28	CPTM - ramo Calmon Viana / Itaquaquecetuba
4	23	Metrô Linha 1 - Azul (Jabaquara, Conceição, São Judas)
5	18	CPTM Linha 7 - Rubi (Lapa, Campo Limpo Paulista)
6	18	Metrô Linha 3 - Vermelha (Tatuapé, República, Carrão)
7	17	Metrô Linha 5 - Lilás (Borba Gato, Brooklin, Largo Treze)
8	14	Metrô Linha 4 - Amarela (Clínicas, Consolação)
9	13	CPTM (Brás, Tamanduateí, São Caetano)
10	11	Trecho de ônibus isolado* (R. Prof. Francisco Pinheiro)
11	11	Metrô Linha 15 - Prata (São Mateus, Sapopemba)
12	11	Metrô Linha 4 - Amarela trecho oeste (Vila Sônia, Higienópolis-Mackenzie)
13	3	Trecho aeroportuário (Aeroporto Guarulhos, Eng. Goulart)
14	2	Travessia hidroviária na Represa Billings

Fonte: Elaborado pelos autores (2026).

* Componente cuja desconexão sugere inconsistência no cadastramento do *feed*; questão retomada na Seção 5.3.

A fragmentação observada reflete uma característica do *feed* GTFS: a integração entre modais é representada operacional e tarifariamente (por meio do Bilhete Único), mas não

topologicamente. Essa lacuna motivou a etapa de integração descrita na Subseção 3.2.2, cujos efeitos são apresentados a seguir.

4.1.3 CARACTERIZAÇÃO DO GRAFO APÓS A INTEGRAÇÃO INTERMODAL

A validação empírica do raio de transferência de 400 metros adotado na construção das arestas intermodais (Subseção 3.2.2) foi conduzida por meio de análise comparativa com raios de 200 e 600 metros. Os resultados, sintetizados na Tabela 3, evidenciam um ponto de saturação no raio de 400 metros.

Tabela 3 — Comparação dos raios de transferência intermodal testados

Raio	Número de arestas de transferência criadas	Componentes resultantes	Estações conectadas	% das estações totais
200	1.308	2	138	69,7%
400	2.624	2	150	75,8%
600	4.877	2	150	75,8%

Fonte: Elaborado pelos autores (2026).

A análise revela três achados relevantes. Primeiro, a estrutura do grafo integrado é insensível à mudança do raio: nos três cenários, o componente principal contém 22.099 nós (99,99% do total) e um componente permaneceu isolado (travessia hidroviária na Represa Billings). Segundo, a transição de 400 para 600 metros não integra nenhuma estação adicional, embora adicione 2.253 arestas que correspondem a conexões redundantes entre estações já integradas e paradas de ônibus a distâncias superiores ao padrão caminhável adotado (EL-GENEIDY *et al.*, 2014). Terceiro, 48 estações do Metrô e CPTM (24,2% do total de 198) permanecem não conectadas mesmo com o raio de 600 metros.

A combinação desses achados sustenta a escolha de 400 metros como ponto de saturação, uma vez que se trata do menor raio que atinge a cobertura máxima de estações alcançável pelo método, sem incorrer no excesso de conexões a distâncias incompatíveis com transferência efetiva a pé.

A análise das 48 estações não integradas revela um padrão estrutural relevante. A totalidade dessas estações pertence à CPTM (nenhuma estação do Metrô ficou desconectada). Mais significativo, todas se localizam em ramais que se estendem para municípios da Região Metropolitana fora do território da capital, conforme detalhado na Tabela 4. A Linha 8-

Diamante, que cruza Osasco, Carapicuíba, Barueri, Jandira e Itapevi, concentra o maior número de estações desconectadas (15); a Linha 10-Turquesa, que percorre o ABC Paulista, tem 9 estações nessa condição; a Linha 11-Coral, em direção a Mogi das Cruzes, também tem 9; a Linha 7-Rubi, em direção a Jundiaí, tem 8.

Tabela 4 — Distribuição das estações não-integradas por linha

Linha	Estações não integradas	Trecho
Linha 8 - Diamante	15	Osasco - Itapevi
Linha 10 - Turquesa	9	ABC paulista (Mauá, Santo André, São Caetano)
Linha 11 - Coral	9	Brás Cubas - Mogi das Cruzes
Linha 7 - Rubi	8	Caieiras - Jundiaí
Linha 12 - Safira	4	Eng. Manoel Feio - Itaquaquecetuba
Linha 13 - Jade	2	Guarulhos-Cecap - Aeroporto-Guarulhos
Linhas 8 + 9	1	Osasco (estação de transferência)

Fonte: Elaborado pelos autores (2026).

Nos municípios da Região Metropolitana atendidos por essas estações da CPTM, o transporte por ônibus é operado por outras concessionárias, cujos dados não integram o *feed* GTFS analisado. As 48 estações desconectadas evidenciam o limite institucional da rede da SPTrans, que se estende sobre trilhos para municípios vizinhos por meio das linhas da CPTM, mas cuja rede de ônibus permanece restrita ao território da capital. O aumento do raio de transferência para 600 metros confirma essa interpretação, ou seja, não foram identificadas paradas de ônibus a 401-600 metros das estações desconectadas, indicando que há ausência da malha de ônibus naquele território. Essa limitação é retomada na Seção 5.3.

Após a integração com raio de 400 metros, o grafo passou a apresentar dois componentes conexos, com o componente principal abrangendo 22.099 (99,99% do total). As 2.624 arestas de transferência adicionadas conectaram 150 das 198 estações de Metrô e CPTM (94 + 104) a pelo menos uma parada de ônibus na vizinhança. Apenas o componente da travessia hidroviária na Represa Billings (dois nós) permaneceu isolado, dada sua localização aquática que segrega naturalmente a malha de ônibus. Os demais 13 componentes do grafo bruto foram integrados ao componente principal. A Tabela 5 sintetiza as características do grafo antes e depois da integração intermodal. As análises de centralidade (Seção 4.2) e as simulações

de resiliência (Seção 4.3) foram conduzidas exclusivamente sobre o componente principal integrado.

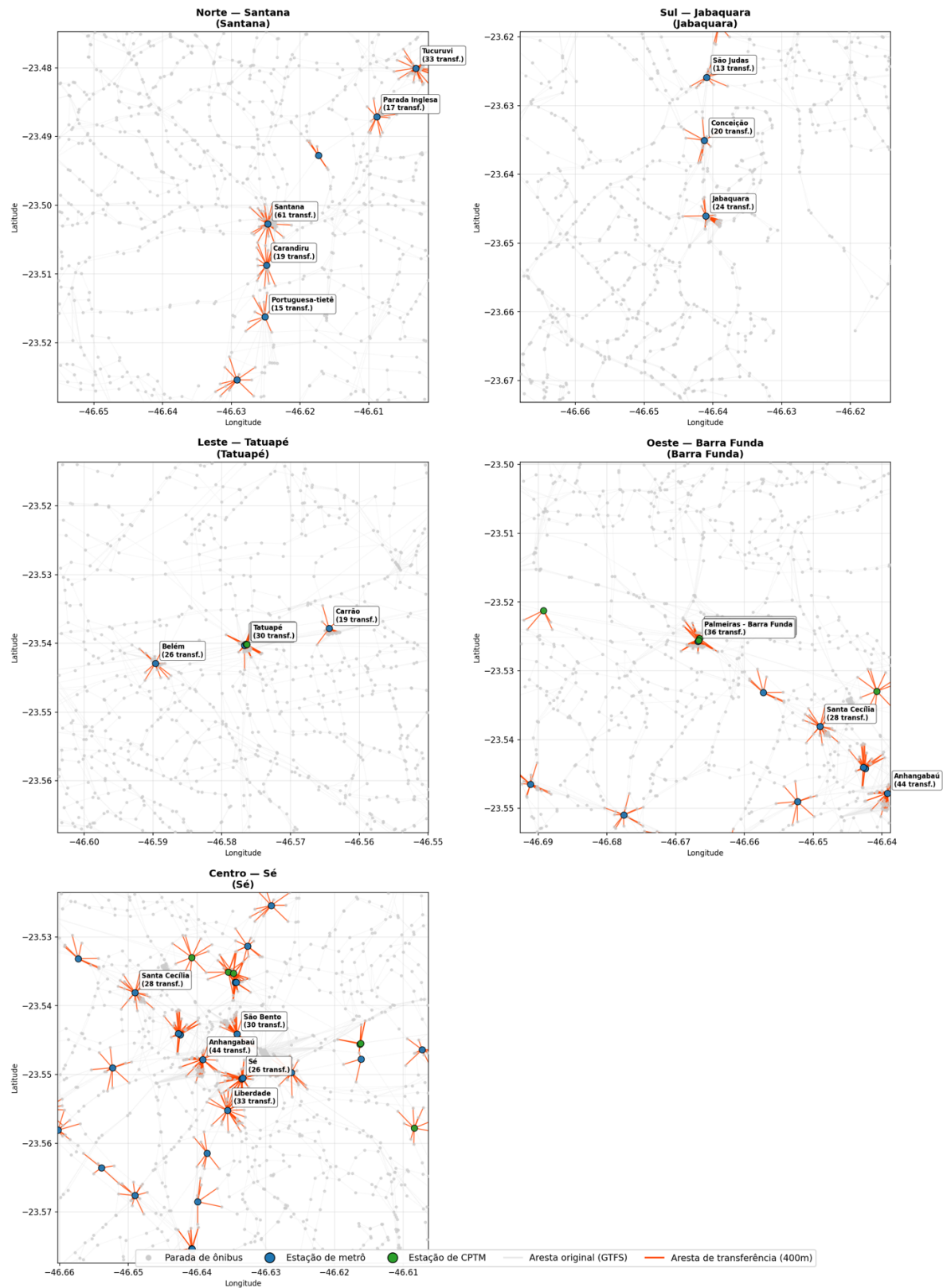
Tabela 5 — Características do grafo antes e após a integração intermodal

Métrica	Antes da Integração	Após a Integração
Nós	22.101	22.099
Arestas	29.540	32.163
Componentes conexos	14	1
Tamanho do componente principal	21.890 (99%)	22.099 (100%)
Grau médio	2,67	2,91

Fonte: Elaborado pelos autores (2026).

A Figura 3 ilustra o resultado da integração intermodal em cinco regiões representativas, centradas em estações estratégicas distribuídas pelas quatro direções cardeais e pelo centro histórico da capital: Sé (centro), Santana (norte), Jabaquara (sul), Tatuapé (leste) e Barra Funda (oeste). A escolha contempla terminais e estações de transferência localizados em diferentes setores urbanos, permitindo visualizar como a integração intermodal se materializa em contextos topologicamente distintos. Cada recorte de 3 km de raio evidencia, em escala apropriada para visualização das transferências, as arestas adicionadas (em laranja) conectando estações sobre trilhos (Metrô em azul e CPTM em verde) a paradas de ônibus (em cinza). A escala de 3 km foi escolhida porque, em escala cidadina, as arestas de transferência de 400 metros correspondem a aproximadamente 0,33% da largura visualizada, enquanto no recorte de 3 km representam cerca de 6,67% da largura, tornando-se claramente perceptíveis.

Figura 3 — Recortes regionais (3 km) das arestas de transferência intermodal em cinco regiões estratégicas



Fonte: Elaborado pelos autores (2026), com base em IBGE (2024) e SPTrans (2026).

4.2 ANÁLISE DE VULNERABILIDADE: PARADAS E ESTAÇÕES CRÍTICAS

A análise de vulnerabilidade da rede objetiva identificar quais paradas e estações exercem papel topológico mais crítico e cuja remoção tende a causar maior impacto na conectividade do sistema. Para essa identificação, foram calculadas três métricas de centralidade sobre o componente principal integrado, conforme descrito na Seção 3.3: *degree centrality*, *betweenness centrality* e *closeness centrality*. Cada métrica captura uma dimensão distinta do papel topológico de um nó na rede, e a comparação entre os *rankings* produzidos por cada uma permite identificar paradas e estações críticas sob diferentes critérios.

Antes da apresentação dos *rankings*, cabe uma observação metodológica relativa à estrutura do *feed* GTFS: estações de integração (entre linhas do Metrô, entre linhas da CPTM ou entre dois modais) são representadas por *stop_id* distintos para cada plataforma ou linha. A Estação Brás, por exemplo, está representada por três *stop_id* (dois da CPTM e um da Linha 3-Vermelha do Metrô); a Estação Luz, por quatro. No grafo construído, esses *stop_id* constituem nós distintos, conectados entre si pelas arestas de transferência intermodal (cf. Subseção 3.2.2). Em consequência, certas estações aparecem mais de uma vez nos *rankings* que se seguem, refletindo a natureza efetivamente multiplataforma desses pontos de articulação.

4.2.1 RANKING POR DEGREE CENTRALITY

A *degree centrality* de um nó é proporcional ao número de conexões diretas que esse nó possui na rede. Em redes de transporte público, paradas com alto *degree* correspondem a *hubs* operacionais (pontos onde múltiplas linhas convergem). A Tabela 6 apresenta as dez paradas com maior *degree* no componente principal integrado.

Tabela 6 — Top 10 paradas por *degree centrality*

#	Parada	Modal	<i>Degree</i>	<i>Betweenness</i>	<i>Closeness</i>
1	Santana	Metrô	63	0,114	0,044
2	Corinthians-Itaquera	CPTM	61	0,203	0,050
3	Largo Treze	Metrô	61	0,035	0,038
4	Corinthians-Itaquera	Metrô	60	0,021	0,046
5	Lapa (Linha 8)	CPTM	59	0,143	0,048
6	Grajaú	CPTM	46	0,104	0,033
7	Pirituba	CPTM	46	0,106	0,044
8	Anhangabaú	Metrô	46	0,007	0,050

9	São Mateus	Metrô	45	0,044	0,034
10	Av. Pe. José Maria	Ônibus	43	0,001	0,037

Fonte: Elaborado pelos autores (2026).

O *ranking* por *degree* é dominado por estações terminais ou de grande convergência operacional. Santana, Largo Treze, Anhangabaú e São Mateus são terminais de linhas do Metrô; Corinthians-Itaquera, Lapa, Grajaú e Pirituba concentram baldeação intensa entre CPTM e linhas de ônibus. A presença de uma única parada de ônibus no top 10 (Av. Pe. José Maria) confirma que, em escala citadina, são as estações sobre trilhos que articulam o maior número de conexões diretas. Importante notar, entretanto, que o *degree* não captura por si só a importância topológica de um nó: a Estação Anhangabaú aparece com grau 46, mas sua *betweenness* (0,007) é uma ordem de magnitude inferior à dos demais nós da tabela, indicando que, apesar das muitas conexões diretas, ela não constitui rota obrigatória para grande parte dos caminhos mínimos da rede.

4.2.2 RANKING POR BETWEENNESS CENTRALITY

A *betweenness centrality* mede a frequência com que um nó aparece nos caminhos mínimos entre todos os pares de nós da rede. Nós com alta *betweenness* funcionam como pontes obrigatórias, pois sua remoção implicaria em desvios significativos para parcela expressiva do tráfego. A Tabela 7 apresenta as dez paradas com maior *betweenness centrality*.

Tabela 7 — Top 10 paradas por *betweenness centrality*

#	Parada	Modal	<i>Degree</i>	<i>Betweenness</i>	<i>Closeness</i>
1	Brás	CPTM	7	0,380	0,052
2	Luz	CPTM	16	0,365	0,053
3	Tatuapé	CPTM	32	0,258	0,051
4	Morumbi	CPTM	6	0,224	0,041
5	Granja Julieta	CPTM	6	0,216	0,040
6	João Dias	CPTM	7	0,215	0,039
7	Corinthians-Itaquera	CPTM	61	0,203	0,050
8	Vila Olímpia	CPTM	8	0,199	0,042
9	Berrini	CPTM	6	0,195	0,042
10	Tatuapé	CPTM	32	0,193	0,051

Fonte: Elaborado pelos autores (2026).

Três achados emergem desta tabela. Primeiro, as dez paradas mais críticas em *betweenness* são todas da CPTM, evidenciando o papel articulador deste modal entre regiões da rede. Segundo, a estrutura da Linha 9-Esmeralda da CPTM, que percorre o eixo do Rio Pinheiros e conecta a região sul ao oeste e norte da capital, concentra cinco das dez paradas mais críticas (Morumbi, Granja Julieta, João Dias, Vila Olímpia e Berrini). Terceiro, e mais relevante, o nó com maior *betweenness* da rede, Estação Brás, apresenta *degree* 7. Esse achado contrasta com o *ranking* por *degree*, que identifica como mais importantes os *hubs* com maior número de conexões diretas. A Estação Brás concentra aproximadamente 38% dos caminhos mínimos da rede, com apenas sete vizinhos diretos, o que corrobora o achado de Song *et al.* (2024) de que estações de baixo *degree* podem ter impacto topológico desproporcional. Trata-se de um nó articulador clássico no paradigma de redes complexas (ALBERT; JEONG; BARABÁSI, 2000): poucos vizinhos diretos, mas posição topológica que torna sua remoção crítica para a integridade da rede.

4.2.3 RANKING POR CLOSENESS CENTRALITY

A *closeness centrality* mede a proximidade média de um nó a todos os demais nós da rede. Valores altos indicam paradas a partir das quais é possível alcançar o restante do sistema com poucos passos, em média. A Tabela 8 apresenta as dez paradas com maior *closeness*.

Tabela 8 — Top 10 paradas por *closeness centrality*

#	Parada	Modal	Degree	Betweenness	Closeness
1	Av. Prestes Maia, 919	Ônibus	7	0,075	0,053
2	Luz	CPTM	16	0,365	0,053
3	Av. Prestes Maia, 940	Ônibus	10	0,047	0,053
4	Av. Tiradentes, 2194	Ônibus	10	0,075	0,053
5	R. Brig. Tobias, 557	Ônibus	7	0,019	0,053
6	Pça. Da Luz	Ônibus	6	0,031	0,053
7	Pça. Da Luz	Ônibus	6	0,032	0,053
8	R. Washington Luis, 98	Ônibus	8	0,020	0,053
9	Brás	CPTM	7	0,380	0,052
10	R. Florêncio De Abreu, 848	Ônibus	7	0,024	0,052

Fonte: Elaborado pelos autores (2026).

Em contraste com os dois *rankings* anteriores, a *closeness* é dominada por paradas de ônibus localizadas no entorno da Estação da Luz, no centro histórico da capital. As paradas Av. Prestes Maia, Av. Tiradentes, R. Brig. Tobias, Pça Da Luz, R. Washington Luís e R. Florêncio de Abreu situam-se todas dentro do raio de 400 metros da Estação Luz e são, em consequência da integração intermodal, conectadas diretamente a ela por arestas de transferência. O resultado reflete uma propriedade geográfico-topológica: o centro histórico de São Paulo é, simultaneamente, o centróide funcional da rede. As distâncias médias até qualquer outra parada do sistema são minimizadas a partir desse polígono central. Cabe notar, contudo, que a *closeness* exibe pouca variabilidade na rede: o nó com maior *closeness* (0,053) é apenas 1,7 vezes superior à mediana (0,032), enquanto o nó com maior *betweenness* (0,380) é cerca de 2.700 vezes superior à mediana dessa métrica (0,000140). A baixa razão entre máximo e mediana indica que a *closeness centrality* tende a ser pouco discriminativa em redes de transporte de grande escala, conforme já apontado por Mouronte-López (2021).

4.2.4 SOBREPOSIÇÃO ENTRE OS *RANKINGS*

A análise comparativa dos três *rankings* de top 10 revela baixa sobreposição entre as métricas, conforme sintetizado na Tabela 9. Apenas uma parada (Corinthians-Itaquera/CPTM) figura simultaneamente nos *rankings* de *degree* e *betweenness*; duas paradas (Luz e Brás) figuram nos *rankings* de *betweenness* e *closeness*; nenhuma parada figura simultaneamente nos três *rankings*.

Tabela 9 — Sobreposição entre os *rankings* de top 10 das três métricas

Interseção	Paradas em Comum
<i>Degree</i> $\cap$ <i>Betweenness</i>	1 de 10
<i>Degree</i> $\cap$ <i>Closeness</i>	0 de 10
<i>Betweenness</i> $\cap$ <i>Closeness</i>	2 de 10
Nos três <i>rankings</i> simultaneamente	0 de 10

Fonte: Elaborado pelos autores (2026).

A baixa sobreposição confirma empiricamente, no contexto da rede da SPTrans, a recomendação de Palk *et al.* (2026) de combinar múltiplas métricas para a análise de

vulnerabilidade. Cada métrica identifica nós críticos por razões estruturalmente distintas: o *degree* capta *hubs* operacionais (terminais e estações de grande convergência); o *betweenness* identifica pontos de passagem obrigatória entre regiões da rede; o *closeness* destaca a centralidade geográfico-funcional do entorno da Estação da Luz. A não-coincidência entre os *rankings* também corrobora o achado de Song *et al.* (2024), de que a importância topológica de um nó nem sempre corresponde a métricas de centralidade tradicionalmente associadas à magnitude do nó.

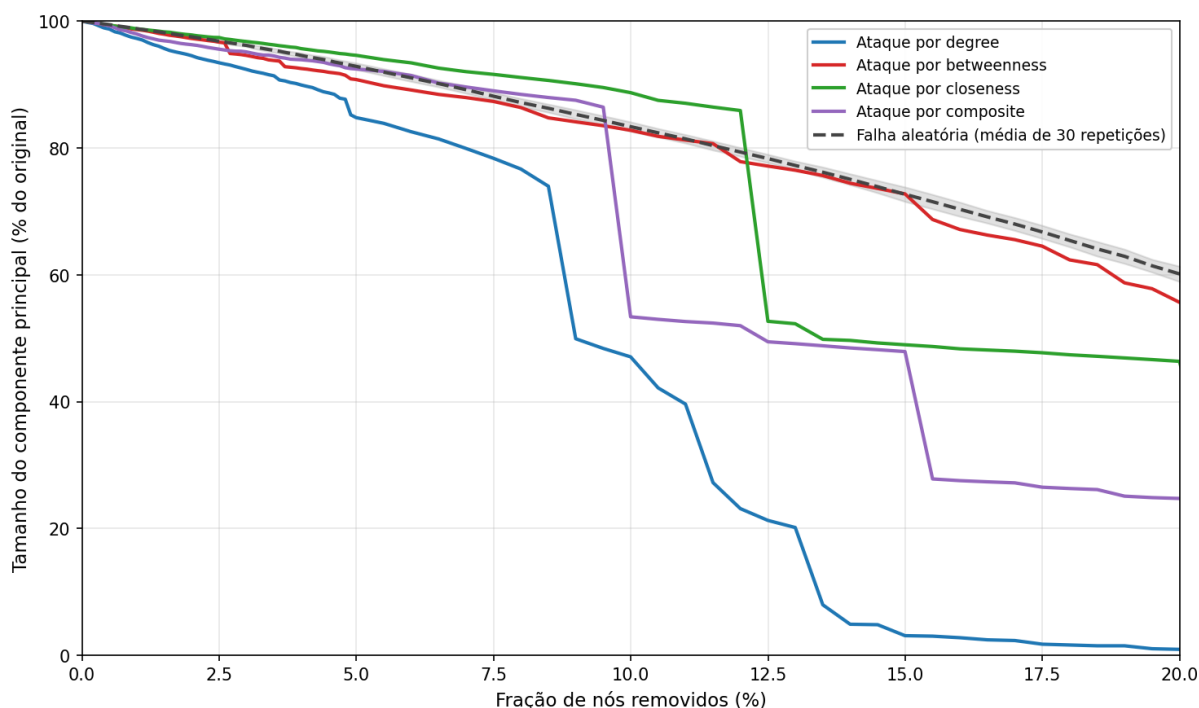
4.3 ANÁLISE DE RESILIÊNCIA: SIMULAÇÃO DE FALHAS

A análise de resiliência avalia como o componente principal integrado responde à remoção progressiva de nós sob dois regimes de perturbação: ataque direcionado por quatro métricas de criticidade e falha aleatória (Subseção 3.4). A apresentação dos resultados está organizada em três frentes: a evolução do tamanho do componente principal (Subseção 4.3.1), a evolução do parâmetro κ de Molloy-Reed (Subseção 4.3.2) e a síntese dos limiares críticos por cenário (Subseção 4.3.3).

4.3.1 FRAGMENTAÇÃO DO COMPONENTE PRINCIPAL

A Figura 4 apresenta a evolução do tamanho do maior componente conexo (LCC) em função da fração de nós removidos, para os cinco cenários simulados. A diferença entre o regime de ataque direcionado e o regime de falha aleatória é acentuada: enquanto a remoção aleatória de 20% dos nós preserva aproximadamente 62% do componente principal, o ataque por *degree* reduz o LCC a menos de 5% para a mesma fração removida.

Figura 4 — Evolução do tamanho do componente principal sob diferentes cenários de remoção



Fonte: Elaborado pelos autores (2026).

Entre as quatro variantes de ataque direcionado, o ataque por *degree* é o mais destrutivo. A 9% de nós removidos, o componente principal já está reduzido à metade; a 13% encontra-se em aproximadamente 30%; a 15% colapsa para cerca de 5%. Os ataques por *closeness* e pelo *ranking* composto produzem padrões semelhantes entre si: ambos exibem um patamar inicial de baixa fragmentação até cerca de 9% de remoção, seguido por uma queda abrupta que estabiliza o LCC em torno de 25% a 45% do tamanho original a partir dos 15%. O ataque por *betweenness*, em contraste, mostra-se o menos destrutivo das quatro variantes direcionadas: degrada o componente principal de forma gradual, e mesmo a 20% de remoção mantém o LCC em torno de 55%.

A menor eficácia do ataque por *betweenness* em relação ao ataque por *degree* merece interpretação. Conforme indicado na Seção 3.4, o *ranking* de criticidade foi calculado no grafo original e mantido fixo ao longo da remoção. Para a métrica *betweenness*, essa decisão tem implicação particular: quando a Estação Brás, com *betweenness* 0,38 no grafo original, é removida, parcela substancial dos caminhos mínimos da rede é redirecionada por nós alternativos, cujo *betweenness* real no grafo modificado torna-se muito superior ao valor que detinham originalmente. O *ranking* fixo, ao não capturar essa redistribuição, conduz a um ataque subótimo. A literatura indica que o ataque por *betweenness* com recálculo a cada

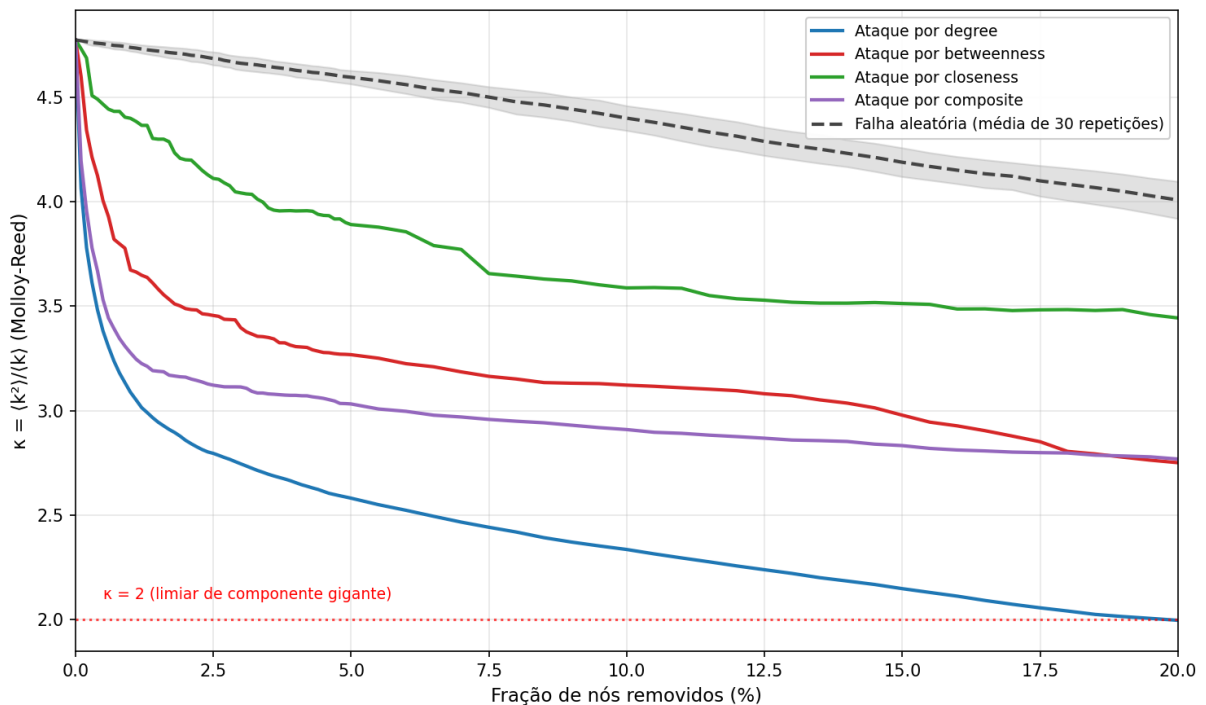
remoção tende a ser substancialmente mais agressivo (HOLME *et al.*, 2002), embora a custo computacional não-trivial. A leitura adequada, portanto, é que o ataque por *betweenness* aqui simulado representa um limite inferior do potencial destrutivo dessa estratégia.

O regime de falha aleatória produz uma curva de degradação aproximadamente linear, com banda de incerteza estreita entre as 30 repetições. Esse padrão confirma o paradigma estabelecido por Albert, Jeong e Barabási (2000): a rede da SPTrans, embora estruturalmente vulnerável a ataques direcionados, é altamente tolerante a falhas distribuídas. A robustez observada no regime aleatório reflete a estrutura esparsa e descentralizada da malha de ônibus, na qual a remoção arbitrária de uma parada raramente compromete a conectividade global.

4.3.2 EVOLUÇÃO DO PARÂMETRO κ DE MOLLOY-REED

A Figura 5 apresenta a evolução do parâmetro $\kappa = \frac{\langle k^2 \rangle}{\langle k \rangle}$ nos cinco cenários. O grafo original apresenta $\kappa = 4,8$, valor consistentemente acima do limiar teórico ($\kappa = 2$) que sinaliza a iminência de fragmentação em grafos aleatórios não correlacionados (MOLLOY; REED, 1995).

Figura 5 — Evolução do parâmetro κ sob diferentes cenários de remoção



Fonte: Elaborado pelos autores (2026).

A análise revela três padrões: Primeiro, o ataque por *degree* é o único cenário que faz κ cruzar o limiar $\kappa = 2$ dentro da janela simulada, atingindo esse valor em torno de 20% de nós removidos. Esse cruzamento coincide com o colapso quase total do componente principal. Segundo, os ataques por *betweenness* e pelo *ranking* composto convergem para $\kappa = 2,75$ no fim da janela, indicando proximidade do limiar, mas sem o cruzar. Terceiro, o ataque por *closeness* mostra-se o menos eficaz em degradar κ entre as variantes direcionadas, mantendo-se em torno de 3,5 ao longo de toda a curva. A falha aleatória degrada κ apenas marginalmente, partindo de 4,8 e mantendo-se acima de 4,0 mesmo a 20% de remoção.

O contraste com os achados da Subseção 4.3.1 é informativo: enquanto o LCC mede a fragmentação observada, κ aproxima uma condição de fragmentação iminente baseada na distribuição de graus. O fato de o ataque por *degree* ser o único a empurrar κ abaixo de 2 reforça que essa estratégia age diretamente sobre a estrutura de conectividade que sustenta o componente gigante, pois ao remover sistematicamente os nós de maior grau, dissolve-se a base estatística que garante a existência de um componente conexo de larga escala.

4.3.3 SÍNTESE: FRAÇÕES CRÍTICAS POR CENÁRIO

A Tabela 10 sintetiza os limiares de fragmentação para cada cenário, expressos como a fração mínima de nós removidos necessária para reduzir o LCC abaixo de 50% e para reduzir κ abaixo de 2. Frações marcadas como "não atingido" indicam que o respectivo limiar não foi alcançado dentro da janela de 20% de remoção.

Tabela 10 — Frações críticas de remoção por cenário

Cenário	Fração crítica para LCC < 50%	Fração crítica para $\kappa < 2$
Ataque por <i>degree</i>	9,0 %	20,0%
Ataque por <i>betweenness</i>	não atingido	não atingido
Ataque por <i>closeness</i>	13,5%	não atingido
Ataque por <i>ranking</i> composto	12,5%	não atingido
Falha aleatória (média de 30 repetições)	não atingido	não atingido

Fonte: Elaborado pelos autores (2026).

A síntese consolida três conclusões de análise de resiliência: a rede da SPTrans é altamente tolerante a falhas aleatórias: a remoção de 20% dos nós ao acaso não basta para

reduzir o componente principal à metade. A estratégia de ataque mais destrutiva é a baseada em *degree*: bastam 9% de nós removidos para fragmentar o componente principal pela metade, e 20% para empurrar a rede para abaixo do limiar de Molloy-Reed. A comparação entre as quatro variantes de ataque direcionado mostra que a combinação de métricas no *ranking* composto não supera o ataque por *degree* isoladamente, indicando que a *degree centrality* é, sozinha, suficiente para orientar a estratégia de remoção de nós mais eficaz na rede analisada.

A combinação desses três achados confirma empiricamente, para a rede da SPTrans, o paradigma estabelecido por Albert, Jeong e Barabási (2000) para redes de distribuição de graus de cauda longa: tolerância elevada a perturbações distribuídas e fragilidade aguda diante de perturbações concentradas em nós de alto grau.

4.4 IMPLICAÇÕES PARA PLANEJAMENTO URBANO

Os achados das Seções 4.1 a 4.3 sustentam um conjunto de implicações para o planejamento e a operação da rede de transporte público da SPTrans. As três principais são articuladas a seguir.

4.4.1 PRIORIZAÇÃO DE PARADAS E ESTAÇÕES PARA PLANOS DE CONTINGÊNCIA

A análise de vulnerabilidade (Subseção 4.2) identificou um conjunto reduzido de paradas e estações que concentram criticidade topológica desproporcional. A Estação Brás, em particular, concentra aproximadamente 38% dos caminhos mínimos da rede com apenas sete vizinhos diretos, configurando-se como nó articulador no sentido descrito por Albert, Jeong e Barabási (2000): sua remoção implicaria reorganização substancial dos caminhos da rede. Luz, Tatuapé, Morumbi e as demais estações no topo do *ranking* de *betweenness centrality* compõem o mesmo grupo.

Esses nós são candidatos naturais para a elaboração de planos de contingência específicos por parte da SPTrans, do Metrô e da CPTM. Eventos previstos, (obras de manutenção, alterações operacionais, grandes eventos urbanos) ou imprevistos (pane elétrica, alagamento, incidente operacional) em qualquer um desses pontos têm potencial de impacto desproporcional sobre o sistema. A análise topológica fornece uma lista nominal e priorizada de nós cujos planos de redundância e desvio merecem detalhamento prévio.

4.4.2 ROBUSTEZ DO SISTEMA DIANTE DE FALHAS DISTRIBUÍDAS

A análise de resiliência (Subseção 4.3) mostrou que a rede da SPTrans é altamente tolerante a falhas aleatórias, pois a remoção arbitrária de 20% dos nós preserva

aproximadamente 62% do componente principal. Esse achado tem implicação positiva para a operação cotidiana: panes localizadas em paradas comuns, que são a maioria da rede, não comprometem a conectividade global do sistema. A estrutura esparsa e descentralizada da malha de ônibus confere ao sistema robustez estatística diante de eventos não concentrados.

A contrapartida é que essa robustez se inverte sob ataques direcionados: bastam 9% de nós removidos por ordem de grau para reduzir o componente principal à metade. A combinação dos dois achados sugere que investimentos em redundância podem ser concentrados em um número pequeno de nós críticos, com retorno desproporcional em termos de resiliência sistêmica.

4.4.3 LIMITES INSTITUCIONAIS DA REDE DA SPTRANS E INTEGRAÇÃO METROPOLITANA

A análise de integração intermodal (Subseção 4.1.3) identificou 48 estações da CPTM, todas localizadas em ramais que se estendem a municípios da Região Metropolitana fora do território da capital, que permanecem desconectadas da malha de ônibus mesmo com raio de transferência ampliado para 600 metros. Esse achado reflete um limite institucional, não metodológico: a SPTrans opera o transporte por ônibus exclusivamente no município de São Paulo, enquanto as linhas da CPTM cruzam fronteiras do município.

Do ponto de vista do planejamento metropolitano, esses 48 pontos representam fronteiras de articulação entre sistemas operados por concessionárias distintas. A análise topológica realizada neste trabalho restringe-se à rede SPTrans, mas evidencia que a integração efetiva do transporte público da Região Metropolitana de São Paulo depende de articulação institucional entre operadores (SPTrans, Metrô, CPTM, EMTU) e operadores municipais dos municípios atendidos pela CPTM. A inclusão dessas fronteiras em análises topológicas futuras requer a integração de *feeds* GTFS de múltiplas operadoras, o que constitui possibilidade de extensão deste trabalho (Subseção 5.4).

5 CONCLUSÃO

Este trabalho aplicou a Teoria dos Grafos à análise topológica da rede de transporte público operada pela SPTrans, com o objetivo de identificar paradas e estações topologicamente críticas e avaliar a resiliência do sistema diante de diferentes regimes de perturbação. As Seções 5.1 a 5.4 a seguir, sintetizam os principais achados, articulam as contribuições do trabalho, reconhecem suas limitações e apontam direções para pesquisas futuras.

5.1 SÍNTESE DOS ACHADOS

A modelagem da rede a partir do *feed* GTFS coletado em 10 de março de 2026 resultou em um grafo com 22.101 nós e 29.540 arestas. A análise dos componentes conexos do grafo bruto revelou 14 componentes desconectados, refletindo a representação topológica não-integrada dos modais no *feed* GTFS. A integração intermodal por meio de arestas de transferência com raio de 400 metros conectou 150 das 198 estações de Metrô e CPTM a paradas de ônibus na vizinhança, resultando no componente principal integrado com 22.099 nós e 32.163 arestas, objeto das análises subsequentes.

A análise de vulnerabilidade identificou três achados centrais: Primeiro, as três métricas de centralidade calculadas (*degree*, *betweenness* e *closeness*) produzem *rankings* de paradas críticas com sobreposição reduzida, com apenas três estações figurando em mais de um *ranking* de top 10 (Brás, Luz e Corinthians-Itaquera) e nenhuma figurando nos três simultaneamente. Segundo, o *ranking* por *betweenness centrality* é dominado por estações da CPTM, com destaque para a Estação Brás, que concentra aproximadamente 38% dos caminhos mínimos da rede com apenas sete vizinhos diretos (configurando-se como nó articulador clássico no paradigma de redes complexas (ALBERT; JEONG; BARABÁSI, 2000)). Terceiro, a *closeness centrality* destaca um aglomerado de paradas no entorno da Estação da Luz, refletindo a centralidade geográfico-funcional do centro histórico paulistano na topologia da rede.

A análise de resiliência confirmou empiricamente, para a rede da SPTrans, o paradigma estabelecido por Albert, Jeong e Barabási (2000) para redes com distribuição de graus de cauda longa: tolerância elevada a falhas aleatórias e fragilidade aguda diante de ataques direcionados. A remoção aleatória de 20% dos nós preserva aproximadamente 62% do componente principal, enquanto o ataque direcionado por *degree centrality* reduz o componente principal a 50% com apenas 9% dos nós removidos. O parâmetro κ de Molloy-Reed cruzou o limiar teórico $\kappa = 2$ apenas sob ataque por *degree*, em torno de 20% de remoção, confirmando que essa estratégia é

a única, entre as quatro variantes testadas, a empurrar a rede para a fragmentação iminente dentro da janela simulada.

A análise da integração intermodal identificou ainda um achado de natureza institucional: 48 estações da CPTM, todas em ramais que se estendem a municípios vizinhos da capital, permanecem desconectadas da malha de ônibus da SPTrans mesmo com raio de transferência ampliado para 600 metros. Esse limite não é metodológico, e sim institucional: reflete a fronteira de atribuição operacional da SPTrans, restrita ao município de São Paulo.

5.2 CONTRIBUIÇÕES

O trabalho oferece contribuições em dois planos: acadêmico e prático. No plano acadêmico, aplica e valida empiricamente, em uma rede metropolitana brasileira de grande escala, o paradigma de robustez estrutural de Albert, Jeong e Barabási (2000) e a recomendação metodológica de Palk *et al.* (2026) de uso combinado de múltiplas métricas de centralidade. A análise contribui para a base empírica da literatura internacional sobre análise topológica de transporte público em contexto latino-americano, sub-representado nos estudos comparativos disponíveis (MASSOBRIO; CATS, 2024).

No plano prático, o trabalho fornece à SPTrans, ao Metrô e à CPTM uma lista nominal e priorizada de paradas e estações cuja remoção produziria maior impacto na conectividade da rede, informação relevante para a elaboração de planos de contingência e para a alocação de investimentos em redundância. A Estação Brás emerge como o nó articulador mais crítico do sistema, candidato natural para detalhamento prévio de planos de redundância e desvio.

A demonstração da viabilidade técnica de conduzir esse tipo de análise a partir de dados abertos no formato GTFS e ferramentas computacionais de uso livre constitui contribuição adicional, particularmente relevante para a replicação da metodologia em outros sistemas de transporte público brasileiros que disponibilizem *feeds* GTFS.

5.3 LIMITAÇÕES

A primeira limitação é temporal, uma vez que a análise se apoia em um único *snapshot* do *feed* GTFS, coletado em 10 de março de 2026, com período de validade declarado entre outubro de 2023 e setembro de 2026. Variações temporais, como extensões de linhas, criação de novos terminais, alterações de itinerário posteriores à coleta, não são contempladas.

A segunda limitação é metodológica e refere-se ao escopo topológico da análise. O trabalho adota representação não-direcionada e não-ponderada do grafo, deliberadamente

abstraindo informações dinâmicas como frequência das viagens, capacidade das linhas e variabilidade dos tempos de percurso. Estudos recentes têm argumentado que a análise puramente topológica fornece uma representação parcial da resiliência do sistema de transporte, devendo ser complementada por abordagens *performance-based* que incorporem demanda, capacidade e tempos reais (TOPAL; DEMIREL; AYDINOGLU, 2026; JAVANMARD *et al.*, 2025). A análise aqui apresentada situa-se na primeira abordagem.

Outra limitação é a institucional: a análise restringe-se ao subsistema operado pela SPTrans, que cobre exclusivamente o município de São Paulo. As 48 estações da CPTM em ramais metropolitanos identificadas na Subseção 4.1.3 permanecem fora do componente principal integrado, refletindo a fronteira institucional da concessão. A análise integrada da rede metropolitana completa requer a articulação de *feeds* GTFS de múltiplas operadoras, fora do escopo deste trabalho.

Por fim, a última limitação é a computacional: a *betweenness centrality* foi calculada por aproximação com amostragem de 500 nós-fonte. Embora prática consolidada em redes de grande escala (MASSOBRIO; CATS, 2024) e com *seed* fixada para reprodutibilidade, valores exatos podem diferir marginalmente dos aproximados. Adicionalmente, as simulações de ataque direcionado utilizaram *ranking* fixo calculado no grafo original, sem recálculo a cada remoção. Para a métrica *betweenness*, conforme discutido na Subseção 4.3.1, essa decisão produz ataques subótimos, e a remoção sequencial com recálculo tenderia a resultados mais agressivos (HOLME *et al.*, 2002).

5.4 TRABALHOS FUTUROS

A primeira extensão deste trabalho é a incorporação de dimensões temporais e dinâmicas à análise. A adoção de GTFS-*Realtime*, padrão complementar ao GTFS estático, permitiria capturar a variabilidade efetiva dos tempos de percurso e atrasos operacionais, abordagem que estudos recentes têm demonstrado capturar fenômenos de acessibilidade invisíveis à análise estática (JAVANMARD *et al.*, 2025).

A segunda, é a integração de *feeds* GTFS de múltiplas operadoras da Região Metropolitana para a construção de um grafo metropolitano completo. Essa integração permitiria reincorporar à análise as 48 estações da CPTM em ramais metropolitanos identificadas como limite institucional do recorte atual.

A terceira é a expansão da análise topológica em direção a abordagens *performance-based*, que incorporem demanda, capacidade e níveis de serviço, conforme advogado por Topal,

Demirel e Aydinoglu (2026). A combinação de análise topológica e *performance-based* forneceria uma caracterização mais completa da resiliência do sistema.

Finalmente, a quarta extensão, é a análise diacrônica, comparando *snapshots* de diferentes períodos para identificar como mudanças na rede, como inaugurações de estações, alterações de linhas, reformas operacionais, afetam suas propriedades topológicas ao longo do tempo.

REFERÊNCIAS

- ALBERT, Réka; JEONG, Hawoong; BARABÁSI, Albert-László. Error and attack tolerance of complex networks. **Nature**, v. 406, n. 6794, p. 378–382, jul. 2000.
- BRANDES, U. A faster algorithm for betweenness centrality. **Journal of Mathematical Sociology**, v. 25, n. 2, p. 163-177, 2001.
- EL-GENEIDY, Ahmed *et al.* New evidence on walking distances to transit stops: identifying redundancies and gaps using variable service areas. **Transportation**, v. 41, n. 1, p. 193–210, 1 jan. 2014.
- FREEMAN, L. C. A set of measures of centrality based on betweenness. **Sociometry**, v. 40, n. 1, p. 35-41, 1977.
- FREEMAN, L. C. Centrality in social networks: conceptual clarification. **Social Networks**, v.1, n.3, p.215-239, 1979.
- HOLME, P.; KIM, B. J.; YOON, C. N.; HAN, S. K. Attack vulnerability of complex networks. **Physical Review E**, v. 65, n. 5, p. 056109, 2002.
- IBGE. **Downloads | Geociências Malhas Municipais** [dataset]. Disponível em: <https://geoftp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2024/>. Acesso em: 10 mar. 2026.
- JAVANMARD, Reyhane *et al.* Using Realtime GTFS to generate easy-to-use transit accessibility measures under travel time uncertainty. **Travel Behaviour and Society**, v. 41, p. 101054, out. 2025.
- MASSOBRIO, Renzo; CATS, Oded. Topological assessment of recoverability in public transport networks. **Communications Physics**, v. 7, n. 1, p. 108, 26 mar. 2024.
- MOBILITYDATA. **General Transit Feed Specification (GTFS) Reference**. Disponível em: <<https://gtfs.org/documentation/schedule/reference/>>. Acesso em: 10 mar. 2026.
- MOLLOY, M.; REED, B. A critical point for random graphs with a given degree sequence. **Random Structures & Algorithms**, v. 6, n. 2-3, p. 161-179, 1995.
- MOURONTE-LÓPEZ, Mary Luz. Analysing the Vulnerability of Public Transport Networks. **Journal of Advanced Transportation**, v. 2021, n. 1, p. 5513311, 2021.
- PALK, Michael; Voß, Stefan; JOVANOVIĆ, Raka. Robustness Assessment of Public Transport Networks in Various Graph Representations: Systematic Review, Decision Support, and Case Study. **Networks**, 2026. DOI: <<https://doi.org/10.1002/net.70021>>.
- SONG, Jinghua *et al.* Assessment and solutions for vulnerability of urban rail transit network based on complex network theory: A case study of Chongqing. **Heliyon**, v. 10, n. 5, 2024. DOI: <<https://doi.org/10.1016/j.heliyon.2024.e27237>>.

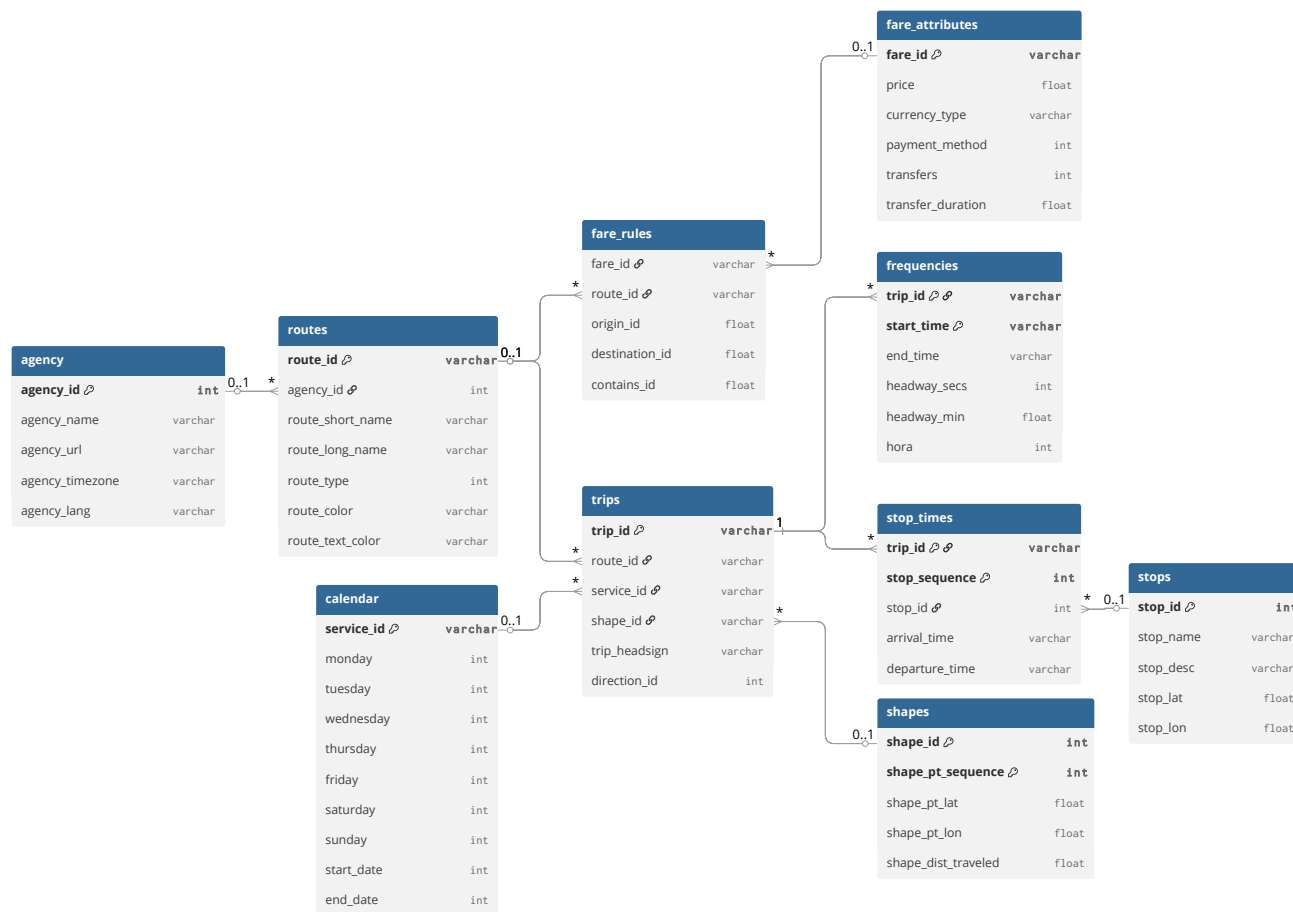
SPTRANS. **Desenvolvedores**. São Paulo: São Paulo Transporte S.A., 2026. Disponível em: <<https://www.sptrans.com.br/desenvolvedores>>. Acesso em: 10 mar. 2026.

SPTRANS. **SPTrans**. Disponível em: <<https://www.sptrans.com.br/sptrans/>>. Acesso em: 26 mar. 2026.

TOPAL, Abdullah Ugur; DEMIREL, Hande; AYDINOGLU, Arif Cagdas. Rethinking the Vision of Transportation Resilience: A Five-Dimensional Framework. **The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XLVIII-4-W19-2025, p. 153–158, 3 mar. 2026.

UN-HABITAT. Metadata on SDGs Indicator 11.2.1: Proportion of population that has convenient access to public transport. Nairobi: **UN-Habitat**, 2018. Disponível em: <https://unhabitat.org/sites/default/files/2020/06/metadata_on_sdg_indicator_11.2.1.pdf>. Acesso em: 7 mai. 2026.

APÊNDICE A - MODELO ENTIDADE-RELACIONAMENTO DO GTFS DA SPTRANS



APÊNDICE B - CÓDIGO-FONTE E NOTEBOOKS

O código-fonte completo desenvolvido neste trabalho encontra-se disponível em repositório público, garantindo a reprodutibilidade integral das análises:

Repositório: <https://github.com/cintia-shinoda/public-transit-graph-theory>

O repositório está organizado em Jupyter Notebooks sequenciais, cada um cobrindo uma etapa do percurso analítico descrito no Capítulo 3:

#	Notebook	Conteúdo
1	1_eda_gtfs.ipynb	Análise exploratória do <i>feed</i> GTFS
2	2_grafo.ipynb	Modelagem do grafo, Integração Intermodal e Análise do Grafo
3	3_vulnerabilidade.ipynb	Métricas de Centralidade (<i>degree</i> , <i>betweenness</i> e <i>closeness</i>) e Identificação de nós críticos
4	4_resiliencia.ipynb	Simulações de ataque direcionado e falha aleatória, Análise de Fragmentação e Cálculo do Parâmetro κ de Molloy-Reed