

# Um Método Centrado em Dados para Melhorar a Segurança de Transportes Coletivos Municipais

Vinicius G. J. Almeida<sup>1</sup>, Thais R. M. B. Silva<sup>1</sup>, Fabrício A. Silva<sup>1</sup>

<sup>1</sup> Instituto de Ciências Exatas e Tecnológicas  
Universidade Federal de Viçosa (UFV) – Campus Florestal  
Florestal, MG – Brasil

{vinicius.jesus, thais.braga, fabricio.asilva}@ufv.br

**Abstract.** *Context-aware routes represent a data-centric area that uses data to offer improvements in a city's flow of vehicles and/or people. Security is one of the contexts that has been highlighted recently, although the literature to date presents solutions mainly for private vehicles. This work proposes a user-centric method for constructing safe routes adapted to the characteristics of buses. Case studies were conducted for the cities of São Paulo and Curitiba, providing routes that avoid places with higher criminal risk, with an increase of, at most, around 12% in their average length, and respecting accessibility for passengers by evaluating roads' inclination.*

**Resumo.** *As rotas cientes de contexto representam uma área centrada em dados que faz uso destes para ofertar melhorias no fluxo de veículos e/ou pessoas em uma cidade. Segurança é um dos contextos que vem se destacando recentemente, embora a literatura até o momento apresente soluções majoritariamente para veículos particulares. Este trabalho propõe um método para construção de rotas seguras adaptado às características de ônibus e centrado em usuário. Estudos de caso foram conduzidos para as cidades de São Paulo e Curitiba, provendo rotas que evitam locais com maior risco criminal, aumentando, no máximo, cerca de 12% no comprimento médio das mesmas, e respeitando a acessibilidade dos passageiros ao avaliar a inclinação das vias.*

## 1. Introdução

A mobilidade urbana, crucial em diversas atividades sociais, descreve a forma de deslocamento realizada por pedestres, ciclistas, usuários de transportes particulares e coletivos, dentre outros, em uma cidade [Aragão 2006]. É notório que, nos últimos anos, a complexidade da mobilidade tem aumentado, impactando negativamente a saúde física e mental dos habitantes de grandes cidades. Uma das formas de se auxiliar no tratamento desta situação é através do uso de soluções de Sistemas de Transporte Inteligente (STI), cujo objetivo, entre outras coisas, é aprimorar o deslocamento de veículos pelas vias de uma cidade [Qureshi and Abdullah 2013].

A mobilidade ciente de contexto pode ser vista como um dos pilares na construção de soluções para STIs. Ela propõe o uso de tipos de dados distintos para oferecer melhorias no deslocamento de veículos, motoristas e passageiros de transportes particulares e coletivos [Vahdat-Nejad et al. 2016]. A partir destes dados, rotas cientes de contexto podem ser construídas e recomendadas, podendo os usuários escolher

o tipo de trajeto desejado, seja ele o mais curto [Madkour et al. 2017], o menos congestionado [Pires et al. 2014], o mais cênico [Denstadli and Jacobsen 2011], dentre outros [Santos et al. 2017].

Dentre os contextos já utilizados na construção de rotas, a segurança tem despertado bastante interesse na literatura nos últimos anos. Neste caso, podem ser construídas e recomendadas rotas seguras, ou seja, rotas que evitam áreas com alto grau de criminalidade. Esse tipo de contexto para rotas pode ser utilizado por diferentes agentes de uma cidade, como pedestres [Boarnet et al. 2005, Bura et al. 2019], ciclistas [Winters et al. 2012], veículos particulares [Santos et al. 2017] e também por veículos públicos, como ônibus urbanos. A falta de segurança para este último caso é um problema importante visto que, apenas no Brasil, foram registrados 53.479 assaltos em ônibus nos últimos 19 anos<sup>1</sup>. No entanto, não existem trabalhos na literatura, até onde os autores puderam investigar, que proponham soluções para construção de rotas seguras para ônibus, em especial considerando a questão do conforto ao acesso por parte dos passageiros.

Neste trabalho é apresentado um método cujo objetivo geral é possibilitar aos gestores municipais do transporte coletivo público identificar e utilizar, se pertinentes, rotas mais seguras para ônibus, com ou sem mudança dos pontos de parada existentes. A solução proposta é centrada em dados, uma vez que constrói rotas com base em diversos repositórios que refletem aspectos da cidade; centrada em usuário, visto que equilibra os ganhos em segurança com os custos de deslocamento para os passageiros; e genérica, já que pode ser instanciada e utilizada por diferentes municípios.

O restante deste trabalho está organizado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados encontrados na literatura; a Seção 3 descreve o método proposto; dois estudos de caso referentes ao uso do método proposto e os resultados podem ser encontrados nas Seções 4 e 5; por fim, a Seção 6 apresenta as conclusões.

## 2. Trabalhos relacionados

Alguns trabalhos da literatura consideram o risco de acidente como um contexto relevante para a determinação das rotas. SafeRNet [Liu et al. 2017] é um *framework* baseado no uso de redes bayesianas que visa diminuir ocorrências de acidentes para veículos particulares. CERVA [Ladeira et al. 2020] é uma solução para oferecer alternativas de rotas baseada em clusterização *Kernel Density Estimation (KDE)* e grafos ponderados. A proposta é baseada na intensidade do tráfego da via juntamente com sua condição de segurança para diferentes janelas temporais. VTq [de Souza and Villas 2020] é um sistema completo para planejamento de rotas com foco em prever a dinâmica futura de fatores urbanos. Todos estes trabalhos, além de possuírem como foco veículos particulares, não apresentaram soluções para segurança como contexto para elaboração de rotas.

Outros estudos apresentam propostas para confecção de rotas para veículos particulares considerando especialmente o índice de segurança. [Félix Mata et al. 2016] apresentam um *framework*, construído através de redes bayesianas e contando com janelas temporais flexíveis, que elabora rotas seguras com base em dados criminais oficiais e do *Twitter*. SafePaths [Galbrun et al. 2016] é um sistema de recomendação de rotas seguras em que os trajetos são gerados a partir de pequenas regiões como ruas e bairros, tendo

<sup>1</sup>NTU: <https://www.ntu.org.br/novo/NoticiaCompleta.aspx?idArea=10idNoticia=1106>

vido utilizadas como base as cidades de Chicago e Filadélfia. Em [Ladeira et al. 2019], o algoritmo *Density Based Spatial Clustering of Application with Noise* (DBSCAN) é utilizado para a clusterização de crimes obtidos para a cidade de Chicago. As rotas seguras são então escolhidas por meio da função *Parzen Window*, capaz de construir rotas seguras escolhendo as vias mais afastadas dos *clusters* criminais. Diferentemente destes trabalhos, este estudo objetiva elaborar rotas seguras para veículos coletivos urbanos.

Por fim, foram encontrados apenas dois trabalhos voltados para veículos coletivos urbanos. [Tompson et al. 2009] apresentam *Hot Routes*, uma solução para ônibus urbanos baseada em clusterização KDE, cujo objetivo é identificar em Londres regiões inseguras, referentes à criminalidade, como pontos de paradas ou trechos do percurso, e exibi-las em um mapa. Esta proposta, diferente do presente trabalho, não propõe novas sugestões de rotas para substituir as identificadas como inseguras. Em [Almeida et al. 2022], foi realizada uma primeira proposta para a construção de rotas seguras para os ônibus da cidade de São Paulo. Porém, ao contrário do mesmo, este trabalho define um método bem elaborado para ser replicado, inclui a inclinação das vias como um critério de impacto na acessibilidade dos passageiros, além de avaliar mais métricas em duas cidades, bem como o impacto socioeconômico da proposta.

### 3. Método para Rotas Alternativas Seguras

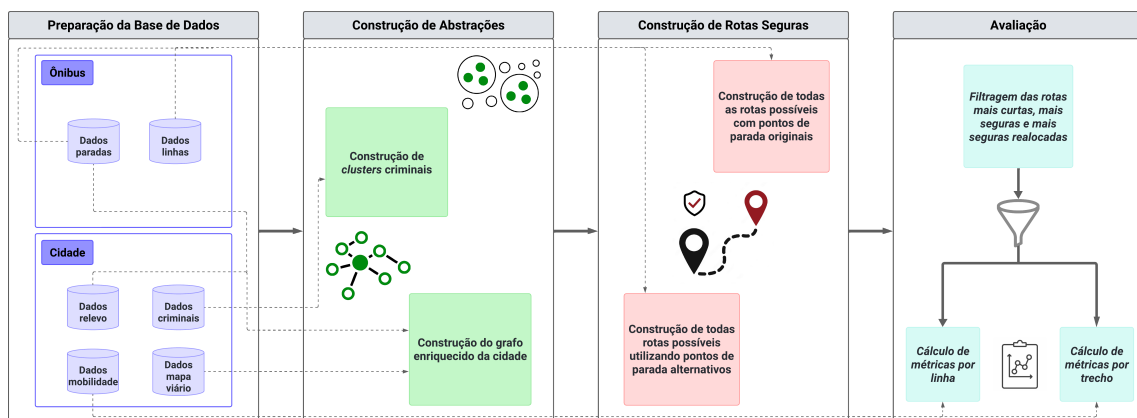


Figura 1. O método proposto e suas etapas.

O método proposto, ilustrado na Figura 1, é dividido em quatro etapas, cada uma delas contendo atividades que permitem a um gestor público construir uma solução personalizada para seu município.

#### 3.1. Preparação da Base de Dados

A primeira etapa do método requer a disponibilidade das bases e respectivos dados mínimos conforme mostrado pela Tabela 1. A coleta e formatação dessas bases varia de município para município, e por isso não é especificada no método proposto.

#### 3.2. Construção de Abstrações

O método necessita de duas abstrações: uma da malha viária da cidade e outra de suas áreas criminais. Para a primeira, deve ser feito uso de uma estrutura de dados que contenha algum tipo de grafo, de preferência orientado, para representar os sentidos das vias

| Base                | Dados   |
|---------------------|---|
| Mapa Viário         | vias (ruas) e conexões entre elas                             |
| Registros Criminais | crimes com descrição (tipo) e localização                     |
| Linhas de Ônibus    | pontos de parada com localização e lista de paradas por linha |
| Relevo              | pontos com localização e altitude                             |
| Mobilidade          | pontos de parada de embarque e desembarque por passageiro     |

**Tabela 1. Dados mínimos necessários para cada base de dados.**

das cidades, uma vez que seu uso facilita o relacionamento entre as ruas e esquinas, sendo facilmente mapeadas como arestas e vértices, respectivamente. Essa abstração permite que informações sejam carregadas de forma a ficarem disponíveis para simular estratégias de deslocamento dentro da cidade posteriormente. Após a construção do grafo, é necessário também enriquecê-lo com dados complementares em seus vértices sobre os pontos de parada existentes e o relevo da cidade.

Para a abstração das áreas inseguras, as informações na base de dados criminais devem ser utilizadas. Porém, normalmente elas são registros únicos, espalhados por toda a cidade. Assim, algoritmos de clusterização são usados para criar agrupamentos e evidenciar tais regiões. É importante, neste caso, utilizar opções cujos *clusters* devem, de alguma forma, ficar próximos (dezenas de metros) da localização dos seus crimes originais agrupados, como é o caso do DBSCAN<sup>2</sup>.

### 3.3. Construção de Rotas Seguras

Para a identificação do grau de periculosidade das possíveis rotas a serem utilizadas pelas linhas de ônibus, foi utilizada a Equação 1 [Ladeira et al. 2019]. Esta leva em consideração o resultado da clusterização de crimes, neste caso, representados como polígonos criminais. Todos os trechos de uma rota  $R_{t(p_i, p_j)}$  entre dois pontos de parada  $(p_i, p_j)$  recebem uma pontuação  $k_{(p_i, p_j)}$ , que quando somadas, representam o valor  $K$ , responsável por indicar o nível de periculosidade da mesma. Desta forma, quanto menor for este valor, mais segura será a rota.

$$k_{(p_i, p_j)} = \frac{1}{m} \sum_{a=1}^m \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x_a - x}{2\sigma^2}\right) \quad (1)$$

Na Equação 1,  $m$  indica a quantidade de vértices (representando esquinas) do trecho da rota entre  $p_i$  e  $p_j$ ;  $\sigma$  é o desvio padrão da quantidade de crimes por *cluster*;  $x_a$  representa a distância de cada vértice em relação ao centro do polígono criminal mais próximo; e  $x$  representa a distância da borda do polígono criminal mais próximo do vértice avaliado em relação ao seu centro. A rota mais segura é construída com base apenas no  $k_{(p_i, p_j)}$  obtido para cada trecho. Desta forma, é escolhida, a cada trecho, a opção que contém o caminho com menor  $k_{(p_i, p_j)}$ .

Quando um ponto de parada de uma linha está dentro de um *cluster* criminal, a segurança da rota é prejudicada, e a solução de identificação de rotas seguras para os coletivos acaba se tornando incoerente. Desta forma, a estratégia utilizada para mitigar este

<sup>2</sup>DBSCAN: <https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html>

problema é considerar a realocação do ponto de parada para outro próximo já existente e que esteja fora de uma área perigosa, caso seja possível. A escolha desta alternativa foi feita para não interferir na infraestrutura da cidade, causando custos extras, além de também não gerar transtornos nas rotinas dos passageiros.

Para realizar as realocações de paradas, foram adotados critérios centrados nos usuários. Com isso, para que um ponto de parada localizado dentro de um polígono criminal seja realocado, a distância deste para o ponto original, em linha reta (distância euclidiana), deve possuir, no máximo, *Raio* metros. Por se tratar de transporte terrestre, também foi considerada a distância real de locomoção pelas vias da cidade, *Distancia*, medida em metros, que resultaria no possível deslocamento real que os passageiros deveriam percorrer a mais, no pior caso, para chegarem até o novo ponto de parada realocado. Os valores de *Raio* e *Distancia* devem ser definidos de acordo com a cidade alvo, e para satisfazer o critério de realocação, basta que o ponto de parada candidato esteja dentro dos limites máximos estabelecidos por esses parâmetros.

Para garantir que os passageiros não terão que lidar com altos aclives/declives até o local do novo ponto de parada realocado, foi utilizado o conceito de impedância, adaptado de [Oliveira and Garcia 2021], originalmente voltado para carrinheiros.

$$i = \begin{cases} \theta^2 \cdot d, & \text{se } \theta > 0 \\ \theta \cdot (-1) \cdot d, & \text{caso contrário} \end{cases} \quad (2)$$

Na Equação 2,  $\theta$  representa o ângulo da via e  $d$  a distância percorrida. Aqui, a ideia utilizada para se obter  $\theta$  foi o uso da aproximação do triângulo retângulo formado pela distância euclidiana percorrida (hipotenusa) em uma via e o ganho de elevação observado (cateto oposto). Com isso, através de trigonometria básica, o ângulo de inclinação pode ser obtido. Logo, é possível obter a impedância  $i$  e então combinar e avaliar diferentes valores, até um limiar coerente com a realidade da cidade alvo.

### 3.4. Avaliação

Esta etapa prevê um processo de filtragem das rotas construídas para cada linha de ônibus da cidade, visando extrair 3 cenários de interesse: Cenário 1 - rotas mais curtas; Cenário 2 - rotas mais seguras mantendo os pontos de paradas originais; Cenário 3 - rotas mais seguras utilizando, se e quando necessário, pontos de parada alternativos. Para cada um desses cenários são avaliadas métricas propostas neste trabalho para apresentar ganhos com segurança e os compromissos em relação à distância e esforço extra necessário. Inicialmente todas as métricas são avaliadas considerando as linhas de ônibus completas. Porém, com a base de dados de fluxo de mobilidade, também é feito o cálculo das mesmas métricas para os trechos mais movimentados da cidade.

## 4. Estudos de Caso

Com o objetivo de avaliar e validar o método proposto, foram feitos estudos de casos para duas cidades: São Paulo e Curitiba. Tais cidades apresentam características distintas em termos de tamanho da malha viária, número de linhas de ônibus e pontos de parada, quantidade de crimes e relevo.

**Preparação da Base de Dados**<sup>3</sup> - As informações dos mapas viários das cidades foram obtidas por meio da biblioteca Python **OSMnx**<sup>4</sup>. As informações de relevo foram obtidas pelo INPE (Instituto Nacional de Pesquisas Espaciais)<sup>5</sup>. Os dados nesta base são distribuídos através de quadrículas, cada uma contendo diversas linhas, que representam diferentes pontos de relevo de uma região.

Para a cidade de São Paulo, os dados referentes às linhas de ônibus foram obtidos de forma gratuita e aberta através da plataforma **InterSCity**<sup>6</sup>. Já para a cidade de Curitiba, os dados, também gratuitos, foram obtidos através da plataforma **URBS**<sup>7</sup>. São Paulo conta com 2.069 linhas, das quais 1.987 puderam ser aproveitadas e, Curitiba com 676, das quais 629 foram utilizadas. Em São Paulo, as linhas possuem em média 42 pontos de parada contra 27 nas de Curitiba.

Em relação aos dados criminais, para a cidade de São Paulo, foi possível obtê-los na plataforma da **Secretaria de Estado da Segurança Pública**<sup>8</sup>. O ano de 2019 foi escolhido por ser o último ano completo de dados sem interferências causadas pela pandemia da COVID-19, e a categoria escolhida foi a de crimes que envolvem furtos e roubos, pois essas impactam significativamente o cotidiano dos passageiros. Foi feita uma limpeza dos dados vazios, de crimes pertencentes às regiões metropolitanas, e de colunas desnecessárias para o método, considerando como informação válida apenas detalhes do crime, data, endereço e latitude/longitude. Também foi necessário aplicar filtros para manter apenas crimes com rubricas relacionadas ao contexto de transporte público, envolvendo situações ligadas ao trânsito e aos passageiros. Ao final desse processo, foram obtidos 212.079 registros criminais geolocalizados. Os registros criminais de Curitiba foram obtidos da plataforma oficial de dados abertos do **Sistema de Gestão da Guarda Municipal de Curitiba**<sup>9</sup>. Todas as etapas de limpeza e filtros por categorias de crimes mencionadas para os dados de São Paulo também foram realizadas, resultando em uma base com 12.448 registros criminais geolocalizados.

Os dados de mobilidade de São Paulo, referentes aos registros de viagens de um único dia típico na cidade, bem como o mapeamento da rede de transportes que operou neste dia, foram fornecidos pela **SPTTrans**<sup>10</sup>. Para estimar o fluxo de mobilidade dos passageiros foi utilizada a sequência de transações dos usuários, obtida através do bilhete único da cidade. Para a cidade de Curitiba, os dados de mobilidade foram cedidos pelo **Serviço Digital de Curitiba**<sup>11</sup>. Como os registros possuíam apenas a partida do passageiro, um funcionário da URBS, que auxiliou este trabalho, sugeriu que os dados de mobilidade fossem gerados com base em uma inferência, respeitando as seguintes regras: apenas bilhetes possuindo somente dois registros diários, com um intervalo de pelo menos 4 horas, poderiam ser tratados como um possível fluxo de origem e destino.

Ao fim deste processo de preparação, todos os dados manipulados foram organi-

<sup>3</sup>Base de Dados disponível em (Omitido para Revisão)

<sup>4</sup>OSMnx: <https://osmnx.readthedocs.io/en/stable/>

<sup>5</sup>TOPODATA: <http://www.dsr.inpe.br/topodata/acesso.php>

<sup>6</sup>InterSCity: [https://interscity.org/open\\_data/](https://interscity.org/open_data/)

<sup>7</sup>Urbanização de Curitiba S/A: <https://www.urbs.curitiba.pr.gov.br/>

<sup>8</sup>SSP: <http://www.ssp.sp.gov.br/transparenciassp/Consulta2022.aspx>

<sup>9</sup>SiGesGuarda: <https://www.curitiba.pr.gov.br/dadosabertos/>

<sup>10</sup>SPTTrans: <https://www.sptrans.com.br/sptrans/>

<sup>11</sup>Serviço Digital de Curitiba: <https://servicodigital.curitiba.pr.gov.br/>

zados e armazenados em arquivos de texto (CSV). Esses arquivos foram salvos no repositório privado do projeto no **GitLab**<sup>12</sup>.

**Construção de Abstrações** - Para o agrupamento dos registros criminais das duas cidades, o algoritmo DBSCAN foi utilizado. Este algoritmo requer dois parâmetros para seu funcionamento:  $\varepsilon$  (raio entre vizinhos) e  $v$  (quantidade mínima de vizinhos), cujos valores, empiricamente obtidos, utilizados para a cidade de São Paulo foram  $v = 14$  e  $\varepsilon = 85m$ , resultando em 2.028 *clusters* criminais. Já para Curitiba, cidade em que a quantidade de crimes registrados é substancialmente menor, os valores  $v = 3$  e  $\varepsilon = 250m$  foram usados, gerando 360 *clusters*. Vale destacar que, para a cidade de São Paulo, os dois maiores *clusters* obtidos, ambos na região do centro, foram removidos como *outliers*, uma vez que com a sua presença se tornava inviável a construção de qualquer tipo de caminho alternativo para os ônibus nesta localidade. Para representar os *clusters* criados no mapa das cidades no formato de polígonos foi utilizada a conversão **Convex Hull**<sup>13</sup>, que consiste em gerar, para um conjunto de pontos clusterizados, através do menor número possível de pontos externos, um polígono convexo que contenha todos os outros pontos dentro. Ao final, o número de polígonos foi de 1.961 para São Paulo e 136 para Curitiba.

O enriquecimento do grafo foi feito utilizando a ferramenta OSMnx, que já oferece a estrutura de dados de um grafo da malha viária das cidades, no qual foram incluídas informações sobre relevo e pontos de parada, mapeados nos nós das esquinas mais próximas. A partir dela podem ser utilizados diversos algoritmos para grafos já disponibilizados.

**Construção de Rotas Seguras** - Para cada uma das linhas de ônibus, de cada uma das cidades, foram identificadas as rotas para os Cenários 1 a 3 (Seção 3.4), conforme estratégia descrita na Seção 3.3. Para o Cenário 3, conforme necessário, foi definido *Raio* = 250m e *Distancia* = 1000 para ambas as cidades. Já para o valor de impedância, foram feitos testes empíricos que consideraram o uso de 180 metros de deslocamento e 9 metros de ganho de elevação, registrado em uma via movimentada e frequentada por diferentes tipos de pessoas e idades. Com isso, o ângulo obtido foi de aproximadamente 2,87°. Desta forma, utilizando o percurso completo, ou seja, percorrendo os 180 metros no trajeto inclinado, a impedância obtida através da Equação 2 foi de aproximadamente 1.500.

**Avaliação** - Conforme previsto pelo método proposto, métricas de avaliação foram geradas tanto para as rotas dos Cenários 1 a 3 para as linhas de ônibus completas das duas cidades, como também para os trechos considerados mais relevantes, de acordo com dados sobre mobilidade urbana utilizados. O Cenário 1 (rota mais curta) foi considerado uma referência neste trabalho, assumindo que os ônibus tradicionalmente se deslocam pela rota mais curta. Sendo assim, para cada métrica, os resultados dos Cenários 2 e 3 são apresentados relativamente ao Cenário 1. Para cada cidade e cada métrica, são calculados os valores de média, mediana e máximo referentes aos valores obtidos para todas as linhas ou todos os trechos considerados.

## 5. Resultados

Na Seção 5.1 são apresentadas métricas calculadas para a linha completa, do início ao fim. Na Seção 5.2 são calculadas métricas considerando apenas trechos mais movimentados.

<sup>12</sup>GitLab - Gerenciador de repositório de software baseado em git: <https://gitlab.com/>

<sup>13</sup><https://www.geeksforgeeks.org/convex-hull-using-divide-and-conquer-algorithm/>

| Métrica | Cidade    | Cenário 2 |       |         | Cenário 3 |        |         |
|---------|-----------|-----------|-------|---------|-----------|--------|---------|
|         |           | máximo    | média | mediana | máximo    | média  | mediana |
| M1      | São Paulo | n.a.      | n.a.  | n.a.    | 71,88%    | 20,00% | 17,78%  |
|         | Curitiba  | n.a.      | n.a.  | n.a.    | 50,00%    | 4,00%  | 1,49%   |
| M2      | São Paulo | n.a.      | n.a.  | n.a.    | 1.499,07  | 390,00 | 365,37  |
|         | Curitiba  | n.a.      | n.a.  | n.a.    | 1.249,57  | 148,00 | 0,00    |
| M3      | São Paulo | -1.649m   | -84m  | 0m      | -2.661m   | -334m  | -222m   |
|         | Curitiba  | -1.608m   | -14m  | 0m      | -1.608m   | -52m   | 0m      |
| M4      | São Paulo | 80,09%    | 5,78% | 4,36%   | 82,46%    | 11,11% | 9,17%   |
|         | Curitiba  | 65,87%    | 4,48% | 2,14%   | 52,00%    | 4,45%  | 2,75%   |
| M5      | São Paulo | 164,84%   | 5,72% | 4,79%   | 152,55%   | 15,54% | 14,21%  |
|         | Curitiba  | 140,87%   | 5,49% | 1,65%   | 140,87%   | 5,53%  | 2,37%   |
| M6      | São Paulo | 272,00m   | 4,26m | 2,34m   | 272,00m   | 6,88m  | 7,41m   |
|         | Curitiba  | 191,00m   | 9,00m | 3,48m   | 191,00m   | 12,00m | 8,38m   |
| M7      | São Paulo | n.a.      | n.a.  | n.a.    | 326,00m   | 18,35m | 16,56m  |
|         | Curitiba  | n.a.      | n.a.  | n.a.    | 69,59m    | 2,55m  | 0,00m   |

Tabela 2. Resultados consolidados dos cenários 2 e 3 relativos ao cenário 1.

Por fim, é feita uma análise socioeconômica do impacto do método na Seção 5.3.

### 5.1. Métricas para Linhas Completas

A Tabela 2 consolida os resultados das sete métricas apresentadas abaixo, através das estatísticas (máximo, média e mediana), para os cenários 2 e 3 relativamente ao cenário 1 (rota mais curta).

**M1: Pontos de Parada Removidos de *Clusters*** - São Paulo e Curitiba contam, em média, com 27,8% e 6,5%, respectivamente, dos pontos de parada de uma linha dentro de *clusters* criminais. Em média, foram removidos 7 pontos de parada por linha (totalizando 20% dos pontos) para São Paulo, e 0,8 pontos por linha (totalizando 4%) para Curitiba. Isso mostra que as novas rotas param menos em *clusters* criminais.

**M2: Impedância Média dos Pontos de Parada Realocados** - Embora o valor máximo permitido de acréscimo de impedância para os pontos de parada tenha sido definido como 1.500, na prática, os resultados para o Cenário 3 foram bem menores, sendo em média obtido um acréscimo de 390 para os novos pontos de parada em São Paulo (aproximadamente 4x menor que o limite) e de 148 para os de Curitiba (mais de 10x menor). Ou seja, a realocação dos pontos de parada não causaram impacto de acessibilidade em termos de inclinação do deslocamento extra dos passageiros do ponto antigo ao novo.

**M3: Soma das distâncias sequenciais dentro dos mesmos *clusters*** - Esta métrica contabiliza as distâncias de pedaços do trajeto onde pares de vértices (adjacentes) do grafo estão contidos dentro de um mesmo *cluster* criminal. Para cada linha de ônibus, a métrica foi calculada para os 3 cenários e, ao final do processo, foram calculadas estatísticas dessas distâncias por cenário, considerando todas as linhas. Conforme a Tabela 2, para São Paulo a redução média foi de 84 metros, enquanto para o cenário 3 (rota mais segura realocada), ela foi ainda melhor, de 334 metros. Em Curitiba, os resultados também seguiram o mesmo comportamento, com redução média de 14 metros para o cenário 2, e 52 metros para o cenário 3. Isso mostra que os ônibus percorreram menos distâncias dentro



de regiões inseguras.

**M4: Acréscimo Total das Rotas** - Uma vez que o fator com maior prioridade dos Cenários 2 e 3 é a segurança dos passageiros, medida pela variável  $K$ , evitar regiões perigosas é mais importante do que o comprimento da rota. Com isso, é possível que esses cenários apresentem aumentos em relação ao comprimento da rota original, definida como mais curta. Como principais resultados, em São Paulo, o Cenário 2 acrescenta em média 1.125 metros (correspondente a 5,78%) ao comprimento original da rota mais curta. Já o Cenário 3 possui uma média de acréscimo maior, de 2.164 metros, que equivale à 11,11%. Para Curitiba, o acréscimo médio para o cenário 2 foi de 597 metros (4,48% em relação à rota mais curta), enquanto o do cenário 3 alcançou 593 metros (4,45%). Isso mostra que o comprimento das rotas aumentou em média, entre 4% e 11%, o que pode ser considerado um compromisso para manter as rotas mais seguras.

**M5: Acréscimo Médio dos Trechos de cada Rota** - Esta métrica considera os deslocamentos entre pares de pontos de parada de uma rota. A partir desses valores, é feita a média para descobrir o comprimento médio dos trechos de cada rota. Para São Paulo, o comprimento médio dos trechos para o cenário 1 é de aproximadamente 491,84m. O acréscimo médio por trecho da rota é de 28,18 metros para o cenário 2 (5,73% em relação ao trecho médio do cenário 1) e 76,41 metros para o cenário 3 (15,54%). Na prática, o acréscimo de tempo para o passageiro não seria muito significativo visto que um ônibus na cidade de São Paulo, em horário de pico, consegue se locomover, na média, a 16km/h<sup>14</sup>. Com isso, no pior caso, considerando um passageiro que embarca no primeiro ponto de parada e desembarca apenas no último, o acréscimo de tempo para o cenário 3 seria de aproximadamente 11 minutos. Já em Curitiba, os resultados também foram semelhantes aos apresentados na métrica **Acréscimo Total das Rotas**, apresentando um aumento médio de 41,09 metros (5,49%) para o cenário 2 e 41,4 metros (5,53%) para o cenário 3, tendo, em média, 748,46m de comprimento médio do trecho para o cenário 1. O acréscimo de tempo para o passageiro de Curitiba seria de aproximadamente 2,5 minutos, considerando uma velocidade média de 26,7km/h<sup>15</sup>.

**M6: Distância dos Nós em Relação aos Centros dos *cluster*** - A partir desta métrica é possível observar que com o método proposto, os ônibus são capazes de distanciar as esquinas que percorrem dos centros de *clusters* mais próximos. Para São Paulo, em média para o cenário 2, os ônibus se distanciam em 4,26 metros dos centros dos *clusters*, e 6,88 metros para o cenário 3, chegando a mais de 200m em alguns casos. Esse resultado ressalta que além das sugestões de realocação dos pontos de paradas, o trajeto todo também está mais seguro. Para Curitiba, as rotas tiveram um distanciamento médio de 9 metros para o cenário 2 e 12 metros para o cenário 3, chegando a mais de 100m em alguns casos.

**M7: Distância dos Pontos de Parada em Relação aos Centros dos *cluster*** - Para avaliar a qualidade do cenário de realocação, também foi criada a mesma métrica exibida anteriormente, porém, considerando apenas os pontos de paradas nas rotas. Para a cidade de São Paulo, o ganho médio de distanciamento dos pontos de paradas dos centros dos *clusters* no cenário 3 foi de 18,35 metros. Já para Curitiba, o ganho foi menor, totalizando

<sup>14</sup><https://www.mobilize.org.br/noticias/11963/velocidade-media-dos-onibus-fica-estagnada-em-16-kmh-na-cidade-de-sp.html>

<sup>15</sup><https://summitmobilidade.estadao.com.br/compartilhando-o-caminho/onibus-de-curitiba-como-a-cidade-popularizou-o-brt/>

2,55 metros em média, uma vez que a média de pontos de paradas das linhas localizados dentro de *clusters* e o número de realocações efetivamente realizadas são significativamente menores do que para São Paulo.

## 5.2. Métricas por Trechos Mais Movimentados

O método proposto também prevê a realização de avaliações considerando apenas as partes mais relevantes do trajeto de cada linha, ou seja, as regiões de cada rota que mais poderiam impactar os usuários do sistema de transporte público. Essas regiões foram definidas utilizando a base de dados de mobilidade da cidade, como os trajetos que mais recebem fluxos de pessoas. Para cada linha foi mapeada a quantidade de pessoas que se locomoveram entre seus trechos de modo a ser possível ranqueá-los, selecionando os mais movimentados.

**Soma das distâncias sequenciais dentro dos *clusters*** - Na cidade de São Paulo, os ônibus, em média, deixaram de andar 5,98 metros dentro de *clusters* criminais para o cenário 2. Já para o cenário 3, o ganho foi ainda maior, sendo de 16,77 metros, em média. Para Curitiba os resultados também foram de melhoria, sendo o ganho de 0,5 metros em média para o cenário 2 e 13,02 metros para o cenário 3. Ou seja, nos trechos mais movimentados, os ônibus também reduziram o percurso em regiões criminais.

**Acréscimo Total das Rotas** - Foram adicionados, em média, 30,26 metros (4% a mais) para as rotas do cenário 2, e 91 metros (12,14%) para as do cenário 3, na cidade de São Paulo. Em Curitiba, em média são 45,6 metros (5,74%) para as rotas do cenário 2, e 45,66 metros (5,39%) para as do cenário 3. Isso mostra que os trechos mais movimentados não sofreram significativamente com o aumento do comprimento nas duas cidades.

**Distância dos Nós em Relação aos Centros dos *clusters*** - Assim como na avaliação para rotas completas, a maior melhoria nesta métrica também se manteve para o cenário 3. Para São Paulo, o distanciamento médio foi de 5,34 metros para o cenário 2 e 18,82 metros para o cenário 3. Ou seja, os ônibus em geral passam mais longe dos centros dos *clusters* criminais. Em Curitiba, a média do cenário 2 foi de 5,81 metros, e para o cenário 3, o distanciamento médio foi de 13,43 metros.

## 5.3. Avaliando Fatores Socioeconômicos

Apesar da avaliação apresentar um equilíbrio satisfatório entre o aumento de segurança das linhas de ônibus e a distância e o esforço adicionais necessários, é ainda importante avaliar se a distribuição dos polígonos criminais pelas cidades está concentrada, em grande maioria, em regiões de classes sociais desfavorecidas, levando a algum tipo de viés em relação aos locais que receberam os maiores números de sugestões de troca.

A avaliação realizada utilizou a base de dados pública do censo IBGE<sup>16</sup>. Os centroides de todos os *clusters* criminais das cidades foram mapeados em delimitações territoriais chamadas setores censitários, conforme utilizado pelo IBGE. Em seguida, foram agrupados por classe social para se entender a relação estabelecida entre os dados socioeconômicos e as regiões inseguras das cidades. Em São Paulo, aproximadamente 48% dos polígonos estão localizados em setores classificados como classe média, 30% classe média-alta, 20% classe alta e apenas 2% classe baixa. Já em Curitiba, tem-se que 40%

<sup>16</sup><https://censo2010.ibge.gov.br/resultados.html>

dos polígonos estão contidos em setores de classe média, 37% classe média-alta e 23% classe alta, não havendo nenhum em setores de classe baixa.

Com o mapeamento da distribuição dos polígonos pelos setores censitários das duas cidades, foi realizado o levantamento das informações socioeconômicas de seus pontos de paradas. Nesta análise, dentre todos os pontos de paradas das duas cidades, apenas aqueles que foram efetivamente realocados serão considerados. Com isso, é possível mapear em quais classes predominaram as substituições, além de ser possível fazer um levantamento das principais regiões das cidades que receberam tais modificações. Em São Paulo, 78% dos pontos realocados estavam situados em locais de classe média, 20% em setores de classe alta e apenas 2% de classe baixa. Em Curitiba grande parte dos pontos realocados também estão dentre as classes melhor situadas economicamente, sendo aproximadamente 47% classe alta, 31,5% classe média-alta, e 21,5% classe média.

Para complementar o resultado, também foram observadas as quantidades de pontos de paradas mantidos em suas respectivas classes sociais. Para as cidades de São Paulo e Curitiba, cerca de 79% e 82%, respectivamente, de todos os pontos de paradas realocados foram mantidos nos mesmos setores a que pertenciam. Dentre os que sofreram mudanças, foi possível perceber que existem diversas combinações de trocas, sendo possível que um ponto de uma região melhor localizada socialmente fosse para uma de classe economicamente menor e vice-versa.

Esses resultados mostram que o método proposto não possui viés para construir os polígonos e distribuí-los pela cidade, pois como pode ser observado, a grande maioria das ocorrências se encontram em regiões com melhores índices socioeconômicos. Além disso, em relação à parte que de fato impacta os passageiros, realocação de pontos de paradas, percebe-se que para ambas cidades, cerca de 80% das realocações foram mantidas em suas respectivas classes sociais, e além disso, dentre as que foram realocadas, todas essas passaram por todos os requisitos que restringem significativamente a descaracterização do ponto de parada original.

## 6. Conclusão e Trabalhos Futuros

Este trabalho apresentou um método para elaboração de rotas alternativas seguras para sistemas de transporte coletivo urbano. Esta solução, centrada em dados, visa melhorar as condições de segurança dos usuários do transporte público, utilizando 5 diferentes bases de dados. As novas rotas seguras propostas apresentam uma alternativa centrada em usuário, uma vez que procuram diminuir a exposição a áreas com altos níveis de criminalidade, controlando prejuízos no que diz respeito às distâncias a serem percorridas tanto pelos ônibus em si como pelos passageiros e levando em consideração o relevo no esforço de deslocamento destes.

O método foi avaliado para duas cidades, São Paulo e Curitiba, utilizando-se dados reais das mesmas. Não só foi possível aplicar sistematicamente as mesmas etapas e suas respectivas atividades previstas pelo método nos dois casos, como os resultados apresentados pelas métricas utilizadas indicaram que, em ambas as cidades, o nível de segurança dos passageiros dentro e fora dos veículos melhorou, sem que eles tenham que realizar deslocamentos desgastantes no embarque ou desembarque.

Desta forma, percebe-se que este é um método que visa apoiar prefeituras, caso

elas decidam modificar rotas e pontos de parada utilizando a segurança e o conforto dos passageiros como critérios.

Como trabalhos futuros, vislumbra-se a aplicação do método proposto em um maior número de cidades, com perfis ainda mais diversificados, além de se avaliar a possibilidade de uso desta solução junto à sistemas autônomos de transporte coletivo.

Além disso, a inclusão de novas informações contextuais nas cidades, como dados de densidade populacional, pode ser interessante para restringir ainda mais os critérios de realocação de pontos de paradas. Isso não apenas aumentaria a viabilidade e a cobertura das novas rotas, mas também reduziria possíveis impactos sociais adversos.

Por fim, construir um piloto do método proposto com o objetivo de coletar *feedbacks* direto dos passageiros e gestores de transporte para validar a aplicabilidade e eficácia do método em condições reais.

### Agradecimentos

Os autores agradecem pelo auxílio financeiro fornecido pela CAPES para a realização deste trabalho.

### Referências

- Almeida, V., Silva, T., and Silva, F. (2022). Se for, vá na paz: Construindo rotas seguras para veículos coletivos urbanos. In *Anais do XL Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 140–153, Porto Alegre, RS, Brasil. SBC.
- Aragão, J. (2006). Curso de gestão integrada da mobilidade urbana. *Secretaria Nacional de Transporte e Mobilidade Urbana da Universidade de Brasília, Brasília, DF*.
- Boarnet, M. G., Day, K., Anderson, C., McMillan, T., and Alfonzo, M. (2005). California’s safe routes to school program: impacts on walking, bicycling, and pedestrian safety. *Journal of the American Planning Association*, 71(3):301–317.
- Bura, D., Singh, M., and Nandal, P. (2019). Predicting secure and safe route for women using google maps. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, pages 103–108.
- de Souza, A. and Villas, L. (2020). Vem tranquilo: Rotas eficientes baseado na dinâmica urbana futura com deep learning e computação de borda. In *Anais do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 351–364, Porto Alegre, RS, Brasil. SBC.
- Denstadli, J. M. and Jacobsen, J. K. S. (2011). The long and winding roads: Perceived quality of scenic tourism routes. *Tourism Management*, 32(4):780–789.
- Félix Mata et al., Miguel Torres-Ruiz, G. G. R. Q. R. Z.-F. M. M.-I. E. L. (2016). A mobile information system based on crowd-sensed and official crime data for finding safe routes: A case study of mexico city. *Mobile Information Systems*, 2016:1–11.
- Galbrun, E., Pelechris, K., and Terzi, E. (2016). Urban navigation beyond shortest route: The case of safe paths. *Information Systems*, 57:160–171.
- Ladeira, L., de Souza, A., Silva, T., Filho, G. R., Peixoto, M. M., and Villas, L. (2020). Cerva: Roteamento contextual para veículos com risco espaço-temporal. In *Anais*

- do XXXVIII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, pages 379–392, Porto Alegre, RS, Brasil. SBC.
- Ladeira, L., Souza, A., Pereira, G., Silva, T. H., and Villas, L. (2019). Serviço de sugestão de rotas seguras para veículos. In *Anais do XXXVII Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 608–621, Porto Alegre, RS, Brasil. SBC.
- Liu, Q., Kumar, S., and Mago, V. (2017). Safernet: Safe transportation routing in the era of internet of vehicles and mobile crowd sensing. *2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC)*.
- Madkour, A., Aref, W. G., Rehman, F. U., Rahman, M. A., and Basalamah, S. (2017). A survey of shortest-path algorithms. *arXiv preprint arXiv:1705.02044*.
- Oliveira, M. V. R. and Garcia, I. C. (2021). Sugestões de rotas personalizadas para caminhões na coleta seletiva de materiais recicláveis. In *Brazilian Symposium on Computer Networks and Distributed Systems*.
- Pires, C., Souza, W., Remboski, T., Aguiar, M., and Ferreira, P. (2014). Acoroute: Sistema de recomendação de rotas evitando congestionamento baseado em comunicação por feromônios. In *Workshop de Computação Aplicada a Gestão do Meio Ambiente e Recursos Naturais*, pages 1–4, Porto Alegre, RS, Brasil. SBC.
- Qureshi, K. N. and Abdullah, A. H. (2013). A survey on intelligent transportation systems. *Middle-East Journal of Scientific Research*, 15(5):629–642.
- Santos, F. A., Rodrigues, D. O., Silva, T. H., Loureiro, A. A. F., and Villas, L. A. (2017). Rotas veiculares cientes de contexto: Arcabouço e análise usando dados oficiais e sensorizados por usuários sobre crimes. In *Anais do XXII Workshop de Gerência e Operação de Redes e Serviços*, Porto Alegre, RS, Brasil. SBC.
- Tompson, L., Partridge, H., and Shepherd, N. (2009). Hot routes: Developing a new technique for the spatial analysis of crime. *Crime Mapping: A Journal of Research and Practice*, 1(1):77–96.
- Vahdat-Nejad, H., Ramazani, A., Mohammadi, T., and Mansoor, W. (2016). A survey on context-aware vehicular network applications. *Vehicular Communications*, 3:43–57.
- Winters, M., Babul, S., Becker, H., Brubacher, J. R., Chipman, M., Cripton, P., Cusimano, M. D., Friedman, S. M., Harris, M. A., Hunte, G., et al. (2012). Safe cycling: how do risk perceptions compare with observed risk? *Canadian journal of public health*, 103:S42–S47.