

Se for, vá na paz: Construindo Rotas Seguras para Veículos Coletivos Urbanos

Vinícius G. J. Almeida¹, Thais R. M. B. Silva¹, Fabrício A. Silva¹

¹ 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 mobility can be defined as the use of different types of data to offer improvements for both drivers and passengers, as well as to the flow of vehicles. In this scenario, one of the services offered will be the use of models built based on such data, to trace routes to be used by different types of vehicles. Among the possible contexts to be considered, safety is one that has been explored recently. Although the works on safe routes found so far have evaluated different solutions, they were mostly tested for private vehicles and evaluated from a computational point of view. This work proposes RACIONAL, a safety route solution for municipal buses, built on well-accepted fundamentals in the literature for private vehicles, but adapted to the specific characteristics of collectives, such as the need to consider bus stops and the impossibility of carrying out constant temporal changes. The evaluation results show that, with few spatio-temporal adjustments, it is possible to offer passengers routes that maintain the coverage of the public transport service, while improving their safety, avoiding areas considered dangerous, without imposing significant increase in distance travelled.*

Resumo. *A mobilidade ciente de contexto pode ser definida como o uso de diferentes tipos de dados para a oferta de melhorias a motoristas, passageiros e para o fluxo de veículos. Um dos serviços relacionados à este cenário é a construção de modelos com base em tais dados para traçar rotas a serem utilizadas por diversos tipos de veículos. Dentre os possíveis contextos a serem considerados, a segurança é um que vem sendo explorado recentemente. Embora os trabalhos sobre rotas seguras encontrados até o momento tenham avaliado variadas soluções, as mesmas foram testadas majoritariamente para veículos particulares e avaliadas sob o ponto de vista computacional. Este trabalho propõe RACIONAL, uma solução de cálculo de rotas seguras para ônibus municipais, construída com base em fundamentos bem aceitos da literatura para veículos particulares, porém adaptados às características específicas dos coletivos, tais como a necessidade de se considerar os pontos de parada e a inviabilidade do uso de mudanças temporais constantes. Os resultados obtidos mostram que, com poucos ajustes espaço-temporais, é possível encontrar rotas que mantêm a cobertura do serviço de transporte público, ao mesmo tempo em que melhora a segurança das mesmas, evitando áreas consideradas perigosas, sem impor aumento significativo da distância percorrida.*

1. Introdução

Devido aos recentes avanços nas áreas de Internet das Coisas (IoT) e sistemas distribuídos, dispositivos capazes de coletar dados de forma prática e eficiente passaram a fazer parte da rotina das pessoas (Atzori et al. 2010). Com isso, novas oportunidades tecnológicas surgiram, possibilitando a criação de modelos de dados que visam inferir, obter conhecimento e auxiliar em tomadas de decisão em muitas áreas (Verbraeken et al. 2020). Dentre essas, as aplicações para cidades inteligentes se apresentam como promissoras, oferecendo oportunidades tais como aquelas ligadas à mobilidade ciente de contexto (Zhang et al. 2011) que, de modo geral, se refere ao uso de dados para oferecer melhorias e comodidades no deslocamento, tanto para motoristas e passageiros quanto para o fluxo de veículos.

Em um futuro próximo, os próprios dispositivos computacionais ligados às aplicações farão, de forma autônoma, a coleta e uso de dados contextuais, aplicando-os na personalização e tomada de decisões (Atzori et al. 2016). Em particular, Sistemas de Transportes Inteligentes (STI) utilizarão um ou mais modelos de dados para traçar rotas a serem utilizadas pelos mais diferentes tipos de veículos (Ladeira et al. 2019). Vários trabalhos na literatura associam dados diversos ao trânsito (e.g., riscos de acidentes nas vias e condições climáticas). Dentre esses, a criminalidade tem se destacado, com propostas que contribuem para identificar vias e regiões com alto índice de registros criminais, possibilitando recomendações de rotas que evitam áreas inseguras de uma cidade.

Embora os trabalhos sobre rotas seguras encontrados até o momento tenham avaliado diferentes soluções, as mesmas foram testadas majoritariamente para veículos particulares, com foco na criação de soluções computacionalmente eficientes que evitem ao máximo regiões perigosas. Entretanto, no recorte dos transportes coletivos, estas soluções poderiam abrir mão de questões como o impacto social causado, neste caso, deixando pessoas com acesso precarizado a este meio de deslocamento. Este trabalho, ao contrário dos demais, avalia a elaboração de rotas seguras para transportes coletivos públicos.

Diante do que foi apresentado, a seguinte pergunta pode ser elaborada: é possível que os ônibus de uma cidade utilizem soluções de rotas seguras equilibrando os ganhos de nível de segurança dos passageiros com o aumento da distância percorrida pelas rotas? Além disso, com uma possível implementação de cálculo de rotas seguras que evitam áreas consideradas perigosas, o sistema ainda seria capaz de manter-se comprometido com a área de cobertura de transporte? A hipótese inicial é de que seja possível, desde que sejam observados limites acerca do que possa ser modificado em termos de pontos de paradas e as trajetórias entre eles.

Este trabalho propõe RACIONAL, uma solução para a construção de rotas seguras para coletivos urbanos baseada em conceitos consolidados utilizados em trabalhos da literatura com foco em veículos particulares, como (Ladeira et al. 2019). Além de encontrar áreas criminais e elaborar para os ônibus rotas que as evitam, RACIONAL propõe o uso de dias da semana e fim de semana como janelas temporais e um algoritmo que permite realocar pontos de paradas para locais mais seguros. A solução desenvolvida foi avaliada considerando-se o sistema de transporte coletivo da cidade de São Paulo, Brasil. Os resultados obtidos mostram que, com poucos ajustes espaço-temporais, é possível encontrar rotas que mantêm a cobertura do serviço de transporte público, ao mesmo tempo em que melhora a segurança das mesmas, evitando áreas consideradas perigosas, sem impor aumento significativo da distância percorrida.

O restante deste trabalho está dividido da seguinte forma: a Seção 2 descreve os principais trabalhos relacionados encontrados na literatura. A solução de determinação de rotas seguras proposta neste trabalho pode ser encontrada na Seção 3. Os resultados que mostram a relação entre segurança e distância percorrida pelos ônibus estão na Seção 4. Por fim, a Seção 5 contém comentários finais e alguns direcionamentos futuros.

2. Trabalhos Relacionados

CERVA (Ladeira et al. 2020) é uma solução para oferecer alternativas de rotas baseada em clusterização KDE e grafos ponderados. A proposta apresentada é baseada na intensidade do tráfego da via juntamente com sua condição de segurança para diferentes janelas temporais do dia. O trabalho conta com bases de dados referentes à cidade de Chicago, incluindo informações criminais oficiais e condições de tráfego provenientes de APIs de sistemas reais, como o *Here Maps*. VTq (de Souza and Villas 2020) é um sistema completo para planejamento de rotas com foco em prever a dinâmica futura de fatores urbanos, tais como mobilidade e criminalidade. Os autores utilizaram na solução uma rede neural e um algoritmo de aprendizado por reforço.

Em (Ladeira et al. 2019), o algoritmo DBSCAN é utilizado para a clusterização de crimes obtidos para a cidade de Chicago. Considerando as rotas mais frequentes seguidas pelos motoristas desta localidade, os autores descrevem uma solução, chamada função *Parzen Window*, capaz de construir rotas seguras escolhendo, entre a origem e o destino, as vias mais afastadas dos *clusters* criminais encontrados. Tais rotas são ainda modificadas em janelas temporais que consideram intervalos de horários ao longo do dia. A proposta de (Ladeira et al. 2019) é utilizada como base para o trabalho proposto neste artigo. Apesar de se tratar de uma solução com foco em veículos particulares e utilizar janelas temporais com intervalos curtos, a estratégia apresentada pelos autores opta por um formato de clusterização que preserva a distribuição original dos crimes, sendo ainda capaz de identificar e ponderar o nível de segurança de cada parte de uma rota por meio de uma única equação compacta, porém efetiva. Além disso, o trabalho escolhido sintetiza de forma satisfatória as escolhas mais recorrentes e melhor avaliadas pelos demais artigos relacionados encontrados na literatura.

SafeRNet (Liu et al. 2017) é um *framework* baseado no uso de redes bayesianas que visa diminuir ocorrências de acidentes para veículos particulares. No estudo foram utilizados critérios como condições climáticas, fluxo de veículos e colisões e condições das vias para propor variações de rotas ao decorrer do dia. Existem muitos estudos que apresentam propostas para confecção de rotas para veículos particulares, considerando especialmente o índice de segurança das mesmas. (Félix Mata et al. 2016) apresentam um *framework*, construído através de redes bayesianas, que elabora rotas seguras com base em um sistema de recomendação híbrido que utiliza dados oficiais e do *Twitter*. O trabalho obteve cerca de 75% de precisão nas recomendações utilizando apenas os dados da rede social. Além disso, também apresenta flexibilidade quanto a janelas temporais. SafePaths (Galbrun et al. 2016) é um sistema de recomendação de rotas seguras que utiliza clusterização *Gaussian Kernel Density Estimation (KDE)*. As rotas são geradas a partir de pequenas regiões como ruas e bairros, tendo sido utilizadas como base as cidades de Chicago e Filadélfia. Os autores argumentam que os resultados mostraram que a solução é robusta em relação a tempo de execução e taxa de atualização dos dados criminais.

Por fim, (Tompson et al. 2009) apresentam *Hot Routes*, uma proposta voltada para ônibus. A solução, baseada em clusterização KDE, tem como principal objetivo identificar regiões inseguras, referentes à criminalidade, como pontos de paradas ou trechos do percurso, na cidade de Londres e exibi-las em um mapa. Esta proposta foi a única encontrada que foca em linhas de ônibus. Entretanto, diferente do presente trabalho, não propõe novas sugestões de rotas para substituir as identificadas como inseguras. Além disso, não possui nenhum método para alterar os pontos de parada para tornar a rota mais segura.

A Tabela 1 apresenta uma comparação do estado da arte com a proposta neste artigo.

	Características			
	Projetado para	Algoritmo	Janela Temporal	Contexto
RACIONAL	Ônibus	DBSCAN e Parzen Window	Sim	Criminal
CERVA (Ladeira et al. 2020)	Veículos Particulares	KDE	Sim	Criminal, Acidentes e Mobilidade
VTq (de Souza and Villas 2020)	Veículos Particulares	LSTM	Sim	Criminal e Mobilidade
Solução base (Ladeira et al. 2019)	Veículos Particulares	DBSCAN e Parzen Window	Sim	Criminal
SafeRNet (Liu et al. 2017)	Veículos Particulares	Redes Bayesianas	Sim	Trânsito (Acidentes)
(Félix Mata et al. 2016)	Veículos Particulares	Redes Bayesianas	Sim	Criminal
SafePaths (Galbrun et al. 2016)	Veículos Particulares	KDE	Não	Criminal
Hot Routes (Tompson et al. 2009)	Ônibus	KDE	Não	Criminal

Tabela 1. Características do trabalho proposto em relação ao estado da arte.

3. RACIONAL: Rotas Alternativas para veículos Coletivos com segurança e distância equilibradas

Com o intuito de alcançar o objetivo deste trabalho, uma proposta para a construção de rotas seguras para linhas de coletivos urbanos chamada RACIONAL foi elaborada e será apresentada a seguir. A proposta utiliza, como será mostrado, fundamentos extraídos do trabalho de (Ladeira et al. 2019). Entretanto, esta solução base de referência da literatura foi originalmente direcionada para veículos particulares. Dessa maneira, os elementos da mesma, utilizados neste trabalho, precisaram ser adaptados ao recorte dos coletivos urbanos. Além disso, a solução aqui proposta ainda contempla uma estratégia original para realocar pontos de parada, de maneira a aumentar a segurança das rotas sem comprometer a cobertura do serviço de transporte ou aumentar significativamente as distâncias percorridas por veículos, em suas rotas, e passageiros até os locais de embarque.

No trabalho atual, a cidade de São Paulo foi escolhida por atender a três requisitos essenciais: disponibilidade no grafo da biblioteca utilizada para traçar rotas, a qual será detalhada a seguir; obtenção de uma base de dados criminal atualizada, com geolocalização, data e breve descrição do crime¹; disponibilização de detalhes de rotas de transporte público, como informações das linhas e seus respectivos pontos de paradas². Porém, a solução é aplicável a qualquer município que tenha esses dados disponíveis. A ferramenta “OSMnx: Python for street networks”³, disponível como um pacote da linguagem Python, foi utilizada na construção do grafo que representa a cidade de São Paulo. Nele, cada esquina da cidade é representada como um vértice, sendo as ligações entre elas modeladas como arestas direcionais representando as vias.

Definições: Uma cidade possui N linhas de ônibus distintas. Cada linha de ônibus é definida como $L_i, 1 \leq i \leq N$. Uma linha (L_i) pode percorrer uma ou mais rotas entre sua origem e seu destino. R_i é o conjunto de todas as rotas R_i^x existentes para a linha L_i sendo, em particular, R_i^c a rota mais curta (i.e., aquela que percorre o caminho com a menor distância) e R_i^s a rota mais segura (i.e., aquela que percorre o caminho com o menor índice de criminalidade). Cada rota R_i^x é composta por uma sequência de m vértices v_j , os quais neste trabalho representam esquinas por onde o veículo deve passar. Assim, $R_i^x = \{v_0, v_1, \dots, v_{(m-1)}\}$, em que m é o número de vértices que compõem a rota. Seja P_i o conjunto de pontos de parada p_z da linha L_i , onde p_z é um vértice no qual o coletivo permite embarque e desembarque de passageiros. Logo, $P_i \subset R_i^x \forall R_i^x \in R_i$. Em outras palavras, todos os pontos de parada de uma linha devem estar presentes em todas as possíveis rotas dessa linha. Por fim, um trecho, $R_{t(p_i, p_j)}$, é um subconjunto de vértices de uma rota que ligam o par de pontos de parada p_i e p_j .

(Ladeira et al. 2019) propuseram um sistema de recomendação de rotas seguras baseado em dados históricos criminais. O trabalho propõe identificar áreas de alta criminalidade pelo agrupamento de dados criminais geolocalizados e utilizar um algoritmo que sugere rotas, em janelas temporais, que evitam passar pelas mesmas. Na solução RACIONAL, considerando-se as particularidades para os ônibus urbanos, a mesma estratégia para identificação de áreas criminais foi utilizada. Em seguida, na etapa de construção das rotas seguras, foi preciso considerar todos os pontos de parada (P_i) dentro das rotas R_i^x de cada linha L_i , ao invés de olhar apenas para os locais de origem e destino dos veículos. Além disso, a semântica das janelas temporais da solução base também foi modificada, visto não ser viável o uso, como propõe o trabalho, de trajetos diferentes para intervalos de horas em um mesmo dia. Dessa maneira, conforme será explicado posteriormente, foram propostas mudanças das rotas apenas entre dias de semana e fim de semana.

3.1. Delimitação de Áreas Criminais

O algoritmo *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN) foi utilizado para a identificação de áreas criminais geograficamente distribuídas pela cidade. Esta escolha ocorreu devido à habilidade do algoritmo de gerar agrupamentos (clusters) com diferentes formatos e tamanhos e por ter bom resultado de clusterização mesmo na presença de ruídos (Santos et al. 2017).

¹Dataset criminal: <https://data.world/maszanchi/boletins-de-ocorrencia-sp-2014>

²Dataset das rotas de ônibus: https://interscity.org/open_data/

³<https://osmnx.readthedocs.io/en/stable/>

O *dataset* criminal utilizado é um arquivo no formato CSV (*Comma-Separated Values*) obtido através da plataforma *data.world*¹. Ele contém aproximadamente 945 mil registros criminais da cidade de São Paulo no ano de 2014. A escolha deste ano se tornou limitada à disponibilidade do atributo *data_ocorrencia_bo* (AAAA/MM/DD), que não estava presente nas últimas versões deste *dataset*, encontradas nos anos de 2015 e 2016, na mesma plataforma. Este arquivo engloba atributos como: número do boletim de ocorrência, data do crime, localização descrita e através de latitude e longitude, rubrica do crime, dentre outras. Com base nessas informações, foi feito um primeiro ajuste: uma vez que dados sem geolocalização não são úteis para o trabalho, os mesmos foram identificados e descartados. Com isso, após essa remoção, a base de dados criminais foi reduzida para aproximadamente 732 mil crimes. Outro ponto importante está relacionado às características dos crimes com o contexto do problema. Por se tratar de um estudo relacionado ao trânsito, alguns tipos criminais foram considerados indiferentes para a segurança de rotas veiculares, como por exemplo: “Oferecer droga a pessoa de seu relacionamento (Art.33,§3º)”. Logo, estes tipos de crimes foram desconsiderados, tendo sido mantidos 4 tipos criminais dos 37 existentes na base de dados, os quais podem ser observados na Tabela 2. Com isso, após esta etapa de seleção de crimes coerentes para o contexto analisado, basta utilizá-los em conjunto com o algoritmo DBSCAN, gerando e demarcando os *clusters* criminais pela cidade.

Tipos de crimes	Quantidade	Porcentagem (%)
Total de crimes válidos	731.899	100,00
Total de crimes selecionados	522.854	71,44
Detalhes dos tipos de crimes selecionados	522.854	100,00
Furto (art. 155)	142.175	27,19
Homicídio simples (art. 121)	9.971	1,90
Roubo (art. 157)	370.532	70,87
Tráfico de entorpecente (Art. 12)	176	0,04

Tabela 2. Distribuição dos tipos de crimes selecionados e utilizados

O DBSCAN possui dois parâmetros: ε (raio entre vizinhos) e v (quantidade mínima vizinhos). Como a base de dados criminal possui crimes registrados no decorrer de um ano, foi estabelecido um número mínimo de ocorrências mensais, coerente com o significado de uma área ser segura ou insegura, para então aplicar para todo o ano. Assim, foi escolhido um valor que representa uma média de pelo menos três crimes a cada mês, o que resulta em pelo menos 36 crimes anuais em cada região avaliada, i.e., $v = 36$. Desta forma, todo agrupamento contendo pelo menos 36 ocorrências criminais foi considerado uma região insegura. Além disso, o parâmetro relacionado à distância máxima entre dois nós de um mesmo *cluster* foi definido, de maneira exploratória, como $\varepsilon = 100m$.

3.2. Identificação de Rotas Seguras

Para a identificação das rotas seguras, foi utilizada a equação da solução base (Equação 1) que leva em consideração o resultado do agrupamento dos dados criminais. Todos os trechos da rota $R_{t(p_i, p_j)}$ recebem uma pontuação $k_{(p_i, p_j)}$, que quando somadas, representam

¹<https://data.world/>

o valor K , responsável por indicar o nível de periculosidade da rota. Desta forma, quanto menor for este valor, mais segura será a rota. Essa equação, denominada *Parzen Window*, é uma forma não paramétrica para estimar a função densidade de probabilidade de uma variável aleatória (Babu and Viswanath 2008).

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

Na Equação 1, m indica a quantidade de vértices (representados por esquinas) do trecho da rota entre p_i e p_j ; σ é o desvio padrão da quantidade de crimes por *cluster*; x_a representa a distância de cada vértice em relação ao centro do *cluster* mais próximo; e x representa a distância da borda do *cluster* mais próximo em relação ao seu centro.

Os dados utilizados nesta etapa sobre as rotas de ônibus da cidade de São Paulo - SP foram retirados da plataforma *Interscity*²(Wen et al. 2018). Dentre as informações contidas, estão presentes as linhas de ônibus para cada dia da semana e também para feriados, além dos pontos de parada distribuídos pela cidade com suas respectivas geolocalizações. Essas informações estão distribuídas através de dois modelos de arquivos XML (eXtensible Markup Language). Para o primeiro tipo, os arquivos são separados por dia da semana. Dentro destes, existem elementos com atributos como o identificador das linhas de ônibus, que totalizam 2.183 linhas, e suas respectivas listas de pontos de parada. Já o segundo modelo é um complemento do primeiro, pois nele estão armazenados os pontos de parada, totalizando 20.041, com suas respectivas coordenadas geográficas. Em média, as linhas de ônibus da cidade de São Paulo possuem 43 pontos de parada, sendo 5 o número mínimo, 132 o máximo e 40 a mediana.

Para a avaliação da solução RACIONAL, foi necessário traçar dois tipos de rotas para cada linha L_i : a rota mais segura R_i^s e a rota mais curta R_i^c . A estratégia seguida utiliza a própria ferramenta OSMnx para gerar todas as opções de rotas R_i para cada linha de ônibus L_i . A rota mais segura é construída analisando-se apenas o $k_{(p_i,p_j)}$ obtido para cada trecho. Desta forma, é escolhida, a cada trecho, a opção que apresenta o caminho com menor $k_{(p_i,p_j)}$. A rota mais curta é utilizada como um parâmetro de comparação para as demais rotas geradas, pois, através dela é possível criar comparações de distâncias percorridas. Sendo assim, para construí-la, o único fator analisado é a menor distância percorrida entre um ponto de parada e outro, desconsiderando o fator criminal. Contudo, como esta rota é gerada da mesma forma que as demais, com a ressalva de escolha da menor distância, é possível identificar seu nível de segurança pelo fator K , da mesma forma como descrito anteriormente para qualquer outra rota. Logo, também torna-se possível classificá-la em termos de segurança para fins de comparação. O Algoritmo 1 formaliza como as rotas R_i^c e R_i^s são calculadas.

3.3. Realocação de Pontos de Parada

Em se tratando de transporte coletivo, os pontos de parada podem ser os locais onde os cidadãos estão mais vulneráveis e suscetíveis a serem vítimas de crimes. Independentemente das possibilidades de rotas criadas, em algumas situações, determinados pontos

²<https://interscity.org/>

Algoritmo 1 Identificação de R_i^c e R_i^s para a linha L_i

Entrada: P_i

```
1:  $R_i^c \leftarrow \emptyset, R_i^s \leftarrow \emptyset$ 
2: for  $index \leftarrow 0$  to  $(len(P_i) - 2)$  do
3:    $origem \leftarrow P_i[index]$ 
4:    $destino \leftarrow P_i[index + 1]$ 
5:    $R_{t(origem, destino)} \leftarrow gera\_rotas(origem, destino)$ 
6:    $R_{t(origem, destino)}^c \leftarrow rota\_mais\_curta(origem, destino)$ 
7:    $R_i^c \leftarrow R_i^c + R_{t(origem, destino)}^c$ 
8:    $niveis \langle K, R^x \rangle \leftarrow \{ \}$  ▷ Tupla de níveis de crimes e rotas
9:   for  $R_{t(origem, destino)}^x \in R_{t(origem, destino)}$  do
10:      $niveis \leftarrow niveis + \{ \langle parzen\_win(R_{t(origem, destino)}^x), R_{t(origem, destino)}^x \rangle \}$ 
11:   end for
12:    $sort(niveis)$ 
13:    $R_i^s \leftarrow R_i^s + niveis[0]$ 
14: end for
Saída:  $R_i^c, R_i^s$ 
```

de paradas seriam os locais menos seguros, pois estariam localizados dentro de um *cluster* criminal. Nesta seção é apresentada uma contribuição original de RACIONAL, que visa alterar os pontos de paradas das linhas, quando possível. Na solução apresentada anteriormente, as opções de rotas criadas são geradas sempre com base nos pontos de parada originais, alterando apenas os caminhos entre paradas consecutivas. Porém, Assim, a estratégia utilizada para mitigar este problema é tentar remover esse tipo de parada, utilizando outra, próxima, para substituí-la.

Para realizar as realocações de paradas, foram adotados alguns critérios. Primeiramente, somente pontos de paradas já alocados para alguma linha de ônibus qualquer poderão ser utilizados, evitando a necessidade de investimento em infraestrutura para se criar uma nova parada. Para que um ponto de parada localizado dentro de um *cluster* seja removido, deve haver algum outro ponto de parada fora de uma região criminal, distante do primeiro a, no máximo, *Raio* metros, em linha reta (ou seja, distância euclidiana). No trabalho atual, $Raio = 250m$. Além disso, por se tratar de transporte terrestre, também foi considerada a distância real de locomoção, isto é, a distância percorrida de uma origem para um destino passando pelas vias existentes entre eles na cidade. Desta forma, esta distância foi definida como $Distancia = 1000$ metros, do ponto original até o ponto candidato, visto que o deslocamento real possui uma distância maior do que aquela em linha reta. Com a aplicação desta estratégia, é possível garantir que os cidadãos não terão que percorrer distâncias maiores que as estabelecidas para conseguir o acesso ao transporte, ou, no pior caso, ficar sem acesso ao novo ponto. Se algum dos critérios não for satisfeito ou se o ponto de parada não estiver localizado em nenhuma região criminal, a realocação não é realizada. O Algoritmo 2 descreve esses passos.

3.4. Janelas Temporais

Por tratar de veículos particulares, em (Ladeira et al. 2019) as alterações em curtos intervalos de tempo não causam impactos significativos aos motoristas. Desta forma, os

Algoritmo 2 Realocação de Pontos de Parada de uma Linha L_i

Entrada: P_i

```
1:  $num\_clusters \leftarrow 5, raio \leftarrow 250, distancia \leftarrow 1000$ 
2: for  $p_z$  in  $P_i$  do
3:   if  $esta\_dentro\_cluster(p_z)$  then
4:      $clusters\_proximos \leftarrow encontra\_clusters(p_z, num\_clusters)$ 
5:      $p_{candidato} \leftarrow vertice\_valido\_fora\_cluster(clusters\_proximos, raio, distancia)$ 
6:     if  $p_{candidato}$  then
7:        $p_z \leftarrow p_{candidato}$ 
8:     end if
9:   end if
10: end for
```

autores avaliaram opções de intervalos para janelas temporais, criando diferentes *clusters* para intervalos de tempo variando entre 1, 2, 4, 8 e 12 horas. Ao final, chegaram à conclusão de que o intervalo de 2 horas seria o melhor para identificar as mudanças na criminalidade das regiões. Essa estratégia não é viável ao recorte de ônibus urbanos, justamente por tornar as alterações nas rotas muito constantes. Na prática, isso resultaria em diferentes versões de uma única linha para diferentes horários do dia, realocando pontos de parada e fazendo com que os passageiros tivessem que conhecer e utilizar um número grande e diversificado de paradas.

Sendo assim, para diminuir os impactos e estresse que poderia ser causado aos passageiros por tantas modificações, neste trabalho foram utilizadas três janelas temporais: **semana**, que utiliza apenas crimes cujas datas estejam dentro do intervalo de segunda a sexta; **fim de semana**, complemento da semana, cria *clusters* apenas com dados de sábado e domingo; e **geral**, que não possui nenhuma separação temporal e é utilizada como parâmetro de comparação com as demais. O próprio padrão de mobilidade da população reforça a escolha por essas janelas temporais, uma vez que para boa parte das pessoas, trabalho e lazer também são divididos nestes intervalos. Vale ressaltar que os dados de feriados disponibilizados no *dataset* não foram utilizados. A Tabela 3 apresenta estatísticas sobre os *clusters* criminais utilizados neste trabalho divididos nas janelas temporais propostas. É possível observar que há uma mudança significativa na distribuição dos crimes, validando a escolha das janelas temporais.

Clusters	Janelas Temporais		
	Geral	Semana	Final de semana
Quantidade	2.030	1.363	214
Média da área (m ²)	54.397,47	45.963,15	32.253,68
Média crimes	103,48	90,24	85,26
Desvio padrão crimes	164,90	108,79	191,28

Tabela 3. Estatísticas dos *clusters* criados por janela.

Diante disso, ao se criar as rotas seguras para os ônibus, será observado se os resultados obtidos possuem interferências dessas janelas temporais.

4. Resultados e Análises

Para a implementação deste trabalho, foi utilizado um *cluster* computacional da Universidade Federal de Viçosa, composto por 6 núcleos e 20GB de memória RAM. Do total de 2.183 linhas de ônibus disponíveis, 2.089 puderam ser adaptadas ao modelo de grafos disponibilizado pela biblioteca OSMnx, e conseqüentemente, utilizadas no decorrer da identificação de rotas e análises. As 94 linhas removidas não possuíam trajetos entre alguns de seus pares de pontos de parada. Com o objetivo de avaliar as principais características da solução de roteamento seguro, foram definidos os cenários apresentados na Tabela 4. Logo, para cada cenário, foram examinadas todas as opções de rotas disponíveis. Os cenários 4 e 5 correspondem à implementação completa da solução RACIONAL, utilizando, respectivamente, as janelas de tempo semana e fim de semana.

Cenário	Tipo de Deslocamento	Janela Temporal	Tem Realocação?
1	Mais curto	Geral	Não
2	Mais seguro	Geral	Não
3	Mais seguro	Geral	Sim
4	Mais seguro	Semana	Sim
5	Mais seguro	Fim de semana	Sim

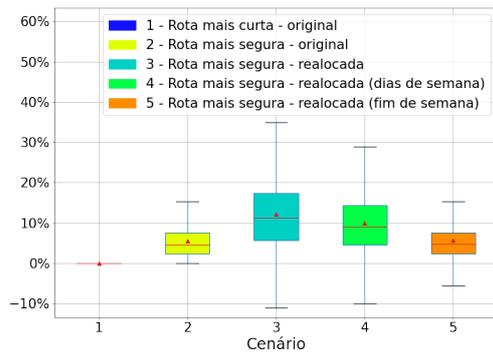
Tabela 4. Cenários implementados no trabalho.

Antes de apresentar as métricas, vale ressaltar que os resultados obtidos foram adquiridos com base em análises relativas ao cenário 1. Como brevemente apresentado, o cenário 1 foi escolhido como referência para a forma atual de como as linhas de ônibus operam dentro da cidade. Esta decisão foi tomada uma vez que o *dataset* de linhas de ônibus possui apenas os pontos de parada. Foi necessário traçar rotas entre eles para gerar soluções iniciais, sendo a do cenário 1 aquela mais coerente para ocupar esta posição, justamente por apresentar o caminho mais curto entre pares de pontos de parada.

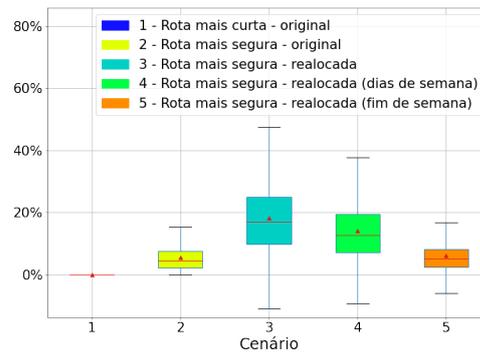
4.1. Comprimento Total e Distância entre Paradas

Comprimento Total Percorrido indica a distância percorrida em uma rota, calculada em relação ao cenário 1. Em outras palavras, mostra a quantidade total em metros percorridos a mais para cada cenário, entre o ponto de origem e o ponto de destino, em relação ao cenário 1. A pergunta a ser respondida pela análise dessa métrica é: criando rotas alternativas mais seguras, é possível manter um comprimento próximo ao original?

Na Figura 1a percebe-se que houve um aumento médio de aproximadamente 5% se comparado o cenário 1 com o 2. Mais precisamente, o comprimento médio para o cenário 1 é 20.076,69m e para o cenário 2 é 21.128,49m. No gráfico também se encontram os cenários 3 a 5, os quais possuem as paradas realocadas, sendo que os cenários 4 e 5 utilizam janelas temporais diferentes. O que pode ser observado é que há um aumento no comprimento da rota em todos os cenários, visando uma maior segurança. No entanto, esse aumento é em sua maioria inferior a 10% em relação à rota original. Outra observação é que a realocação de pontos de parada leva a um maior aumento (cenários 3 a 5), mas as janelas temporais foram importantes para mitigar esse aumento, ao se usar dados criminais de semana (cenário 4) e fim de semana (cenário 5).



(a) Média do comprimento total



(b) Média do comprimento entre pontos de parada

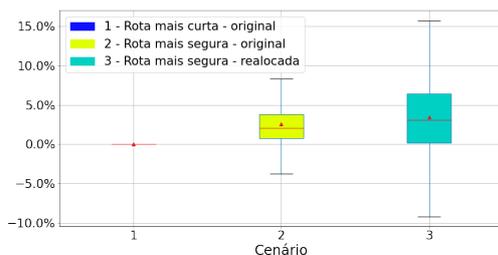
Figura 1. Média - Comprimentos total e entre pontos de paradas.

Em sintonia com o que foi apresentado, *Comprimento Médio entre Pontos de Parada* é uma métrica que representa a distância média que o ônibus terá que percorrer entre dois pontos de parada adjacentes. A Figura 1b mostra que o comportamento é o mesmo da métrica anterior, já que para termos uma rota mais segura, é possível que o ônibus tenha que se deslocar por um caminho diferente do mais curto. Além disso, a realocação de pontos de parada leva a um aumento nessa distância, sendo que as janelas temporais também foram importantes para mitigar o aumento.

4.2. Distâncias até os centros dos clusters

A métrica *Média das Distâncias dos Vértices em Relação aos seus Respectivos Centros de Clusters mais Próximos* é utilizada para identificar, em média, o quão próximo ao centro de uma região perigosa o ônibus está passando. Ou seja, quanto mais longe desse centro, melhor. Na Figura 2a percebe-se que tanto o cenário 2 quanto o cenário 3 possuem em média uma melhoria de aproximadamente 2,5% se comparado com o cenário 1.

Através do gráfico na Figura 2b é possível notar que os cenários 4 e 5 possuem um conjunto de valores médios mais distintos, se comparados ao cenário 1. Parte disto se deve ao fato de serem rotas construídas sob *clusters* diferentes (Tabela 3) em função das janelas temporais. Esses resultados mostram que as janelas temporais são fundamentais não apenas para reduzir o comprimento das rotas seguras, mas também para torná-las mais seguras, passando mais longe dos centros dos *clusters*.



(a) Cenários 2 e 3



(b) Cenários 4 e 5

Figura 2. Média das distâncias dos nós em relação aos centros de clusters.

4.3. Distância dentro de clusters

Distância percorrida dentro de clusters criminais é uma métrica relacionada a quanto um ônibus trafega dentro de *clusters* criminais durante o trajeto. Ou seja, quanto menor, melhor. Para isso foram feitas algumas adaptações. Como não é possível identificar as intercessões entre vias e *clusters* na ferramenta *OSMNx*, foi considerado que para um trecho estar dentro de uma região perigosa ele deve começar e terminar em esquinas localizadas dentro de um *cluster*. Com isso, diante da rota completa, cada par de esquinas que satisfizer estas condições será contabilizado para esta métrica.

A Figura 3 também compara os demais cenários com o cenário 1. De acordo com essa figura, percebe-se que os cenários mais seguros percorrem uma distância menor dentro de regiões perigosas, por isso a porcentagem para estes cenários é negativa. Mais uma vez, nota-se os benefícios da proposta atual de realocar paradas quando possível, já que os cenários 3 a 5 possuem os melhores resultados. Além disso, as janelas temporais também tiveram um efeito positivo (cenários 4 e 5).

4.4. Quantidade de Paradas em Clusters Criminais

Por fim, *Quantidade de Paradas em Clusters Criminais* é uma métrica que faz um comparativo entre as quantidades de pontos de paradas originais e os novos propostos localizados nas áreas inseguras demarcadas. Antes de apresentar essa comparação, é preciso ressaltar que diante dos cinco cenários apresentados, os resultados são os mesmos para os cenários 1 e 2, justamente por possuírem as configurações originais da rota.

Na Tabela 5 é possível notar que com a implementação da possibilidade de realocação de pontos de parada, vários pontos originais foram retirados de regiões inseguras e realocados para regiões fora de *clusters* criminais. Além disso, a título de comparação, a média de pontos de paradas, considerando realocação, em regiões com *clusters* criminais é apresentada na última coluna da Tabela 5. Através dessa, é possível observar que se comparado à solução original, em média, aproximadamente 25% dos pontos de parada estão localizados dentro de *clusters*. Com as propostas sugeridas, há um ganho significativo para os cenários 3, 4 e 5, reduzindo este valor para 8,12%, 4,17% e 0,61%, respectivamente. Mais uma vez, destaca-se a importância tanto da realocação proposta, quanto do uso de janelas temporais apropriadas para as rotas seguras. Estes resultados são ainda mais relevantes, considerando que os pontos de paradas são os locais de maior vulnerabilidade dos cidadãos em termos de criminalidade, justamente pelo fato do ônibus estar parado para embarque e desembarque de passageiros.

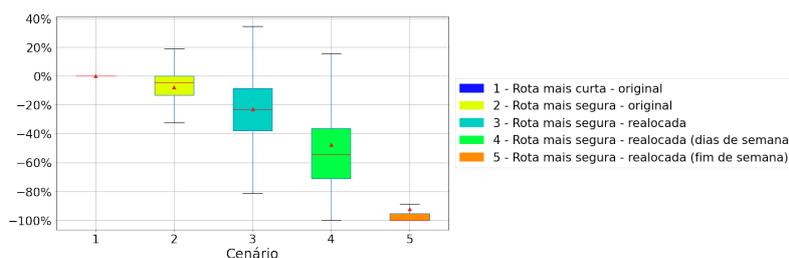


Figura 3. Soma das distâncias sequenciais percorridas dentro de clusters.

Cenários	Mediana	Máximo	Média	Desvio padrão	% de pontos de parada em <i>clusters</i> da rota
1 e 2	9	45	10,35	6,76	25,46%
3	2	17	2,94	2,89	8,12%
4	1	10	1,56	1,84	4,17%
5	0	8	0,25	0,80	0,61%

Tabela 5. Estatísticas de pontos de parada localizados em *clusters* criminais.

5. Conclusão e Trabalhos Futuros

Este artigo descreveu a solução RACIONAL, uma proposta para elaboração de rotas alternativas seguras para ônibus urbanos. Para isso, foi adaptada uma solução da literatura para o contexto de transportes coletivos urbanos. Além disso, foi feita uma proposta original para alterar os pontos de parada, e para utilizar janelas temporais separadas em dias de semana e fim de semana.

É possível notar, por meio das métricas avaliadas, como os cenários 2 a 5 obtiveram ganhos quando comparados ao modelo de rotas original (Cenário 1). Percebe-se que a proposta de realocação de pontos de paradas teve um impacto positivo significativo, diminuindo em até 25 pontos percentuais a quantidade de paradas localizadas dentro de *clusters*. Além disso, distanciou cada vértice da rota do centro do *cluster* mais próximo e resultou em trajetos até 40% menores em áreas criminais para rotas em dias de semana.

Implantar o conceito de janelas temporais, adaptadas ao contexto de ônibus e mobilidade urbana, também contribuiu significativamente para os resultados. Sem a separação em janelas, algumas rotas ou trechos poderiam ser modificadas sem necessidade. Todos os ganhos ressaltados puderam ser obtidos a um custo de aumentar a média do comprimentos totais das rotas em aproximadamente 12%, no seu pior caso. Isso mostra que as alterações não comprometeram significativamente o trajeto existente, ficando dentro dos limites estabelecidos e garantindo o acesso dos passageiros para os novos pontos de parada recomendados.

Como trabalhos futuros, é possível avaliar como rotas podem se tornar mais seguras se os pontos de parada puderem ser alocados em qualquer localização, e não apenas em pontos de parada previamente existentes. Além disso, também é importante avaliar o impacto social em termos do número de cidadãos afetados pelas mudanças de rotas. Para este último, uma alternativa seria construir análises baseadas em padrões de mobilidade e horários de pico da região avaliada, de forma a balancear a disponibilidade do transporte com as novas distâncias percorridas pela população.

Referências

- [Atzori et al. 2010] Atzori, L., Iera, A., and Morabito, G. (2010). The internet of things: A survey. *Computer Networks*, 54(15):2787–2805.
- [Atzori et al. 2016] Atzori, L., Iera, A., and Morabito, G. (2016). Understanding the internet of things: definition, potentials, and societal role of a fast evolving paradigm. *Ad Hoc Networks*, 56.

- [Babu and Viswanath 2008] Babu, V. S. and Viswanath, P. (2008). An efficient and fast parzen-window density based clustering method for large data sets. In *International Conference on Emerging Trends in Engineering and Technology*, pages 531–536.
- [de Souza and Villas 2020] 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.
- [Félix Mata et al. 2016] 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 et al. 2016] Galbrun, E., Pelechrinis, K., and Terzi, E. (2016). Urban navigation beyond shortest route: The case of safe paths. *Information Systems*, 57:160–171.
- [Ladeira et al. 2020] 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 et al. 2019] 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 et al. 2017] 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)*.
- [Santos et al. 2017] 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 sensoriados por usuários sobre crimes. In *Anais do XXII Workshop de Gerência e Operação de Redes e Serviços*, Porto Alegre, RS, Brasil. SBC.
- [Tompson et al. 2009] 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.
- [Verbraeken et al. 2020] Verbraeken, J., Wolting, M., Katzy, J., Kloppenburg, J., Verbeelen, T., and Rellermeyer, J. S. (2020). A survey on distributed machine learning. *ACM Comput. Surv.*, 53(2).
- [Wen et al. 2018] Wen, M., Rosa, T. d. O., Souza, M. C., Aleixo, R. P., Alves, C., Sá, L., Santana, E. F. Z., and Kon, F. (2018). Criação de modelo para simulação de movimentação de Ônibus a partir de dados reais. In *Proceedings of the 1st Brazilian Workshop on Smart Cities*, WBCI 2018, pages 1–10.
- [Zhang et al. 2011] Zhang, J., Wang, F.-Y., Wang, K., Lin, W.-H., Xu, X., and Chen, C. (2011). Data-driven intelligent transportation systems: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 12(4):1624–1639.