



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE CIÊNCIA E TECNOLOGIA
CURSO DE ENGENHARIA DE TRANSPORTES

MILENA FERREIRA DA CONCEIÇÃO

**ESTUDO DA VIABILIDADE DA UTILIZAÇÃO DE DADOS
WI-FI PÚBLICOS NA RECONSTRUÇÃO DE MATRIZES
ORIGEM/DESTINO DE PASSAGEIROS**

TRABALHO DE CONCLUSÃO DE CURSO

APARECIDA DE GOIÂNIA
2020

UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE CIÊNCIA E TECNOLOGIA
CURSO DE ENGENHARIA DE TRANSPORTES

MILENA FERREIRA DA CONCEIÇÃO

**ESTUDO DA VIABILIDADE DA UTILIZAÇÃO DE DADOS
WI-FI PÚBLICOS NA RECONSTRUÇÃO DE MATRIZES
ORIGEM/DESTINO DE PASSAGEIROS**

Monografia apresentada na disciplina de Trabalho de Conclusão de Curso da graduação em Engenharia de Transportes da Faculdade de Ciência e Tecnologia da Universidade Federal de Goiás.

Orientador: Prof. Dr. Marcos Paulino Roriz Junior

APARECIDA DE GOIÂNIA
2020

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Da Conceição, Milena Ferreira
Estudo da Viabilidade da Utilização de Dados Wi-Fi Públicos
na Reconstrução de Matrizes Origem/Destino de Passageiros
/ Milena Ferreira da Conceição.
2021, 79 f.: il.

Orientador: Prof. Dr. Marcos Paulino Roriz Junior.
Trabalho de Conclusão de Curso (Graduação) – Universidade Federal de Goiás,
Campus Aparecida de Goiânia, Engenharia de Transportes, Goiânia, 2021.
Bibliografia. Anexos. Apêndice.

Inclui siglas, mapas, fotografias, abreviaturas, símbolos, gráfico, tabelas,
algoritmos, lista de figuras, lista de tabelas.

1. Matriz O/D. 2. Wi-Fi. 3. Estimção. 4. Viabilidade. 5. Comparação da matriz O/D.
Roriz Junior, Marcos Paulino. Estudo da Viabilidade da Utilização de Dados Wi-Fi Públicos
na Reconstrução de Matrizes Origem/Destino de Passageiros

CDU 625

UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE CIÊNCIA E TECNOLOGIA
CURSO DE ENGENHARIA DE TRANSPORTES

MILENA FERREIRA DA CONCEIÇÃO

**ESTUDO DA VIABILIDADE DA UTILIZAÇÃO DE DADOS
WI-FI PÚBLICOS NA RECONSTRUÇÃO DE MATRIZES
ORIGEM/DESTINO DE PASSAGEIROS**

Monografia apresentada na disciplina de Trabalho de Conclusão de Curso da graduação em Engenharia de Transportes da Faculdade de Ciência e Tecnologia da Universidade Federal de Goiás.

Aprovada por:

Prof. Dr. Marcos Paulino Roriz Junior
(Orientador)

Prof. Dr. Ronny Marcelo Aliaga Medrano
(Examinador Interno)

Prof. Dr. Carlos Eduardo Sanches de Andrade
(Examinador Interno)

Data: ____/____/____



TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO NO REPOSITÓRIO INSTITUCIONAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio do Repositório Institucional (RI/UFG), regulamentado pela Resolução CEPEC nº 1204/2014, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do Trabalho de Conclusão de Curso de Graduação (TCCG):

Nome completo do autor: MILENA FERREIRA DA CONCEIÇÃO

Título do trabalho: ESTUDO DA VIABILIDADE DA UTILIZAÇÃO DE DADOS WI-FI PÚBLICOS NA RECONSTRUÇÃO DE MATRIZES ORIGEM/DESTINO DE PASSAGEIROS

2. Informações de acesso ao documento:

Concorda com a liberação total do documento [] SIM [] NÃO¹

Assinatura do(a) autor(a)²

Ciente e de acordo:

Assinatura do(a) orientador(a)²

Data: ___/___/___

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo. Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

² A assinatura deve ser escaneada.

Dedico este trabalho primeiramente a Deus, que me sustentou para continuar a cada dia essa jornada. A minha mãe Nazaré que me apoiou em todas as minhas decisões, e sempre esteve comigo mesmo nos momentos mais difíceis. Ao meu avô João que sempre foi meu exemplo de pessoa e de profissional. A minha irmã Mirela por estar sempre comigo e me apoiar nos dias difíceis. Aos meus amigos da primeira turma de Engenharia de Transportes, eu amo vocês.

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas graças a Deus, não sou o que era antes.” (Martin Luther King)

AGRADECIMENTOS

A Deus por ser meu refúgio e fortaleza em todos os momentos da minha vida, por sua graça ter alcançado minha vida, e por ser meu socorro bem presente nos momentos em que parecia não ter saída.

A minha mãe Nazaré por estar sempre comigo e com meus irmãos, cuidando, protegendo e garantindo que nossos sonhos possam ser realizados, e por essa fonte de amor e carinho, mesmo nos momentos mais difíceis.

Meu avô João e meu pai Ramon, por sempre estarem comigo nessa caminhada me apoiando e dando condições para que os meus sonhos se tornassem realidade.

Meus irmãos Mirela e Daniel por me dar forças para prosseguir, sempre estarem comigo me apoiando me ajudando.

A minha família em especial minha avó Dio e minha tia Socorro pelo suporte mesmo de longe, por sempre se importarem comigo, vocês são minha força para cada dia querer seguir em frente. Aos demais tios, tias e primos por sempre me apoiarem nessa caminhada.

Aos meus irmãos de fé Paulo Júnio, Marcela, Gustavo, aos meus pastores José Cândido e Márcia por me ajudarem espiritualmente nessa jornada, me aconselhando e orientando.

A família que formei aqui na UFG Adail, Ana Paula, Camila, Daniele, Gabriela, Pedro.

Aos meus amigos que me ajudaram na pesquisa de campo Adail, Daniele, João Paulo, Débora, Andressa, Mateus, Rafael, John, Kamilla e Paulo.

A meus professores de Engenharia de Transporte por todo ensinamento ao longo desses anos, em especial meu orientador Dr. Marcos Paulino Roriz Junior pela paciência, compreensão, pelos ensinamentos valiosos e auxílio para o desenvolvimento do trabalho.

MUITO OBRIGADO!

RESUMO

A compreensão dos deslocamentos dos passageiros do transporte público é um recurso importante, pois auxilia no planejamento e na operação dos sistemas de transporte. A partir de tais deslocamentos é possível construir uma matriz de origem e destino (O/D), que quantifica e sintetiza a mobilidade associada aos usuários do sistema. Entretanto, a construção de tal matriz pode ser demasiadamente custosa em termos operacionais e financeiros. Recentemente tem-se buscado alternativas para tal estimação, entre elas destaca-se o uso de dados Wi-Fi, emitido pelos *smartphones* dos passageiros. Entretanto, a especificação da tecnologia Wi-Fi não padroniza a frequência de envio de tais dados, o que pode dificultar a captura dos deslocamentos dos passageiros. Isto dificulta a captura da matriz ponto a ponto, uma vez que pode ocorrer do passageiro subir em um ponto de parada, mas ter seu dado capturado em estações subsequentes. A fim de avaliar este quesito, este trabalho flexibiliza a estrutura da matriz O/D para diferentes granularidades, isto é, formas de agrupar os pontos de parada (*e.g.*, zonas de tráfego e setores censitários). Assim, o trabalho visa investigar como a matriz O/D estimada a partir de dados Wi-Fi se assemelha da obtida manualmente ao considerar diferentes divisões territoriais e os próprios limites da tecnologia. Para isso, foi realizado um estudo de caso na cidade de Aparecida de Goiânia com intuito de explorar as matrizes O/D geradas a partir da tecnologia Wi-Fi. Os resultados encontrados sugerem que divisões maiores, como macrozonas e zonas de tráfego, apresentam uma matriz mais similar com a real. Entretanto, também ressalta-se cautelas neste resultados devido aos limites das divisões territoriais e da própria tecnologia Wi-Fi.

Palavras-chave: Matriz O/D. Wi-Fi. Estimação. Viabilidade. Comparação da matriz O/D.

ABSTRACT

Understanding the flow of public transport passengers is an important resource, as it helps the planning and operation of transport systems. From such trajectories it is possible to build an origin and destination (O/D) matrix, which quantifies and synthesizes the mobility associated with the system's users. However, the construction of such matrix can be too costly in operational and financial terms. Recently, alternatives for such estimation has surged, among them we can highlight the use of Wi-Fi data, which are passively emitted by the passenger's smartphones. However, the Wi-Fi technology specification does not standardize the frequency that such data are sent, which can make it difficult to capture passengers' movement. Precisely, this makes it difficult to capture a point by point matrix, since it may happen that a passenger that boarded in a given stop point can have his data captured only in subsequent stations. In order to understand this issue, this work makes the structure of the O/D matrix flexible for different granularities, that is, instead of using a point by point schema it groups the stop stations in regions (e.g., traffic zones and census sectors). Thus, the work aims to investigate how the O/D matrix estimated from Wi-Fi data is similar to that obtained manually when considering different territorial divisions and the limits of the technology. Based on that, a case study was carried out in the city of Aparecida de Goiânia in order to explore the O/D matrices generated from the Wi-Fi technology. The results found suggests that larger divisions, such as macrozones and traffic zones, present a matrix more similar to the real one. We also highlight some cautioning remarks that emerged with respect to the Wi-Fi technology itself and when using territorial divisions instead of point by point schema.

Keywords: Matrix O/D. Wi-Fi. Estimation. Viability. Comparison of the O/D matrix.

SUMÁRIO

LISTA DE TABELAS.....	XI
LISTA DE FIGURAS	XII
LISTA DE ABREVIATURAS E SÍMBOLOS	XIII
1 INTRODUÇÃO	1
1.1 PROBLEMA	4
1.2 HIPÓTESE.....	5
1.3 OBJETIVO GERAL E ESPECÍFICO	6
1.4 CONTRIBUIÇÕES	6
1.5 DISTRIBUIÇÃO DO TRABALHO.....	6
2 REVISÃO BIBLIOGRÁFICA	7
2.1 DIVISÃO DO ESPAÇO.....	7
2.1.1 DIVISÕES TERRITORIAIS DAS CIDADES	7
2.1.1.1 ZONAS DE TRÁFEGO	9
2.1.1.2 MACROZONAS	9
2.1.1.3 BAIRROS	10
2.1.1.4 SETORES CENSITÁRIOS	11
2.2 DETERMINAÇÃO DA MATRIZ ORIGEM/DESTINO	13
2.2.1 MÉTODOS DIRETOS	14
2.2.2 MÉTODOS INDIRETOS	15
2.2.2.1 WI-FI.....	16
2.3 MÉTODOS DE COMPARAÇÃO DAS MATRIZES	20
2.3.1 R-QUADRADO (R^2).....	20
2.3.2 RAIZ DO ERRO MÉDIO QUADRÁTICO (RMSE) E ERRO MÉDIO QUADRÁTICO (%RMSE).....	20
2.3.3 DISTÂNCIA DE HELLINGER	21
2.3.4 DEFASAGEM ENTRE AS MATRIZES	21
2.4 TRABALHOS RELACIONADOS	22
3 METODOLOGIA.....	25
3.1 ANÁLISE DAS POLÍTICAS DE DIVISÃO TERRITORIAL DAS CIDADES	25
3.2 DIVISÃO E DETALHAMENTO DA LINHA EM ANÁLISE	26
3.3 CAPTURA DOS DADOS WI-FI.....	28
3.4 PERFIL DE TRANSMISSÃO DOS SMARTPHONES	29
3.5 PROJEÇÃO DO DESLOCAMENTO DOS PASSAGEIROS COM BASE NA TECNOLOGIAWI-FI.....	30
3.6 COMPARAÇÃO DAS MATRIZES	31

4	ESTUDO DE CASO	34
4.1	ANÁLISE DAS POLÍTICAS DE DIVISÃO TERRITORIAL.....	34
4.2	DIVISÃO E DETALHAMENTO DA LINHA EM ANÁLISE	35
4.3	CAPTURE DAS CARACTERÍSTICAS DOS DADOS WI-FI	40
4.4	MATRIZ O/D MANUAL.....	43
4.5	RECONSTRUÇÃO DA MATRIZ O/D USANDO DADOS WI-FI.....	40
4.6	COMPARAÇÃO DAS MATRIZES O/D.....	54
5	CONSIDERAÇÕES FINAIS.....	56
	<u>REFERÊNCIAS BIBLIOGRÁFICAS</u>	59

LISTA DE TABELAS

Tabela 2.1. Textos Relacionados	24
Tabela 3.1. Trecho do perfil de transmissão do smartphone C8:F3:19:60:48:77	30
Tabela 3.2. Exemplo da origem da matriz O/D	32
Tabela 3.3. Exemplo do destino da matriz O/D	24
Tabela 4.1. Número de pontos de parada para macrozonas	38
Tabela 4.2. Número de pontos de parada para zonas	39
Tabela 4.3. Número de pontos de parada para setores censitários	40
Tabela 4.4. Smartphones usados no trabalho	41
Tabela 4.5. Dados sobre os smartphones	41
Tabela 4.6. Matriz O/D manual dos smartphones para as macrozonas	43
Tabela 4.7. Matriz O/D manual dos smartphones para as zonas de tráfego	44
Tabela 4.8. Matriz O/D manual dos smartphones para os setores censitários	44
Tabela 4.9. Dados sobre os smartphones	46
Tabela 4.10. Quantidade média de pacotes enviados por cada smartphone	46
Tabela 4.11. Matriz O/D manual dos smartphones para as macrozonas	51
Tabela 4.12. Matriz O/D manual dos smartphones para as zonas de tráfego	52
Tabela 4.13. Matriz O/D manual dos smartphones para as zonas de tráfego	53
Tabela 4.14. Comparação entre as origens das matrizes O/D para diferentes granularidades	54
Tabela 4.15. Comparação entre os destinos das matrizes O/D para diferentes granularidades	55

LISTA DE FIGURAS

<i>Figura 2.1. Exemplo da relação hierárquica da divisão do espaço com o conceito de zona de tráfego</i>	<i>9</i>
<i>Figura 2.2. Exemplo de divisões territoriais do município de Aparecida de Goiânia</i>	<i>12</i>
<i>Figura 2.3. Elementos de uma matriz O/D</i>	<i>14</i>
<i>Figura 2.4. Exemplo de transmissão de pacotes de sondagem (probe-request) pelos smartphones</i>	<i>17</i>
<i>Figura 2.5. Exemplo da taxa de envio de pacote</i>	<i>19</i>
<i>Figura 3.1. Etapas da metodologia</i>	<i>25</i>
<i>Figura 3.2. Exemplo da intersecção das zonas com a linha de ônibus</i>	<i>27</i>
<i>Figura 3.3. Equipamentos.....</i>	<i>29</i>
<i>Figura 3.4. Exemplo de comparação de matriz</i>	<i>31</i>
<i>Figura 4.1. Divisões espaciais da cidade de Aparecida de Goiânia</i>	<i>34</i>
<i>Figura 4.2. Linha 975</i>	<i>35</i>
<i>Figura 4.3. Distância entre os pontos de parada da linha 975</i>	<i>36</i>
<i>Figura 4.4. Histograma da distância entre pontos de parada da linha 975</i>	<i>36</i>
<i>Figura 4.5. Raio de influência da linha 975 para as diferentes divisões territoriais na cidade de Aparecida de Goiânia.....</i>	<i>37</i>
<i>Figura 4.6. Exemplo dos pontos de parada na borda dos setores censitários</i>	<i>38</i>
<i>Figura 4.7. Boxplot da frequência de transmissão dos smartphones</i>	<i>42</i>
<i>Figura 4.8. Matriz O/D manual ponto a ponto</i>	<i>43</i>
<i>Figura 4.9. Matriz O/D manual para as diferentes divisões territoriais</i>	<i>45</i>
<i>Figura 4.10. Espacialização dos sinais emitidos pelo smartphone A0:4F:85:D7:12:DC</i>	<i>47</i>
<i>Figura 4.11. Espacialização dos sinais emitidos pelo smartphone C8:F3:19:60:48:77</i>	<i>48</i>
<i>Figura 4.12. Espacialização dos sinais emitidos pelo smartphone E4:58:E7:9D:CF:EF</i>	<i>48</i>
<i>Figura 4.13. Espacialização dos sinais emitidos pelo smartphone 68:76:4f:0a:60:da</i>	<i>49</i>
<i>Figura 4.14. Espacialização dos sinais emitidos pelo smartphone 78:A8:73:95:37:A6.....</i>	<i>50</i>
<i>Figura 4.15. Matriz O/D ponto a ponto obtida</i>	<i>50</i>
<i>Figura 4.16. Matriz O/D obtida para as macrozonas.....</i>	<i>51</i>
<i>Figura 4.17. Matriz O/D obtida para as zonas de tráfego</i>	<i>52</i>
<i>Figura 4.18. Matriz O/D obtida para os setores censitários.</i>	<i>54</i>

LISTA DE ABREVIATURAS E SÍMBOLOS

ABREVIATURAS

APC	Contagem Automática de Passageiro
GPS	<i>Global Positioning System</i>
IEEE	<i>Engenheiros Eletricistas e Eletrônicos</i>
KML	<i>Keyhole Markup Language</i>
MAC	<i>Media Access Control</i>
O/D	Origem/Destino
RMTC	Rede Metropolitana de Transporte Coletivo
RSSI	<i>Received Signal Strength Indication</i>
SBE	Sistema de Bilhetagem Eletrônica

SÍMBOLOS

\mathcal{P}	Pacotes capturados
\mathcal{F}	Função de classificação

1 INTRODUÇÃO

O planejamento urbano é uma atividade fundamental para o desenvolvimento de aglomerados populacionais (GUERRA, 2011). Na ausência deste planejamento, as cidades crescem descontroladamente, o que pode ser percebido em diversas capitais brasileiras (SOUZA, 2014). O mesmo está ligado a ordenação do espaço, como uso e ocupação do solo, que conjuntamente com o transporte pode explicar e condicionar o desenvolvimento das cidades (CUNHA, 2005).

Neste contexto, o transporte público urbano é um elemento importante para os deslocamentos dos aglomerados populacionais. Este sistema de transporte possibilita a ligação em massa entre diferentes usos e ocupações de solo, podendo assim auxiliar na resolução de problemas recorrentes dos centros urbanos, tais como o congestionamento e a poluição (RODRIGUES; SORRATINE, 2014).

Entretanto, recentemente o transporte público urbano vem perdendo usuários ano após ano (SILVA; SILVA, 2018). Dentre os fatores para tal perda, pode-se destacar a lotação dos veículos, o tempo de espera e o aumento no valor da tarifa (GUIMARÃES; COELHO FILHO, 2018). Além de fatores econômicos, estes problemas são decorrentes de uma falha no planejamento e entendimento do transporte público urbano (SOUZA, 2014).

Estudos indicam que a compreensão dos deslocamentos dos passageiros é um elemento chave para auxiliar no planejamento e na otimização do sistema de transporte público (HOLTERMANN, 2016). Tal tarefa consiste em conhecer a movimentação dos passageiros, ou seja, conhecer o número de viagens realizadas entre locais de origem e de destino em um determinado período de tempo (BERTONCINI, 2007). Após a obtenção desses deslocamentos pode-se construir uma matriz de origem e destino dos mesmos, daí o nome matriz O/D, para quantificar e sintetizar a mobilidade associada aos passageiros no sistema (CACERES; WIDEBERG; BENITEZ, 2009).

As matrizes O/D dos passageiros são elementos de entrada essenciais para diversos métodos do planejamento de transporte, desde o modelo clássico de quatro etapas a modelos sofisticados de simulação, que possibilitam modificar e otimizar rotas existentes ou acrescentar novas rotas visando a melhoria da qualidade do transporte público (PEIXOTO, 2013). De fato, conforme Loureiro *et al.* (2011), através da matriz O/D pode-se escrever os fluxos de viagens realizados existentes e potenciais ao longo das regiões nas cidades. Dessa forma, a identificação de onde os passageiros embarcam e

desembarcam nos veículos durante seus deslocamentos são etapas cruciais para compreender e propor melhorias no transporte.

A estimativa dessa matriz pode ser feita de diferentes formas (HAKEGARD; MYRVOLL; SKOGLUND, 2018). Tradicionalmente, a forma mais comum de determinação da matriz é através do método de pesquisa direto, que é realizada em campo. Dentre tais modos, Richardson *et al.* (1995) destaca: as pesquisas de observação, a pesquisa domiciliar e as de abordagem. Na primeira, os pesquisadores realizam a contagem do “sobe e desce” dos passageiros nos pontos de parada. Já na segunda, o pesquisador dirige-se até o domicílio e aplica um formulário com intenção de descobrir as atividades de deslocamento dos moradores. E por fim, ressalta-se a pesquisa de abordagem, onde o entrevistado é abordado no meio de suas atividades e responde a respeito dos deslocamentos realizados.

Entretanto, tais técnicas estão cada vez mais difíceis de serem realizadas, devido à dificuldade de acesso aos entrevistados, isto é, a dificuldade em entrevistar os passageiros e dos mesmos terem disponibilidade para responderem às perguntas. Além disso, os custos e o tempo de execução dessa técnica são onerosos, restringindo a sua aplicação para prazos mais longos (GUERRA, 2011), uma vez que é necessário utilizar diversos pesquisadores para cobrir uma única linha de pesquisa. Desta forma, com o passar do tempo, a matriz O/D obtida manualmente se torna defasada, pois não apresenta uma repetibilidade de captura dos dados ao longo dos anos, devido seu alto custo de execução, ao passo que as cidades e, conseqüentemente, a linha do sistema de transporte comumente sofrem modificações ao longo do tempo (PELLETIER; TRÉPANIER; MORENCY, 2011).

Devido às dificuldades para obtenção da matriz O/D de forma direta, pesquisadores têm buscado explorar alternativas para obtenção ou atualização das matrizes existentes. Entre tais alternativas pode-se destacar a utilização da bilhetagem eletrônica. Tal sistema consiste em uma forma de pagamento da tarifa do transporte público que pode ser realizado no interior do veículo ou antecipadamente, por meio de um bilhete comum ou magnético, ficha plástica ou metálica e cartão plástico com informação magnética (FERRAZ; TORRES, 2004).

Entretanto, esses dados são projetados para um propósito específico, que é a cobrança da tarifa, que se limita ao local de validação do embarque, visto que o mesmo raramente precisa validar ao descer do veículo ou entre as viagens (*e.g.*, em terminais).

Desta forma, os dados acabam não contemplando o destino final dos passageiros (DUNLAP; HENRICKSON; WANG, 2016).

Há algumas maneiras de resolver o problema de obtenção do destino do passageiro utilizando a bilhetagem eletrônica. Dentre essas, pode-se destacar o método apresentado por Nielsen (NIELSEN, 1998), que cruza diversos dados de embarque ao longo do dia (GUERRA, 2011). Especificamente, utiliza-se o próximo ponto de validação como destino do ponto prévio. Entretanto, existe a dificuldade em capturar as linhas utilizadas no deslocamento, visto que ao cruzar dois pontos pode-se ter múltiplas rotas disponíveis para atingir tal destino, por exemplo, em terminais de transbordo.

Em paralelo, a comunidade científica começou a pesquisar a viabilidade de se usar os próprios dispositivos móveis dos passageiros como meio para obtenção desses dados (DUNLAP; HENRICKSON; WANG, 2016; MISHALANI; MCCORD; REINHOLD, 2016; JI; ZHAO; DU, 2017; EL-TAWAB *et al.*, 2017; KURKCU; OZBAY, 2017; HAKEGARD; MYRVOLL; SKOGLUND, 2018 ; HIDAYAT; TERABE; YAGINUMA, 2018; EL-TAWAB *et al.*, 2019; AFSHARI *et al.*, 2019). A possibilidade de rastrear dispositivos móveis, identificando o tráfego de dados Wi-Fi, atraiu interesse significativo, uma vez que o uso de dispositivos móveis pessoais (*e.g.*, *smartphones*, *tablets* e computadores portáteis) vem experimentando um crescimento nos últimos anos e desempenhando um papel importante na vida das pessoas. No ano de 2018, o Brasil ocupou o 4º lugar dos países que mais utilizaram o *smartphone*, sendo o primeiro da América Latina (TELECO, 2018). Esse interesse surgiu devido à necessidade de obter grande quantidade de dados estatísticos, uma vez que existem relatos que uma parte significativa das pessoas podem deixar seus dispositivos com o Wi-Fi habilitado (HAKEGARD; MYRVOLL; SKOGLUND, 2018).

Especificamente, têm-se estudado a utilização de sinais Wi-Fi emitido pelos *smartphones* dos passageiros. Isto decorre, pois, os sinais são constantemente transmitidos pelo *smartphones* no processo de descoberta de rede.

Como o pacote de dados é transmitido em público, existem sensores que podem capturar tais pacotes, rastrear o dispositivo e então detectar o deslocamento feito ao longo do tempo. Isto porque os pacotes emitidos pelos dispositivos podem ser identificados pelo endereço MAC (*Media Access Control*) que é um identificador exclusivo do dispositivo, semelhante a uma impressão digital (AFSHARI *et al.*, 2019).

Os pacotes de sondagem de rede possuem um identificador do emissor (MAC) e um indicador de potência do sinal da onda recebida, denominada *Received Signal*

Strength Indication (RSSI), que identifica a proximidade que o emissor está do receptor. Trabalhos recentes utilizam a quantidades de pacotes e a potência dos sinais recebidos, para estimar o número de passageiros no ônibus (RORIZ JUNIOR; MEDRANO, 2018).

1.1 PROBLEMA

Apesar de resultados significativos, o sucesso do método da construção da matriz O/D a partir dos dados está diretamente associado ao fato de o Wi-Fi dos dispositivos estarem ligados. O método é proporcional a quantidade de dados gerados, isto é, se o aparelho emitir pouco pacotes, tende-se a obter uma estimativa insatisfatória. Sabe-se que a quantidade de pacote transmitido está relacionado a utilização do *smartphone*, por exemplo, quanto mais o usuário utiliza seu *smartphone*, mais dados Wi-Fi são emitidos e menor é o tempo da taxa de envio de pacote (JAMIL *et al.*, 2016). Entretanto, se o usuário não utilizar seu dispositivo, por exemplo, o mesmo estiver no bolso ou com a tela desligada, a taxa de envio seguirá um padrão de envio conforme o modelo do *smartphone* (FREUDIGER, 2015).

A marca e modelo do *smartphone* influencia nessa taxa de envio, Dunlap *et al.* (2016), realizou um experimento com diversos aparelhos e constatou que alguns dispositivos apresentavam uma frequência maior dependendo da versão. Isto decorre devido o padrão do Wi-Fi não estabelecer uma frequência mínima para transmissão dos sinais de presença, deixando a especificação da taxa de envio a cargo das empresas (CROW *et al.*, 1997; GAST, 2005). Por exemplo, um smartphone com o Android KitKat 4.4.2 transmite em média um pacote de presença Wi-Fi a cada 72 segundos, enquanto que um aparelho com o sistema iOS 8.1.3 o faz a cada 330 segundos (FREUDIGER, 2015).

Como consequência disso, pode existir uma dessincronização entre a posição de envio do sinal de presença Wi-Fi com a de embarque do passageiro (PARADEDA; KRAUS JUNIOR; CARLSON, 2019; RORIZ JUNIOR; MEDRANO, 2018). Este fator complica a captura da matriz O/D, uma vez que o passageiro pode embarcar em um ponto de parada e o pacote de presença Wi-Fi do seu dispositivo ser transmitido e capturado apenas após um determinado período (*e.g.*, 330 s) em um ponto diferente (MIKKELSEN *et al.*, 2016). Semelhantemente, o problema também afeta a obtenção do par de destino, uma vez que o passageiro pode desembarcar em um ponto de parada e ter tido o último pacote Wi-Fi enviado há três minutos.

Tais problemas são inerentes a tecnologia Wi-Fi, visto que são causados pela falta de padronização da frequência de envio dos pacotes de presença na especificação (GAST,

2005). Esta limitação dificulta a captura de uma matriz O/D ponto a ponto (PARADEDADA; KRAUS JUNIOR; CARLSON, 2019), pois a demora para a transmissão dos pacotes Wi-Fi pode dessincronizar a captura dos pontos de embarque e desembarque do passageiro. Os resultados preliminares de Sousa Júnior *et al.* (2019) mostram que a maioria dos pares O/D diferem em até 4 pontos de parada, isto é, a diferença entre o ponto de embarque e desembarque real do passageiro daquele estimado utilizando a tecnologia Wi-Fi é de 4 pontos na maioria dos casos. Por exemplo, quando o passageiro embarca, pode ocorrer de o *smartphone* do usuário demorar para transmitir o pacote com o sinal de presença Wi-Fi e o mesmo ser observado somente no quarto ponto de parada.

Considerando esses fatores, percebe-se uma dificuldade para implantar a tecnologia Wi-Fi para a captura de uma matriz O/D ponto a ponto no sistema de transporte público. A precisão dos dados obtidos pela tecnologia Wi-Fi diverge da especificidade ponto a ponto das linhas do sistema de transporte público. Entretanto, é importante ressaltar que a utilização da tecnologia pode possibilitar um acompanhamento em aproximadamente tempo real da movimentação dos passageiros (CUNCHE; KAAFAR; BORELI, 2014; ORANSIRIKUL *et al.*, 2016). Esse acompanhamento, mesmo com as limitações apontadas, pode capturar as mudanças e anomalias substanciais nas linhas do sistema, que por sua vez pode ser usada para disparar uma análise mais profunda, como uma pesquisa manual da mesma pelas operadoras.

Desta forma, esse estudo visa explorar como a matriz O/D obtida pela tecnologia Wi-Fi se assemelha com aquela obtida manualmente ao considerar uma divisão espacial menos granular da linha do sistema de transporte, isto é, ao invés de dividir a linha em ponto a ponto usar divisões espaciais maiores, tais como bairros, setores censitários, macrozonas e zonas de tráfego (KNEIB, 2008). Precisamente, o trabalho visa investigar a similaridade da matriz O/D obtida pela tecnologia Wi-Fi ao relaxar a divisão espacial da linha em espaços maiores.

1.2 HIPÓTESE

Ao considerar a limitação da dificuldade de captura de uma matriz O/D ponto a ponto com a tecnologia Wi-Fi, supõe-se que uma subdivisão da linha de ônibus em análise de forma menos granular, em relação a divisão ponto a ponto, pode mitigar os efeitos colaterais da tecnologia, como o atraso no envio de dados, e apresentar uma matriz com maior similaridade a realizada pelos passageiros. Espera-se que ao considerar que a linha de ônibus pode ser dividida em áreas maiores, como bairros, setores censitários,

macrozonas e zonas de tráfego, os resultados das matrizes O/D obtidas usando Wi-Fi e a matriz manual aumentarão e assim possibilitarão o uso da tecnologia para estimações futuras do sistema de transporte público.

1.3 OBJETIVO GERAL E ESPECÍFICO

O objetivo geral deste estudo é verificar a viabilidade da reconstrução da matriz O/D a partir da utilização e captação de dados Wi-Fi dos *smartphones* dos clientes que utilizam o transporte público, considerando as diferentes divisões urbanas, ou seja, divisões menos granulares que a divisão ponto-a-ponto como, por exemplo, as zonas de tráfego, bairros, setores censitários e macrozonas. Este estudo tem como objetivo específico a comparação da similaridade entre as matrizes O/D construídas com diferentes granularidades entre si.

1.4 CONTRIBUIÇÕES

A pesquisa se concentra em investigar a viabilidade de se utilizar os dados Wi-Fi, para estimar a matriz O/D entre diferentes divisões territoriais. De modo a facilitar o processo de conduzir pesquisas com O/D. Apesar de isto ser abordado em outros estudos, tal tecnologia visa contribuir através de construção e validação da matriz estimada.

1.5 DISTRIBUIÇÃO DO TRABALHO

As próximas seções deste trabalho serão divididas da seguinte forma, a Seção 2 apresentará a revisão bibliográfica, que aborda conceitos básicos das divisões territoriais das cidades e conceitos que envolvem as metodologias de estimação da matriz O/D e formas de comparação entre as mesmas. Ainda na revisão bibliográfica são apresentados os trabalhos relacionados com o tema. Já na Seção 3 é apresentada a metodologia do trabalho, contemplando assim o método proposto para estimação da matriz O/D e como foi feito o experimento. Na Seção 4 é apresentado o experimento. E por fim na Seção 5 são apresentadas as considerações finais do trabalho.

2 REVISÃO BIBLIOGRÁFICA

Este trabalho se baseia nos conceitos relacionados à tecnologia Wi-Fi e aos métodos de obtenção de matrizes O/D, que por sua vez estão atrelados a divisão territorial da cidade em locais de origens e destinos. Com intuito de familiarizar o leitor com os conceitos e premissas adotadas, além dos trabalhos atuais na área, esta revisão bibliográfica divide-se em duas partes. A primeira descreve concisamente os conceitos fundamentais do trabalho, já a segunda explora os principais trabalhos relacionados à temática.

2.1 DIVISÕES DO ESPAÇO

Devido as dificuldades encontradas na obtenção da matriz O/D ponto a ponto, como por exemplo, o atraso de envio de pacote. Esta seção apresentará as diferentes divisões territoriais das cidades, de modo a auxiliar na estimação dos pares de origem e destinos.

2.1.1 DIVISÕES TERRITORIAIS DAS CIDADES

As linhas de ônibus ligam diferentes áreas de aglomerados urbanos. Como mencionado anteriormente, as limitações da tecnologia Wi-Fi dificultam uma análise pontual do deslocamento dos passageiros nas linhas. Nesse sentido, o estudo visa explorar se as divisões menos granulares da linha podem contribuir para a captura desta matriz. Assim, inicia-se este estudo com a compreensão de como as divisões territoriais se deram no país e como elas se entrelaçam com as linhas do sistema de transporte.

De acordo com Cigolini (2009), a maneira como conhecemos de divisão territorial atual teve início no final do século XX. Tais divisões visavam acompanhar a evolução urbana dos municípios que começaram a se formar nesse período. Também se observa que se pode estimular o desenvolvimento econômico e social de uma região ou de partes da cidade através da delimitação dos espaços urbanos, como relatado em Boisier (2009).

Atualmente as divisões territoriais elaboradas pelo Estado brasileiro seguem diretrizes (*e.g.*, IBGE) ou resoluções legais (*e.g.*, Plano Diretor) (HIDALGO, 2013). Por exemplo, o setor censitário deve ser delimitado por resoluções municipais, como o plano diretor. Mas na ausência, o próprio órgão censitário, em colaboração com o município, delimita essas regiões (IBGE, 2013). Tendo em vista isso, as divisões apresentadas neste trabalho seguirão a definição daquelas descritas nas diretrizes e resoluções federais, estaduais e municipais disponibilizadas por setores da administração pública.

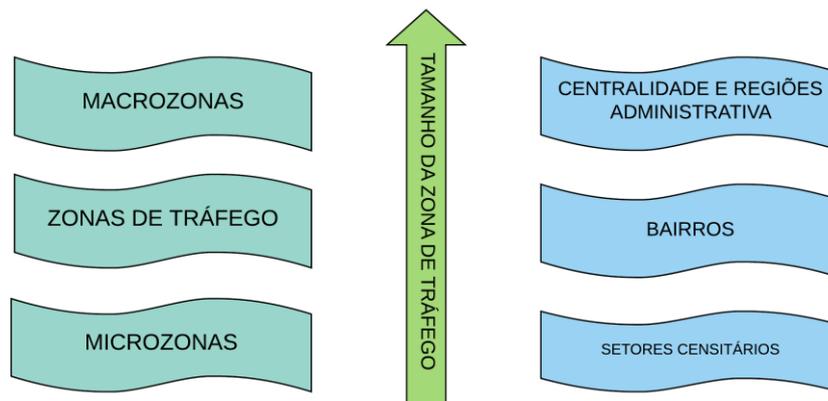
Dentre os trabalhos citados acima destaca-se a resolução do plano diretor. Segundo o Estatuto da Cidade (2002), o mesmo é definido como um instrumento básico da política de desenvolvimento e expansão urbana. O Estatuto da Cidade (2002) estabelece que municípios com mais de 20.000 habitantes devam elaborar e revisar o plano diretor a cada 10 anos. O plano diretor servirá como um facilitador do planejamento urbano já que é uma ferramenta de ordenamento social que visa identificar quais os problemas enfrentados pela cidade e apontar ações para solucioná-los.

Como abordado anteriormente, o plano diretor é uma fonte de informação para o planejamento urbano. Desse modo, o mesmo deve apresentar as divisões territoriais atuais e futuras para cada cidade, ou seja, cada cidade ficará responsável por definir uma metodologia para determinar essas divisões. No entanto, ao adotar metodologias diferentes, cada cidade poderá apresentar aspectos únicos em suas divisões, tornando assim difícil uma caracterização comum para todas as cidades (TACO; TEDESCO; GUERRA, 1997).

Desse modo, o trabalho visa adotar o princípio de divisão por zonas de tráfego, principalmente ao fato de que as zonas de tráfegos podem variar a sua granularidade (resolução). Tal decisão é baseada em Ortuzar (2011), que destaca que as zonas de tráfego devem apresentar uma compatibilidade com as outras divisões administrativas como por exemplo, os bairros, e os setores censitários. Ao escolher divisões já existentes pode-se aproveitar os dados já disponíveis, para complementação da pesquisa.

Com intuito de exemplificar estes conceitos, a Figura 2.1 ilustra a relação hierárquica entre as divisões espaciais. Como dito, dependendo da granularidade desejada, as zonas podem assumir tamanhos macros (Macrozonas) ou locais (Microzonas). Para exemplificar isso, o lado direito mostra divisões espaciais comumente atreladas ao conceito utilizado. Por exemplo, as centralidades e regiões administrativas são comumente atreladas a macrozonas, enquanto que as microzonas podem ser equiparadas aos setores censitários ou até mesmo os bairros. As subseções a seguir descrevem concisamente a relação destas subdivisões.

Figura 2.1: Exemplo da relação hierárquica da divisão do espaço com o conceito de zona de tráfego.



Fonte: Autor (2020).

2.1.1.1 ZONAS DE TRÁFEGO (ZTs)

As Zonas de Tráfego podem ser definidas como frações da área urbana que apresentam características e uso do solo semelhantes entre si, condicionando a uma ocupação ordenada e funcional (KNEIB, 2008). Para Campos (2013) as zonas de tráfego têm a finalidade de estabelecer a quantidade de fluxo gerado pelos movimentos básicos.

Ortuzar (2011) ressalta que um dos problemas no planejamento de transporte é determinar o nível de resolução (granularidade) e de detalhes a serem adotados. Tendo em vista esse aspecto, o autor destaca dois fatores que influenciarão os limites dessas áreas. Primeiramente, toma-se como base o caráter do estudo, ou seja, se o estudo é estratégico, por exemplo, com intuito de previsão futura da demanda na cidade, utiliza-se zonas maiores, enquanto estudos detalhados, como intervenções locais, é preferível que as mesmas sejam menores e mais detalhadas. Segundo, ressalta-se os recursos disponíveis para o estudo. Nota-se que, quanto maior a necessidade de detalhamento e número de zonas maior a exatidão do estudo, entretanto mais recursos serão necessários para captura dos respectivos dados.

2.1.1.2 MACROZONAS

A macrozona é o primeiro nível de definição espacial para o plano diretor, ou seja, é o referencial espacial para o uso e a ocupação do solo na cidade, em concordância com as estratégias de políticas urbanas (BRASIL, 2002). De acordo com a Constituição Federal de 1988 as macrozonas devem compreender todo o território municipal, contemplando as zonas urbanas e rural. Além disso, as macrozonas devem ser formadas

por frações da área urbana com características semelhantes de ocupação, economia, meio ambiente e a infraestrutura urbana (PLANO DIRETOR DO RIO DE JANEIRO, 2011).

As macrozonas são ferramentas utilizadas no planejamento urbano, que visam a compreensão das diferentes realidades de um município, orientam o planejamento e auxiliam em definições de políticas públicas principalmente aquelas que estão relacionadas com o processo de urbanização (PLANO DIRETOR DE CAMPINAS, 2006).

Com base nos trabalhos pesquisados as macrozonas não são comumente utilizadas para a elaboração das matrizes O/D, visto que a granularidade dessas divisões territoriais é bem maior que as demais, tornando assim difícil a visualização das movimentações entre as diferentes partes de uma cidade.

Algumas cidades, como por exemplo Aparecida de Goiânia, usam o conceito de divisão por macrozonas para definir as centralidades e regiões administrativas no plano diretor da cidade. Tal questão retorna ao fato de que cada cidade pode adotar sua metodologia para as divisões territoriais.

2.1.1.3 BAIRROS

O conceito e alcance da noção do que é um bairro é bastante discutido pela comunidade acadêmica, sejam por engenheiros, geógrafos ou arquitetos. Desta forma, esta seção apresenta algumas das principais definições do mesmo.

Souza (1987) tem uma percepção diferente sobre a definição de bairros, o autor tem a preocupação com a questão da identidade do indivíduo a esta unidade territorial, em seu trabalho realizado no interior de São Paulo o autor ressalta que os elementos físicos e laços afetivos estão intimamente ligados a população do bairro (BEZERRA, 2011).

Wilheim (1982) define bairros como uma unidade urbana, que representa a espacialidade da população, em seu trabalho o autor cita que a cidade de São Paulo na época contava com 900 “sociedade de moradores” cuja territorialidade é facilmente estabelecida.

Já Lefebvre (1975) interpreta o bairro com uma forma concreta do espaço e do tempo na cidade, que atua como um módulo social de maior convergência entre o espaço geométrico e o espaço social entre o quantificado e o qualificado (BEZERRA, 2011).

De acordo com Cazollato (2005), as vantagens da utilização desse tipo de divisão se dão que grande parte das cidades brasileira apresentam uma divisão territorial semelhante a bairros facilitando assim a coleta e análise dos dados. Tais resultados

poderão auxiliar os órgãos públicos na elaboração de ações voltado ao planejamento urbano e auxílio na fiscalização de ações públicas.

Deve-se ressaltar que apenas as cidades planejadas possuem bairros com limites bem definidos, ou seja, a maioria das cidades brasileiras possuem os limites definidos arbitrariamente, o que leva a concluir que o conceito de bairro como citado anteriormente está confundido com própria evolução cultural e econômica da cidade. Como consequência um bairro pode possuir diferentes naturezas de tráfego, o que pode gerar dificuldades na elaboração de ações específicas para cada bairro. Por fim, destaca-se que apesar dos limites rígidos dos bairros, nem todos municípios fornecem tal delimitação em formato digital. Este é o caso do município de Aparecida de Goiânia e muitos outros municípios. Isto dificulta a utilização desta subdivisão em modelos computacionais.

2.1.1.4 SETORES CENSITÁRIOS

De acordo com Kneib (2008), os setores censitários podem ser definidos com uma unidade territorial para fins de controle cadastral da coleta de dados do Censo Demográfico. Para o IBGE os setores censitários devem possuir limites físicos identificáveis que respeitem os limites políticos-administrativos do quadro urbano e rural legal e de outras estruturas territoriais de interesse. Além disso, deve possuir um quantitativo de domicílios adequado para operação censitária (IBGE, 2013).

Se o município não possuir legislações urbanística capaz de realizar tais divisões o próprio IBGE realizará essas definições com a devida aprovação municipal. Essas informações para que ocorra a divisão são pautadas principalmente em definições legais para caracterizar a área urbana do município. Contudo, não se encontra informações na bibliografia do instituto que explicitem quais critérios foram utilizados para delimitação do que denominam “perímetro para fins de coleta censitária” (PERA; BUENO, 2016).

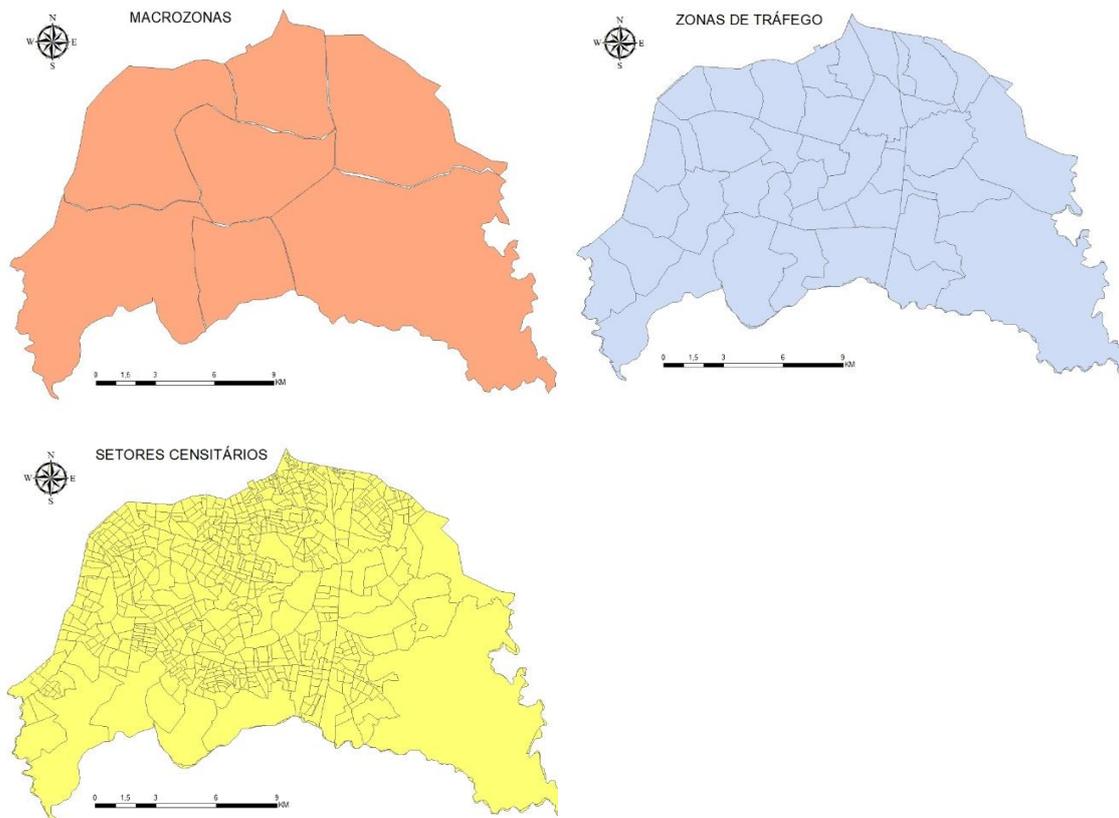
Desta forma, o setor censitário é desejável pois, é uma divisão territorial que todo município deve ter devido as pesquisas censitárias. Isto facilita a análise dos dados visto que a divisão espacial é conhecida. Além disso tem a vantagem que a mesma é facilmente encontrada em formato digital.

A publicação Metodologia do Censo 2010 (IBGE, 2013) faz apenas uma menção que para a realização das bases cartográficas que orientaram a delimitação dos setores censitários de 2010 foram realizadas reuniões periódicas que foram chamadas de “Comissões Municipais de Geografia e Estatística”. Tal iniciativa foi criada pelo próprio IBGE e foi constituída com representantes municipais que tinha a função de acompanhar

o preparo da base cartográfica municipal. Essas bases podem ser adquiridas no site do IBGE.

A Figura 2.2 apresenta alguns exemplos das divisões citadas acima, para a cidade de Aparecida de Goiânia, dentre elas pode-se destacar macrozonas, zonas de tráfego e setores censitários. Pode -se notar que a macrozona divide o município de Aparecida em 7 áreas, já a zona de tráfego divide a cidade em 40 áreas, e por fim os setores censitários são divididos em mais de 600 áreas.

Figura 2.2: Exemplo de divisões territoriais do município de Aparecida de Goiânia.



Fonte: Autor (2020).

Dados: Setores Censitários (IBGE), macrozonas e zonas de tráfego (plano diretor de Aparecida de Goiânia).

2.2 DETERMINAÇÃO DA MATRIZ ORIGEM/DESTINO

Com o entendimento de quais são as divisões do espaço, pode-se então prosseguir para compreender como ocorre a movimentação dos passageiros entre as diferentes áreas. De acordo com Caceres *et al.* (2008), as matrizes O/D são utilizadas para quantificar e sintetizar a mobilidade associada a pessoas e bens. De modo que essas matrizes fornecem informações sobre o número de viagens realizadas a partir de uma região de origem e uma de destino durante um intervalo de tempo.

As definições das áreas irão depender da resolução e do detalhamento adotado pelo estudo. Ainda segundo Caceres *et al.* (2008), as áreas que compõem as origens e destinos de uma matriz, ou seja, as regiões onde se iniciam e finalizam as viagens são definidas por critérios sociais e econômicos ou agregados populacionais que correspondem com áreas que apresentam relação de causa com os movimentos de transporte que ocorre entre eles.

Nesse contexto, Guerra (2011) afirma que uma maneira de representar a demanda por viagem no transporte público é através do número de passageiros que viajam de uma região de origem i para uma região de destino j , ou seja, neste cenário a demanda é equivalente ao número de usuários que se deslocam de uma área para outra. Note que não há necessidade de registro da viagem por todas as zonas percorridas, por exemplo, um passageiro pode deslocar da primeira zona de tráfego para a quinta. Suponha que este trajeto necessariamente percorra a segunda e quarta zona. Neste caso, o registro da viagem não será duplicado nestas zonas. Especificamente, a produção e a atração da viagem serão computadas apenas na primeira e quinta zona respectivamente

A Figura 2.3 esquematiza os elementos de uma matriz O/D desta forma. Observe que as colunas e as linhas representam respectivamente as zonas de tráfego de origem (O_i) e as zonas de tráfego de destino (D_j). Com a soma das linhas ou das colunas para cada zona (1 a n), pode-se encontrar a quantidade de viagens atraídas (A_j) ou produzidas (P_i) por cada zona de tráfego. Já os elementos que compõem a matriz d_{ij} representam a demanda com origem na zona de tráfego i e destino em uma zona de tráfego j .

Figura 2.3: Elementos de uma matriz O/D.

O/D	1	2	...	n	
1	d_{11}	d_{12}	...	d_{1n}	P_1
2	d_{21}	d_{22}	...	d_{2n}	P_2
⋮	⋮	⋮	...	⋮	⋮
n	d_{n1}	d_{n2}		d_{nn}	P_n
	A_1	A_2	...	A_n	A/P

Fonte: Adaptado de GUERRA (2011).

A construção dessa matriz pode ser realizada de forma direta ou indireta. As subsecções seguintes descrevem concisamente as etapas envolvidas em cada forma.

2.2.1 MÉTODOS DIRETOS

Os métodos diretos usados para construção da matriz O/D são baseados em pesquisas de campo. Ferreira (1999) afirma que mesmo para coleta de pequenas amostras do fluxo dos passageiros O/D, os métodos requerem uma quantidade significativa de recursos financeiros e de tempo para execução.

Entre os diversos métodos diretos de pesquisa de matriz O/D apresentado em trabalhos anteriores (GUERRA, 2011; RICHARDSON *et al.* 1995; MESQUITA; AMARAL, 2016), pode-se destacar: a pesquisa de entrevista domiciliar, a pesquisa de abordagem ao entrevistado e a pesquisa de observação, todas realizadas manualmente.

As pesquisas domiciliares podem ser classificadas em dois tipos: a de entrevista pessoal e a de autopreenchimento. A pesquisa de entrevista pessoal é aquela realizada na residência do entrevistado em que o pesquisador faz uma série de perguntas anotando as respostas em um formulário próprio. Já as pesquisas domiciliares de autopreenchimento, são aquelas feitas pelo entrevistado sem a ajuda de um pesquisador (RICHARDSON *et al.* 1995). A primeira esbarra em uma série de desafios, desde o custo até questões de segurança, uma vez que os pesquisadores precisam deslocar-se até a residência do entrevistado, que implica em custo de deslocamentos do pesquisador. Já o segundo tipo de pesquisa, possui uma baixa taxa de retorno, uma vez que o próprio entrevistado precisa responder o questionário sem a ajuda de um pesquisador, e por motivos diversos acaba não tendo um interesse ou meios para respondê-los.

Por outro lado, as pesquisas de abordagem são aquelas em que as pessoas são abordadas no curso de uma atividade, podendo assim a entrevista ser realizada no interior

do veículo do transporte público, em vias públicas, terminais e locais de trabalho. Os entrevistados recebem um questionário dentro do veículo que pode ser recolhido ali ou pode ser enviado pelos correios (RICHARDSON *et al.* 1995). As vantagens dos métodos de pesquisa domiciliares e de abordagem é que ambas podem obter informações detalhadas dos passageiros como, por exemplo, informações socioeconômicas.

Por conta destes problemas, as pesquisas comumente realizadas para obtenção dos deslocamentos nos transportes são as de observação devido ao tempo de execução, que podem ser menores que as demais (FERREIRA, 1999). Como consequência do menor tempo de execução, obtêm-se um custo menor. As pesquisas de observação podem ser diretas ou indiretas.

Alguns exemplos das pesquisas de observação direta são: descrição de pesquisa de transporte como, por exemplo, uso de técnicas como gravação de vídeo e veículos instrumentais, usado principalmente para contagem de tráfego. Também pode-se destacar a pesquisa de contagem de embarque e desembarque mais conhecida como “sobe/desce”, essa pesquisa pode ser utilizada em algumas situações para identificação da matriz O/D (FERREIRA, 1999).

As técnicas de observação indiretas também são bastante utilizadas em estudos de transporte. Entre elas se destacam a estruturação de bancos de dados governamentais (PNAD, CAGED, cadastros de imóveis e de veículos), além disso destaca a utilização em mapas digitais, como por exemplo, google *Earth*. (RICHARDSON *et al.* 1995).

Os métodos diretos de pesquisa geralmente precisam de grande volume de mão-de-obra e tempos de tabulação excessivamente longos e trabalhosos, dessa forma os métodos indiretos vêm como uma forma de aquisição de dados obtidos em campo de forma simples como será demonstrado na próxima seção.

2.2.2 MÉTODOS INDIRETOS

Os métodos indiretos visam extrair uma matriz O/D através de outros elementos presentes no transporte público, como os bilhetes eletrônicos e dados celulares utilizados pelos passageiros.

O Sistema de Bilhetagem Eletrônica (SBE) é definido como uma maneira de aquisição de créditos de viagem antecipada, por meio da gravação em cartões eletrônicos que são lidos em equipamentos validadores instalados nos ônibus e terminais (ANPT, 2012). Recentemente alguns trabalhos como Mesquita e Amaral (2016) e Guerra (2011) vêm utilizando o SBE como forma de obtenção de dados para estimação da matriz O/D.

Na metodologia de Mesquita e Amaral (2016) é demonstrado que os dados são adquiridos em forma de planilha eletrônica contendo informações das validações dos passageiros. Em seguida, realiza-se um pré-processamento dos dados para filtrar apenas os campos necessários para a geração da matriz O/D, como: número de identificação de cada passageiro, data e horário da validação e a localização geográfica de cada validação.

Alguns trabalhos como o de Guerra (2011) enriquecem os dados obtidos com o SBE com dados prévios de matrizes O/D obtidos manualmente. Essas matrizes, também chamadas se matrizes sementes, visam servir de meio para calibração do modelo adotado.

Embora os dados possam ser tratados e modelados para que fiquem conforme o desejado, o SBE apresenta limitações como, por exemplo as origens dos passageiros são associadas ao local de validação do cartão passageiro, o destino é obtido com base na próxima validação do passageiro uma vez que o passageiro não precisa validar novamente o seu cartão na saída do transporte público, tornando assim os dados descontínuos e fragmentados (GUERRA,2011).

Tendo em vista as limitações apresentadas, principalmente o fato de não saber onde a pessoa sobe/desce do veículo, a tecnologia Wi-Fi vem como uma tentativa de obter dados de embarque e desembarque dos passageiros de forma mais precisas comparada com o SBE.

2.2.2.1 WI-FI

Devido ao uso crescente de dispositivos móveis pessoais e a correspondente dependência de comunicação sem fio (Wi-Fi) em espaços públicos, a tecnologia vem sendo estudada como uma alternativa para obtenção de grandes quantidades de dados, principalmente na área do transporte.

Isto é possível devido ao funcionamento do protocolo Wi-Fi. O Protocolo IEEE 802.11, conhecido como Wi-Fi, foi desenvolvido e padronizado pelo Instituto dos Engenheiros Eletricistas e Eletrônicos (IEEE), o qual define e possibilita diretrizes para a comunicação sem fio (Wi-Fi) entre os dispositivos eletrônicos (CROW *et al.*, 1997).

Para cada atividade desenvolvida na conexão dos dispositivos o protocolo estabelece técnicas para sua realização. Dentre essas técnicas pode-se destacar o processo de descoberta da rede, que analisa quais dispositivos estão próximos por meio de pacotes de rede denominado *probe request* (RORIZ JUNIOR; MEDRANO, 2018).

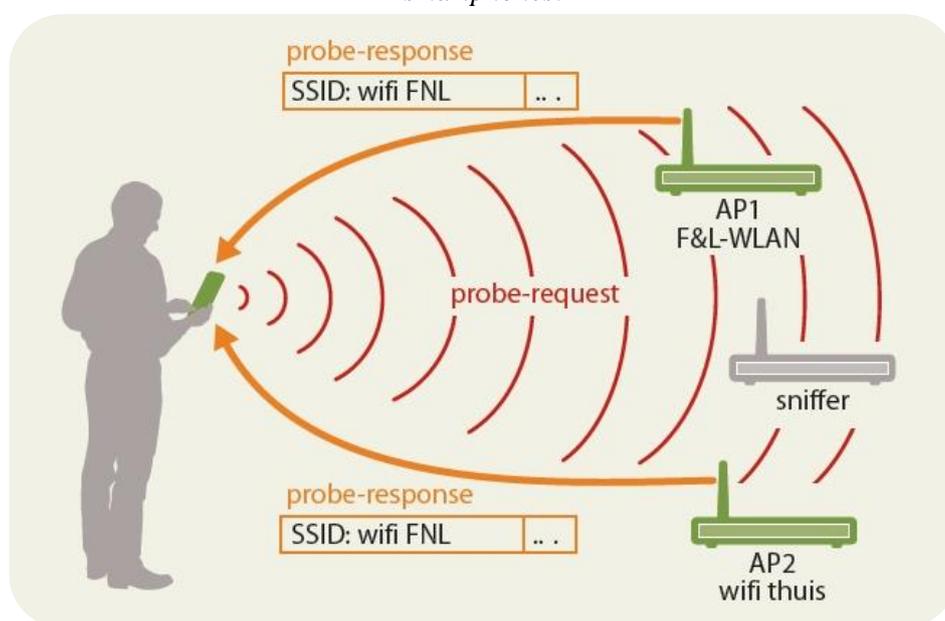
Esses pacotes de rede (*probe request*) são emitidos publicamente em uma faixa de onda próxima do emissor, todos os dispositivos interessados em conectar-se poderão

recebê-los. Especificamente, os dispositivos móveis ficam enviando periodicamente este pacote para sondar, daí o nome *probe*, as redes na sua vizinhança. Cada pacote de rede incluem um identificador de dispositivo exclusivo, uma espécie de impressão digital, do *smartphone* chamado de MAC (Média Access Control) (FREUDIGER, 2015).

Cada pacote inclui um identificador de potência do sinal recebido, denominado *Received Signal Strength Indication* (RSSI). Quando o pacote é recebido é observado a potência, então o receptor computa o identificador RSSI, em decibéis, que representa o quão próximo estão os aparelhos. Este indicador é comumente utilizado para indicar a potência da rede em aparelhos, sendo usado várias barras para indicar que a rede se encontra próxima. Especificamente, o valor do RSSI varia de 0 até $-\infty$, de forma que quanto mais próximo de zero, mais próximo estará o *smartphone* (RORIZ JUNIOR; MEDRANO, 2018).

Os pacotes de rede são enviados periodicamente em faixas de ondas públicas, desse modo qualquer computador que estiver monitorando o espectro da onda poderá receber as informações. Para o processo de captura de rede são utilizados computadores que são conhecidos como *sniffers*, eles podem utilizar esses pacotes para identificar se há ou não a presença de um *smartphone* nas proximidades (FREUDIGER, 2015). A Figura 2.4 representa uma pessoa com um *smartphone* que periodicamente envia pacotes para que o *smartphone* se conecte a um ponto de acesso (roteador) e nesse processo com a presença do *sniffer* é possível obter informações da localização do usuário.

Figura 2.4: Exemplo de transmissão de pacotes de sondagem (*probe-request*) pelos *smartphones*.



Fonte: Adaptado de (RORIZ JUNIOR; MEDRANO, 2018).

Para que ocorra o processo de captura de dados é necessário que o usuário do transporte público possua um aparelho *smartphone* e que o mesmo esteja com Wi-Fi ligado. Independentemente do *smartphone* estar conectado a uma rede, o dispositivo irá realizar a atividade de sondagem periodicamente. Isto é feito para atualizar as redes ao redor do usuário. No entanto, isso só pode ocorrer se o Wi-Fi estiver ligado. (FREUDIGER, 2015).

Para o processo de identificação da origem e destino do passageiro é necessário que o *sniffer* esteja localizado no interior do ônibus. Inicialmente o *smartphone* do passageiro enviará pacotes de rede para descobrir e atualizar as listas de redes próximas. Quando o passageiro entrar no veículo esses pacotes de redes estão no raio em que o *sniffer* poderá capturar seu sinal. Dessa forma, pode-se concluir que quando o *sniffer* capturar o primeiro sinal emitido pelo *smartphone* dentro do veículo, ali será considerado a origem do passageiro, assim como o último sinal recebido pelo *sniffer* será considerado como destino.

Uma das dificuldades de estimação das origens e destinos dos passageiros é identificar quais pacotes de sondagem estão dentro ou fora do veículo, visto que o *sniffer* pode capturar alguns sinais de *smartphone* que estão próximo do ônibus, como por exemplo, em um cruzamento onde vários veículos ficam próximos uns dos outros. Para a identificação de quais *smartphone* estão dentro ou fora do veículo, alguns modelos de classificação de pacotes de sondagem realizam apenas a comparação com limites arbitrários como por exemplo a potência do RSSI se o pacote recebido está abaixo do limiar superior, mas tal abordagem é problemática. Uma alternativa mais segura é o aprendizado de máquina realizado por Roriz Junior e Medrano (2018), nele são utilizados conceitos para estimação dos limites automaticamente a partir dos próprios dados.

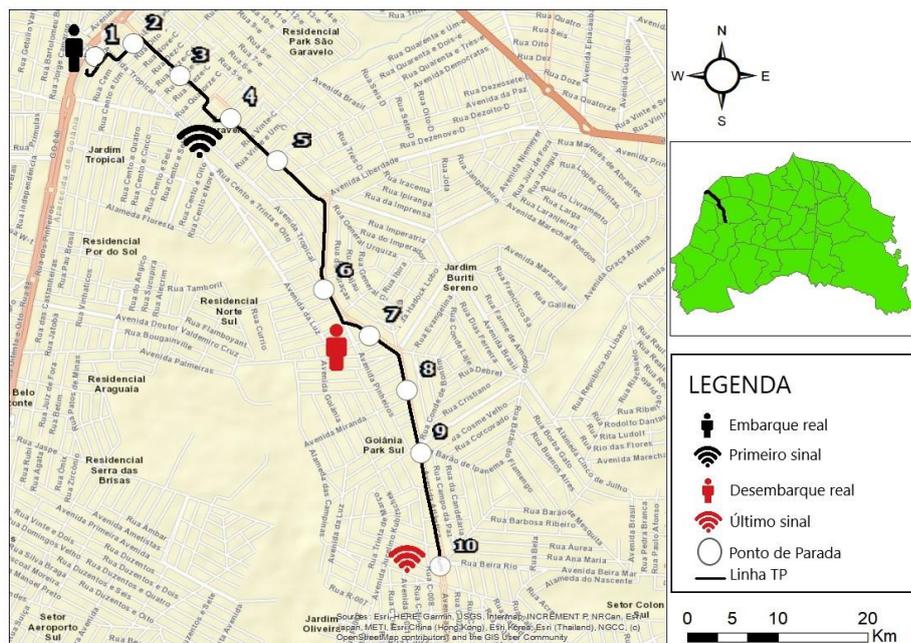
A principal dificuldade enfrentada na estimação da origem e do destino dos passageiros é o atraso do envio de pacotes de sondagem, uma vez que cada *smartphone* possui uma taxa diferente de envio de pacotes, ou seja, pode ocorrer que o passageiro embarque em um ponto, mas devido a marca e modelo do *smartphone* seus pacotes só serão enviados e, conseqüentemente, recebidos posteriormente. Desta forma, o primeiro pacote será capturado em um local adiante da linha de ônibus, situado próximo a um ponto possivelmente diferente daquele em que o passageiro embarcou no veículo.

Um exemplo do atraso de envio de pacotes é citado no trabalho de Freudiger (2015), o autor aponta em seus resultados que o Android KitKat 4.4.2 apresenta uma taxa de envio de pacote de 72 segundos enquanto o iOS 8.1.3 apresenta uma taxa de 330 segundos.

Para exemplificar melhor como ocorre o atraso de envio de pacote considere uma linha do transporte público hipotética com 10 pontos de paradas e um passageiro com seu *smartphone*. Suponha que o embarque real do passageiro seja no primeiro ponto de parada e desembarque no sétimo ponto de parada, devido ao atraso de envio de pacote pelo *smartphone* é possível que o pacote seja enviado alguns segundos depois. A Figura 2.5 ilustra como seria o recebimento dos pacotes ao longo da linha.

A taxa de envio de pacote influenciará diretamente na estimação da matriz O/D dos passageiros uma vez que para inferir que o passageiro embarcou naquele ponto será considerado o primeiro recebimento do pacote de sinal.

Figura 2.5: Exemplo da taxa de envio de pacote.



Fonte: Autor (2020).

2.3 MÉTODOS DE COMPARAÇÃO DAS MATRIZES

Em posse da obtenção da matriz O/D manual e da sintetizada por dados Wi-Fi pode-se iniciar o processo de comparação das duas. É válido destacar que existem diferentes medidas para tal comparação, entre elas se destacam as apresentadas abaixo.

2.3.1 R-QUADRADO (R^2)

De acordo com Washington *et al* (2011) o R-quadrado é um dos mais comuns e utilizados nas pesquisas, é uma medida estatística de quão próximos os dados estão da linha de regressão ajustada. É utilizado para comparar os pares de origem e destino de duas matrizes O/D. Os valores de R^2 variam de 0 a 1, com valores mais altos indicando menor diferença entre as matrizes O/D. A fórmula de cálculo do R^2 está representada abaixo:

$$R^2 = 1 - \frac{SS_E}{SS_T} \quad (1)$$

$$SS_E = \sum_i (OD_{i,j}^1 - OD_{i,j}^2)^2 \quad (2)$$

$$SS_T = \sum_i (OD_{i,j}^1 - \overline{OD})^2 \quad (3)$$

Onde $OD_{i,j}^1$ é o par i, j da primeira matriz O/D, $OD_{i,j}^2$ é o par i, j da segunda matriz O/D e \overline{OD} é a média dos pares das matrizes O/D.

Considerando que quanto maior o valor do R mais similar serão as matrizes, a linha de regressão deve estar perto de uma linha de 45 graus através da origem.

2.3.2 RAIZ DO ERRO MÉDIO QUADRÁTICO (RMSE) E ERRO MÉDIO QUADRÁTICO (% RMSE)

O erro médio quadrático (RMSE) é o erro quadrático médio percentual da raiz (% RMSE) são usados para avaliar a proximidade das matrizes O/D. O % RMSE demonstra a variabilidade da demanda de forma mais evidente, ou seja, se duas matrizes O/D forem idênticas, a percentagem seria igual a zero. Os valores podem ser calculados das seguintes formas.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (OD_{i,j}^1 - OD_{i,j}^2)^2}{N}} \quad (4)$$

$$\%RMSE = \frac{RMSE}{\left(\frac{\sum_{i=1}^N OD_{i,j}^1}{N}\right)} * 100 \quad (5)$$

Onde $OD_{i,j}^1$ é o par i, j da primeira matriz O/D e $OD_{i,j}^2$ é o par i, j da segunda matriz O/D.

2.3.3 DISTÂNCIA DE HELLINGER (dHel)

O matemático Ernst David Hellinger (1883-1950) em um dos seus trabalhos apresentou um quantificador simples e de simetria clara que tem a finalidade de medir a distância entre distribuições de probabilidades discretas que ficou conhecida como distância de Hellinger (Hellinger, 1909).

Dessa forma a distância de Hellinger é definida para as distribuições de probabilidades da seguinte forma:

Seja $P = \{p_1, \dots, p_k\}$, e $Q = \{q_1, \dots, q_k\}$, duas distribuições de probabilidades de Ω , a distância de Hellinger está definida para o caso discreto como:

$$dHel(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^k (\sqrt{p_i} - \sqrt{q_i})^2} \quad (6)$$

A distância máxima 1 é alcançada quando P atribui a probabilidade zero a cada conjunto ao qual Q atribui uma probabilidade positiva e vice-versa. Quando o valor for próximo de 0 maior será a probabilidade. A vantagem de utilização dessa métrica é a utilização de valores percentuais para estimação da similaridade entre as matrizes. Visto que um dos objetivos do trabalho é identificar a similaridade entre a matriz manual e a usando sinal Wi-Fi, tal estimativa só será possível por meios percentuais, dessa forma a distância de Hellinger auxiliará nessa análise.

2.3.4 DEFASAGEM ENTRE AS MATRIZES

Além dos métodos citados anteriormente também pode ser utilizado o método usando a defasagem entre as matrizes reais e a obtida usando Wi-Fi. Tal método visa a utilização do somatório da diferença entre os pontos de embarque para a matriz real e matriz usando Wi-Fi dividido pela quantidade de smartphones utilizados na pesquisa. Tal somatório também é válido para o desembarque. A Equação 7 apresenta numericamente a equação descrita acima.

$$\sum_{i=1}^n \left| \frac{Si - Oi/D_i}{N_{\text{smartphone}}} \right| \quad (7)$$

Onde:

Si é o ponto de parada de embarque/desembarque onde o *smartphone* i foi identificado pelo Wi-Fi,

O_i/D_i é o ponto de parada real onde o *smartphone* i realmente embarcou ou desembarcou

Observe que quanto menor o resultado mais similar serão as matrizes.

2.4 TRABALHOS RELACIONADOS

Recentemente, vários pesquisadores investigam a aplicabilidade das tecnologias de comunicação sem fio, como o Wi-Fi e Bluetooth, para melhor compreender a movimentação dos passageiros de um sistema de transporte público. Por exemplo, a tecnologia Wi-Fi vem ganhando destaque devido à quantidade de dados gerados, sendo assim superior à outras tecnologias como, por exemplo, o Bluetooth (EL-TAWAB *et al.*, 2017). Esse fluxo de dados permite não somente identificar a presença do passageiro, mas acompanhar o trajeto do mesmo ao longo do sistema de transporte.

De acordo com Dunlap *et al.* (2016), os leitores de Wi-Fi instalados dentro dos ônibus possibilitam a estimação das origens e destinos dos passageiros de ônibus. As matrizes O/D foram estimadas usando sensores de Bluetooth e Wi-Fi. Observou-se que as matrizes obtidas pela tecnologia Wi-Fi foram superiores em termos quantitativos as obtidas por Bluetooth. Entretanto, observa-se uma ausência em relação as limitações inerentes a tecnologia Wi-Fi, como o atraso do envio de pacotes e o impacto na captura dos pares de origem e destino dos passageiros.

A Tabela 2.1 apresenta os principais trabalhos encontrados sobre a estimação da matriz O/D e suas tecnologias para tal estimação. O trabalho de Hidayat *et al.* (2018) usa a tecnologia Wi-Fi de para identificação dos não passageiros do transporte público, o trabalho tem destaque devido a técnicas utilizadas pelos autores para estimação de quem não está embarcado no veículo, uma vez que neste trabalho será necessário fazer uma separação de quem está no veículo e quem está fora.

Dos trabalhos apresentados na Tabela 2.1 observou-se que a ausência de uma comparação entre a matriz obtida pela tecnologia com a real. Dentre os trabalhos, alguns realizaram uma comparação entre diferentes tecnologias, como, por exemplo, o Mishalani *et al.* (2016), Dunlap *et al.* (2016), Ji *et al.* (2017), Hakegard *et al.* (2018), que compararam a tecnologias Wi-Fi com Bluetooth e APC.

Uma das principais limitações apresentada nos trabalhos foram a de realização das pesquisas, como, por exemplo, o trabalho de El-Tawab *et al.* (2017) devido ao tempo de emissão de pacotes de Bluetooth serem menores que Wi-Fi houve sobreposição de dados. Já no Trabalho de Hakegard *et al.* (2018) ocorreu erros na hora de execução do

experimento. No trabalho de Hidayat *et al.* (2018) ocorreu falsas leituras de Wi-Fi que impossibilitou a realização de estudos mais aprofundados.

Como dito, os resultados dos trabalhos apresentam alguns aspectos importantes para a pesquisa como citado acima, dentre eles se destaca a preocupação com a taxa de envio, visto que os trabalhos não apresentam questionamentos sobre a mesma. Dentre as diferentes metodologias aplicadas, em todas foram possíveis estimar a matriz O/D, embora não comparada com a matriz real, mas comparadas entre si pode-se observar que apresentaram uma similaridade significativa entre as matrizes.

Tabela 2.1: Textos Relacionados

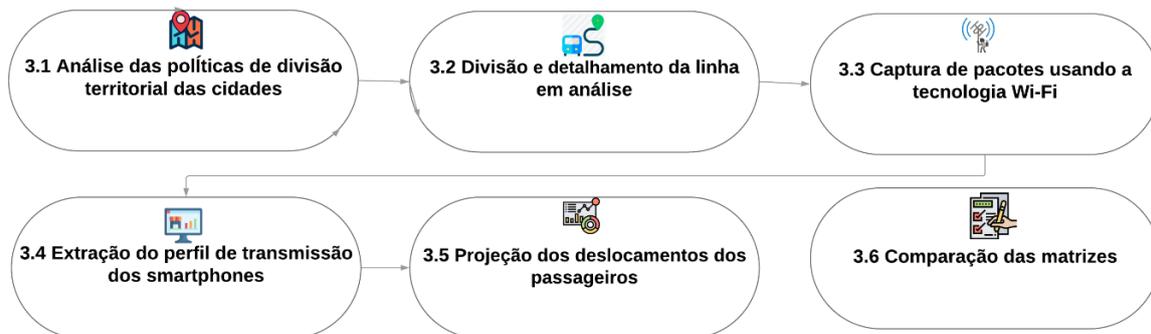
Autores	Tecnologia	Resolução da Matriz (Granularidade)	Comparação	Considera atraso de envio	Premissas	Técnica
Mishalani et al. 2016	Wi-Fi e APC	Ponto a ponto	Entre Tecnologias	Não	Combinação das tecnologias para geração da matriz O/D	Algoritmos Próprio
Dunlap et al. 2016	Wi-Fi, Bluetooth e GPS	Ponto a ponto	Entre Tecnologias	Não	Sobreposição de viagem	Algoritmos Próprio
Ji et al. 2017	Wi-Fi e Farebox	Ponto a ponto	Entre Tecnologias	Não	É possível estimar a matriz O/D através de dados como velocidade média do veículo	Modelo Bayesiano
El-Tawab et al. 2017	Wi-Fi Bluetooth	Ponto a ponto	Não realizou comparação	Não	Instalou em cada ponto de parada um aparelho de coleta de dados	Nó inteligente
Hakegard et al. 2018	Wi-Fi e APC	Ponto a ponto	Entre Tecnologias	Não	Combinação das tecnologias para geração da matriz O/D	Modelagem estatística
Afshari et al. 2019	Wi-Fi	Ponto a ponto	Entre Tecnologias	Não	Ao utilizar os clusters para estimação da matriz O/D é possível aumentar a precisão dos resultados	Clusters
El-Tawab et al. 2019	Wi-Fi	Ponto a ponto	Não realizou comparação	Não	Instalou em cada ponto de parada um aparelho de coleta de dados	Nó inteligente
Zhou et al. 2019	Wi-Fi	Ponto a ponto	Não realizou comparação	Não	É possível estimar a matriz O/D através de dados como velocidade média do veículo	SVM
Jalali, 2019	Wi-Fi e Bluetooth	Ponto a ponto	Manual	Não	É possível criar uma matriz O/D real online	Clusters

3 METODOLOGIA

Considerando os limites da tecnologia Wi-Fi, este trabalho apresenta uma metodologia para estimar a viabilidade de se obter a matriz O/D do sistema de transporte público coletivo, utilizando dados oriundos dessa tecnologia, através da flexibilização dos pontos de paradas em agrupamentos com base em divisões maiores (*e.g.*, macrozonas e zonas de tráfego) e da análise da frequência dos dados recebidos.

As etapas da metodologia consistem em analisar as políticas de divisão territorial das cidades, para delimitar as formas de agrupar (subdivisões) os pontos de parada de origem e destino, seguida pela intersecção com a linha de ônibus em análise. Após essas etapas, realiza-se a captura, de forma manual, dos sinais Wi-Fi. Posteriormente, realiza-se o processamento da base de dados para extrair as características de transmissão, como a quantidade e a frequência de envio de pacotes. Em posse destes dados, realiza-se a projeção dos deslocamentos na linha de transporte para obtenção da matriz. Por fim, é feita uma comparação entre a matriz sintética da obtida manualmente. Os passos desta metodologia estão ilustrados na Figura 3.1.

Figura 3.1: Etapas da metodologia.



Fonte: Autor (2020).

3.1 ANÁLISE DAS POLÍTICAS DE DIVISÃO TERRITORIAL DAS CIDADES

A primeira parte da metodologia consiste em analisar a política de divisão espacial da cidade para delimitar as formas de agrupamentos de pontos de parada possíveis, *e.g.*, por zonas de tráfego ou por bairros. Como citado anteriormente, esta informação é comumente relatada em normativas pública, como nos planos diretores municipais e setores censitários. É preferível a escolha de delimitações digitalizadas, uma vez que será necessário realizar o processamento delas computacionalmente. A etapa contempla a escolha de uma ou mais resolução de divisão territorial, uma vez tem como objetivo investigar quais estruturas possuem maior similaridade com o deslocamento observado pelos passageiros.

Para exemplificar esse passo, considere o caso de Aparecida de Goiânia, que teve seu primeiro Plano Diretor realizado em 2002 e recentemente foi revisado (2014/ 2015). Este traz consigo as delimitações das centralidades e regiões administrativas, além das zonas de tráfego do município. Também serão consideradas as divisões do espaço por setores censitários.

Os arquivos de centralidades e regiões administrativas conjuntamente com as zonas de tráfego serão reproduzidos em conformidades com as do Plano Diretor de Aparecida de Goiânia. Já os setores censitários serão adquiridos no banco de dados do IBGE. Todos esses arquivos estarão em um formato vetorial para que possa servir de entrada em sistemas de informação geográfica (como ArcMap e QGis) e programas de computação estatística (como a plataforma R e o MATLAB).

3.2 DIVISÃO E DETALHAMENTO DA LINHA EM ANÁLISE

Após a determinação de quais divisões territoriais serão utilizadas na pesquisa, é necessário um estudo de quais linhas de ônibus serão analisadas. Semelhantemente as divisões territoriais, é desejável que a linha de ônibus esteja descrita em formato digital, preferencialmente vetorial, como, por exemplo, em *Keyhole Markup Language* (KML) ou *shapefile*, pois estes formatos facilitam o processamento computacional realizado adiante.

Em posse das delimitações das divisões territoriais e das linhas em análise em formato vetorial, pode-se prosseguir para o processamento dos dados preliminares. Nesta etapa é necessário a realização da intersecção dos arquivos de divisão territorial com as linhas escolhidas. Recomenda-se realizar o processo de intersecção em um programa de sistema de informação geográfica, como QGis e ArcGis, para obter o resultado da intersecção em formato vetorial (como *shapefile*).

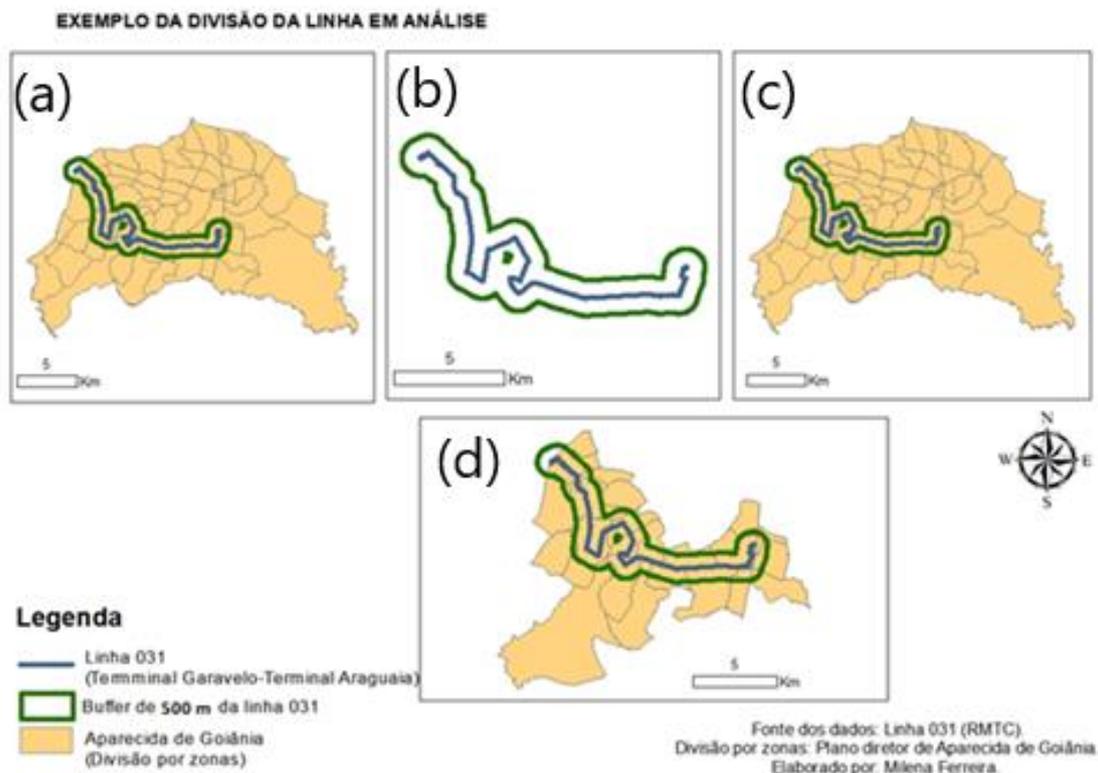
Inicialmente para realização da intersecção é necessário realizar a criação de um *buffer* na linha de transporte, que é um processo que expande (daí o nome *buffer*) o polígono em uma distância específica ao redor das feições selecionadas. A partir disso é possível obter as áreas que são influenciadas pelo polígono selecionado. Essa distância pode ser determinada de acordo com o estudo. Os autores Ragazzo e Lima (2013) afirmam que a área de influência de uma linha de transporte público é de 500 metros, por ser a distância aceitável para o deslocamento do passageiro de sua residência até o ponto de parada. Tal valor pode ser adotado para criação do *buffer*.

Ao adicionar a divisões da cidade e criar o *buffer* da linha pode-se realizar a intersecção dos dois arquivos vetoriais. Nessa etapa a ferramenta realizará uma separação de quais áreas são influenciadas pela linha, ou seja, quais áreas são interceptadas pela linha escolhida,

considerando o raio de influência. O resultado dessa intersecção será utilizado como entrada para as outras fases do método. Por exemplo, servirá para contabilizar a matriz manual e de entrada para a construção da matriz O/D dos dados Wi-Fi.

A Figura 3.2 exemplifica como funciona o processo de intersecção e como é o resultado. O exemplo é da linha 031 da rede de transporte coletivo de Aparecida de Goiânia. A primeira parte da figura (a) mostra a cidade de Aparecida de Goiânia dividida em zonas de tráfego, logo em seguida é apresentada a linha com um *buffer* (b). A fim de ilustrar o processo, a próxima ilustração (c) mostra a sobreposição da linha com o *buffer* na subdivisão do município. Por fim a última ilustração (d) demonstra o resultado da intersecção da linha com as áreas representadas, a saber apenas as zonas que são tocadas pelo *buffer*.

Figura 3.2: Exemplo de intersecção das zonas com a linha de ônibus.



Fonte: Autor (2020).

Com base nesse processo, pode-se então realizar um detalhamento das características espaciais de cada parte da linha. A saber, será contabilizado o número de pontos de paradas distribuído em cada divisão territorial. Tal medida está diretamente relacionada a flexibilização da matriz, uma vez que divisões maiores conterão um quantitativo maior de pontos de parada. Também é extraído o espaçamento médio entre os pontos de parada da linha. Tal informação será útil, pois a mesma pode ser correlacionada com a frequência de recebimento dos dados dos *smartphones* dos passageiros (como descrito a seguir).

3.3 CAPTURA DOS DADOS WI-FI

Em posse das análises feitas na linha de estudo, inicia-se o processo para a construção da matriz utilizando os sinais Wi-Fi emitidos pelos aparelhos. Para tal, primeiramente, é necessário fazer a captura de dados, isto é, um rastreamento dos pacotes Wi-Fi recebidos dentro do ônibus. Para o processo de captura de dados será necessário o uso de um computador intermediário (*sniffer*), que ficará no interior do veículo e será responsável por receber todos os pacotes enviados pelos *smartphone*.

Os *sniffers* são sistemas embarcados portáteis que podem ser introduzidos dentro dos ônibus. Especificamente, neste trabalho utilizou-se a plataforma de *hardware* Raspberry Pi 3B para construção deste equipamento, devido o mesmo possuir por padrão um *chip* de monitoramento Wi-Fi e a dimensão reduzida do sistema. O dispositivo, possui uma dimensão de 85,6 x 53,98 x 17 mm como é ilustrado na Figura 3.3. Através deste *chip*, o equipamento poderá monitorar e capturar os dados Wi-Fi emitidos por aparelhos próximos.

É importante destacar que deverá ser acoplado ao *sniffer* um sensor de GPS, que servirá para atrelar a posição do veículo ao recebimento do pacote Wi-Fi. Tal procedimento permite construir um rastro de dados enviados por um determinado aparelho, que conseqüentemente corresponde ao trajeto de um passageiro da respectiva linha.

Desta forma, ao iniciar o processo de coleta de dados do *sniffer* deverá ser adquirido uma base de dados com as seguintes características para cada dado de entrada:

- Endereço MAC (Média Access Control), que é único para cada dispositivo, semelhante a impressão digital do aparelho;
- Indicador de Força do Sinal Recebido (RSSI) que é um valor numérico, que indica a potência do sinal Wi-Fi entre o *smartphone*, medido em decibéis por miliwatt (dBm);
- Latitude e Longitude do *sniffer* ao capturar um pacote de sonda;
- Velocidade do veículo, em m/s, que indica o quão rápido o ônibus se encontra quando o pacote é capturado;
- Horário do *sniffer* que mostra o instante (*timestamp*) em que o pacote foi coletado;

Como forma de controle, nessa pesquisa serão realizados o tratamento apenas dos dados dos *smartphones* conhecidos previamente (endereço MAC). Precisamente, após a coleta de dados os pacotes enviados por endereços MAC desconhecidos serão retirados da análise, restando apenas os dados conhecidos.

Figura 3.3: Equipamentos.



Fonte: Autor (2020).

3.4 PERFIL DE TRANSMISSÃO DOS SMARTPHONES

Nessa etapa será realizado o processamento dos dados adquiridos usando a tecnologia Wi-Fi a fim de se obter o perfil de transmissão dos *smartphones* e, conseqüentemente, do deslocamento dos passageiros. Como citado anteriormente vários fatores influenciam na taxa de transmissão dos pacotes de sondagem, como por exemplo, a marca, o modelo e a utilização ou não do aparelho na hora que está coletando os dados (FREUDIGER, 2015). Especificamente, a forma de uso do aparelho impacta no quantitativo de dados gerados. Para essa pesquisa será utilizado o pior caso de uso, onde o *smartphone* encontra-se em repouso e com a tela desligada (*e.g.*, no bolso do passageiro). Sabe-se que esta situação de uso é aquela que produz o menor número de pacotes de sondagem, assim teremos menos dados capturados e conseqüentemente, menos informações do deslocamento do passageiro.

Para obtenção do perfil do *smartphone* e, conseqüentemente, do deslocamento, será feito uma análise de comparação entre os tempos que os pacotes foram recebidos. A fim de ilustrar tal análise, considere a Tabela 3.1 que exemplifica um trecho do recebimento de pacotes de um *smartphone* com endereço MAC C8:F3:19:60:48:77. Pode-se observar que o dispositivo apresentou uma frequência de transmissão de dados irregulares. Por exemplo, os dois primeiros pacotes possuem uma diferença de 40 segundos, já o terceiro e quarto possui um espaçamento de 191 segundos. A partir desses dados é feito uma extração da média de envio do dispositivo do passageiro. Tal frequência será útil para compreender como a mesma poderá impactar o atraso entre o embarque real do passageiro e a captura do dado conseqüente.

Tabela 3.1: Trecho do perfil de transmissão do *smartphone* C8:F3:19:60:48:77.

Timestamp do pacote	RSSI (dBm)	Latitude	Longitude	Velocidade do ônibus (m/s)
00:00:16	-56	-16,668009896	-49,238210815	17,6
00:00:56	-55	-16,668018289	-49,238159895	0
00:02:16	-50	-16,668023095	-49,238194414	0
00:05:27	-49	-16,668049474	-49,238289892	37,5

Fonte: Autor (2020).

3.5 PROJEÇÃO DO DESLOCAMENTO DOS PASSAGEIROS COM BASE NA TECNOLOGIA WI-FI

Com os valores de frequência de transmissão de cada aparelho é possível ainda realizar alguns cálculos que auxiliarão na projeção do atraso dos pontos de paradas para cada *smartphone*. A projeção desse atraso será feita com base em três variáveis: a frequência de envio dos pacotes, a velocidade operacional da linha em estudo e o espaçamento médio entre os pontos. A velocidade operacional pode ser obtida com base em dados históricos da linha como também em velocidades teóricas. Por exemplo, Ferraz e Torres (2004) afirmam que a velocidade operacional teórica para ônibus deve variar de 10 a 40 km/h. Já o espaçamento médio pode ser obtido a partir de uma análise espacial da linha com base na espacialização digital, como mencionando na seção 3.2.

Em posse desses dados, pode-se projetar a distância média percorrida pelo ônibus em função do tempo (frequência conjuntamente com tempo de embarque e desembarque) e velocidade de operação. Para tal, utiliza-se a fórmula elementar de velocidade média, ilustrada na Equação 8:

$$V_{op} = \frac{d}{t} \quad (8)$$

Onde: V_{op} é a velocidade média de operação em m/s da linha em estudo,

d é a projeção do deslocamento em metros,

t é a soma de dois elementos, o primeiro é a frequência de transmissão de cada *smartphone* e o segundo é o tempo que o passageiro leva para embarcar e desembarcar no transporte público. Ferronato (2005) afirma que esse valor é de 15 segundos para pontos de paradas comum. Quando não se leva em consideração esse valor no cálculo assume-se o pior caso, ou seja, que o veículo não irá parar em nenhum ponto de parada. Para a tecnologia Wi-Fi

esse seria o pior caso, uma vez que terá menos tempo para transmissão dos dados. Assim, neste trabalho desconsidera-se o tempo gasto no embarque e desembarque em cada ponto de parada.

Esse cálculo simples é importante pois a partir desse valor é possível projetar em média, qual será a posição do trajeto da linha que o primeiro pacote de sondagem do *smartphone* será recebido pelo *sniffer*. Esse dado, em conjunto com o espaçamento médio entre os pontos de parada, pode então informar o atraso médio entre o ponto de embarque e o identificado.

Para exemplificar essa etapa, considere que o dispositivo de um passageiro envia em média um pacote de sondagem a cada 200 segundos. Considerando uma velocidade média de operação de $25 \text{ km/h} = 6,944 \text{ m/s}$, pode-se então estimar que em tal cenário o pacote percorreria $d \cong 6,944 \cdot 200 = 1388$ metros. Ao alinharmos a projeção da distância percorrida com um espaçamento de 400 metros entre os pontos de parada, os dados indicariam que a informação de origem daquele passageiro seria capturada, no mínimo, após quatro estações do local de embarque.

Essa defasagem (ou atraso) entre o local de embarque e o local estimado será objeto de estudo da próxima etapa do método, onde será feita uma comparação entre o local de embarque real e manual. Observe que o agrupamento (*e.g.*, zona de tráfego ou setor censitário) de vários pontos de parada em uma mesma origem ou destino atenua os erros, uma vez que em tal cenário, não haveria “erro” caso o ponto de embarque ou desembarque esteja no mesmo agrupamento onde o dado foi capturado.

3.6 COMPARAÇÃO DAS MATRIZES

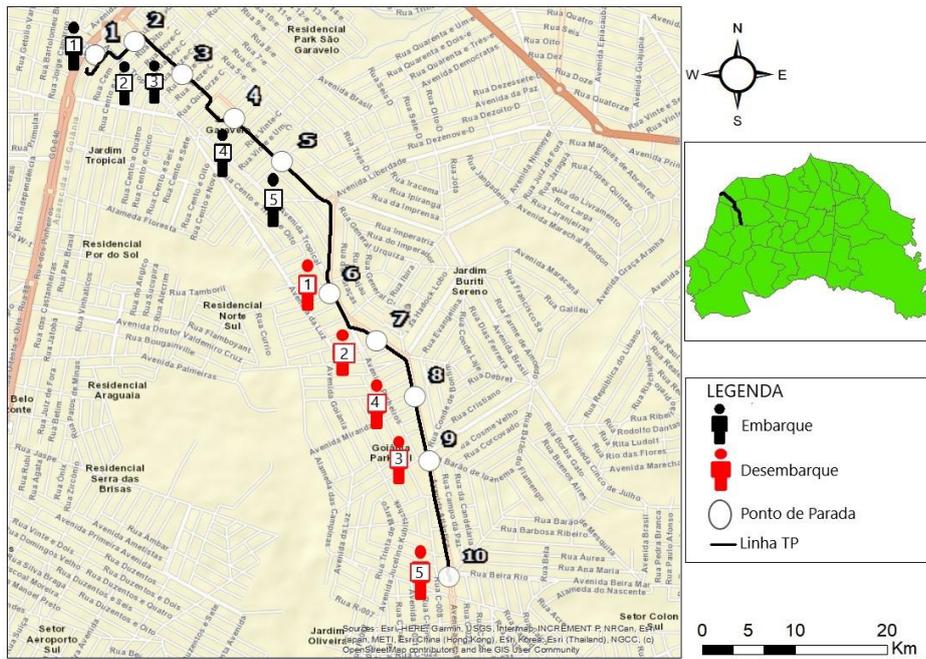
Após a obtenção dos dados das seções anteriores é necessário realizar uma comparação entre as matrizes construídas, para isso será levado em consideração a defasagem entre o embarque/ desembarque real e o obtido usando a tecnologia Wi-Fi. Essa etapa é importante para validação do método adotado na pesquisa, ou seja, a comparação indicará se as matrizes apresentam similaridade entre si e entre a matriz O/D real.

Para essa etapa será utilizado a comparação das matrizes entre o local de embarque e desembarque real e a defasagem em relação a obtida, ou seja, como temos a matriz O/D real que será a mesma para todos *smartphones* podemos realizar um somatório entre a origem e destino real e a obtida com dados Wi-Fi e dividir pelo número de *smartphone* como descrito na Equação 7 para a origem e destino dos passageiros. Tal metodologia pode ser aplicada tanto para pontos de parada como para áreas maiores.

Para exemplificar usaremos uma linha de transporte público como mostra a Figura 3.4, com dez pontos de paradas e cinco passageiros. Com a Figura 3.4 temos a matriz O/D real,

agora vamos estimar hipoteticamente a origem da matriz usando os dados Wi-Fi como mostra a Tabela 3.2, o resultado será apresentado logo abaixo.

Figura 3.4: Exemplo de comparação de matriz.



Fonte: Autor (2020).

Tabela 3.2: Exemplo da origem da matriz O/D.

Passageiro/Matriz	1	2	3	4	5
Ponto Real	1	3	3	4	5
Ponto usando tecnologia Wi-Fi	2	5	4	5	5

Fonte: Autor (2020).

Usando a Equação 7 e os dados da Tabela 3.2 podemos encontrar que o valor da similaridade será:

$$\sum_{i=1}^n \left| \frac{S_i - O_i}{N_{\text{smartphone}}} \right| = \frac{|S1 - O1| + |S2 - O2| + |S3 - O3| + |S4 - O4| + |S5 - O5|}{5}$$

$$\frac{1 + 2 + 1 + 1 + 0}{5} = \frac{4}{5} = 0,8.$$

Então a similaridade entre a origem da matriz O/D real e usando dados Wi-Fi é de 0,8. Note que o valor esperado para a diferença entre as matrizes O/D é zero, ou seja as matrizes seriam idênticas entre si. Dessa forma à medida que esse valor aumenta, também se aumenta a discrepância média entre os pontos de paradas observados. Nesse caso significa que em média a matriz obtida possui uma discrepância de 0,8 pontos de paradas entre os locais de embarque.

Para o destino também será utilizado a Equação 7, assim como na origem será feito uma estimação dos pares do destino usando a tecnologia Wi-Fi, como mostra a Tabela 3.3.

Tabela 3.3: Exemplo do destino da matriz O/D.

Passageiro/Matriz	1	2	3	4	5
Ponto Real	6	7	9	8	10
Ponto usando tecnologia Wi-Fi	4	3	5	6	7

Fonte: Autor (2020).

Com os valores acima podemos obter o valor da similaridade entre os destinos.

$$\sum_{i=1}^n \left| \frac{S_i - D_i}{N_{\text{smartphone}}} \right| = \frac{|S1 - D1| + |S2 - D2| + |S3 - D3| + |S4 - D4| + |S5 - D5|}{5}$$

$$\frac{2 + 2 + 2 + 0 + 0}{5} = \frac{6}{5} = 1,2.$$

Logo à similaridade entre o destino da matriz O/D real e usando dados Wi-Fi é de 1,2. O que significa que a discrepância entre as matrizes é de 1,2 pontos de parada entre os locais de desembarque.

4 EXPERIMENTO

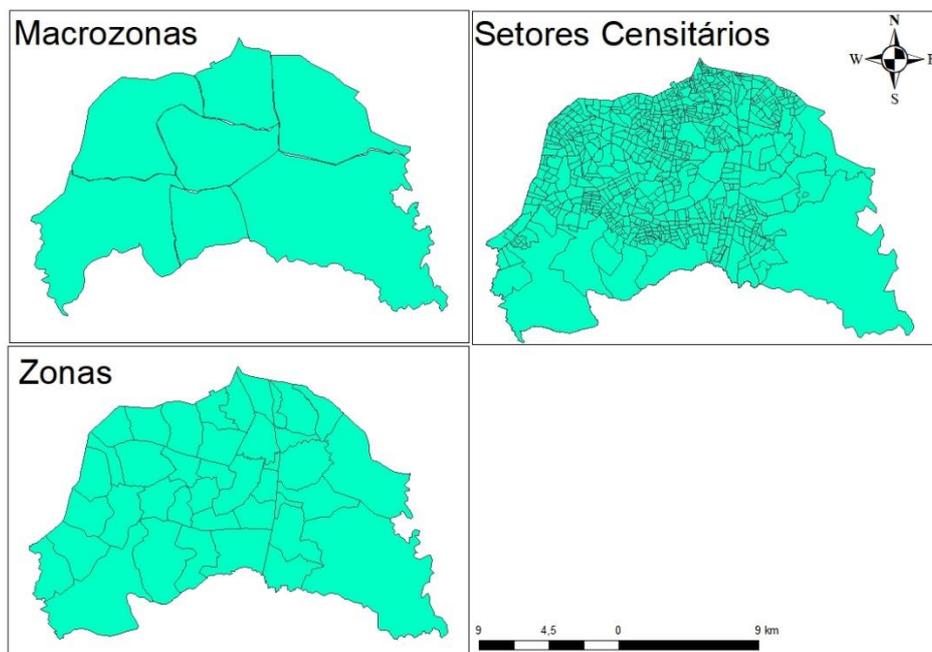
Neste capítulo será apresentado um estudo de caso que ilustra a aplicação do método em uma linha de ônibus do município de Aparecida de Goiânia. Com esse estudo, objetiva-se explorar as matrizes O/D geradas a partir da tecnologia Wi-Fi, em especial compreender como a granularidade das divisões territoriais irá impactar na acurácia da matriz obtida.

4.1 ANÁLISE DAS POLÍTICAS DE DIVISÃO TERRITORIAL

Como abordado na metodologia, esta etapa consiste em analisar as políticas de divisão espacial da cidade para definir as possíveis granularidades da matriz O/D. Escolheu-se a cidade de Aparecida de Goiânia como objeto de estudo devido a mesma possuir várias divisões territoriais predefinidas digitalmente, tais como macrozonas e setores censitários, bem como a mesma ser acessível aos pesquisadores, uma vez que a instalação do curso de Engenharia de Transportes da UFG se encontra neste município. Para o estudo chegou-se à conclusão de que as macrozonas, as zonas de tráfego e os setores censitários seriam os mais adequados, visto que as outras divisões territoriais disponíveis, por exemplo os bairros, não apresentam compatibilidade com as descritas na literatura nem possuem versões digitalizadas disponíveis.

As subdivisões para macrozonas e zonas de tráfego são descritas no plano diretor do município, enquanto que os setores censitários foram obtidos através da plataforma do IBGE. A Figura 4.1 ilustra as respectivas subdivisões espaciais do município.

Figura 4.1: Divisões espaciais da cidade de Aparecida de Goiânia.

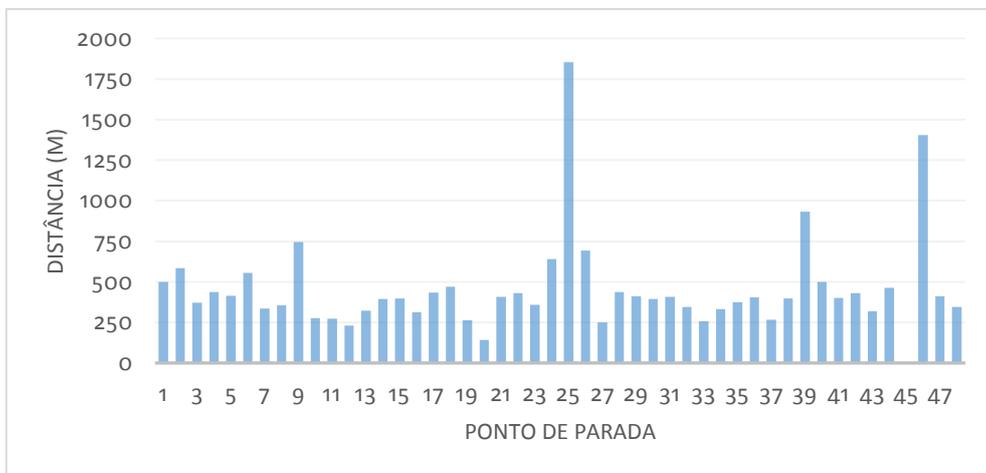


Fonte: Autor (2020).

Como mencionado no Capítulo 3, para determinação das subdivisões influenciadas pela linha 975, houve a necessidade da criação de um *buffer*. A área de influência escolhida foi de 500 metros devido a dois fatores, a distância média entre os pontos de parada e valores referenciais da literatura. Utilizando ferramentas de processamento de informação geográfica, observou-se que os pontos de parada da linha apresentam uma distância média de 458 m. Como segundo fator, ressalta-se que a literatura adota 500 m como a distância de caminhada até o ponto de parada (RAGAZZO e LIMA, 2013).

O espaçamento médio entre os pontos de parada da linha foi calculado com base na média da distância entre duas estações consecutivas. Para cada par dos 48 pontos de parada, calculou-se a distância entre eles, ilustrado na Figura 4.3. Com base nos dados do distanciamento entre os pontos de parada foi possível calcular a distância mínima, média e máxima entre tais pares, cujos valores são 142 m, 458 m e 1854 m.

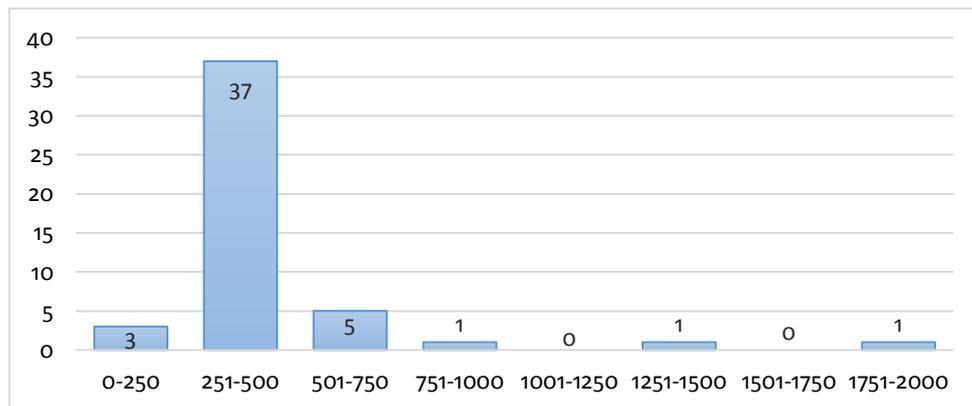
Figura 4.3: Distância entre os pontos de parada da linha 975.



Fonte: Autor (2020).

A Figura 4.4 expressa a frequência de distribuição para cada intervalo de espaçamento entre os pontos de parada.

Figura 4.4: Histograma da distância entre pontos de parada da linha 975.

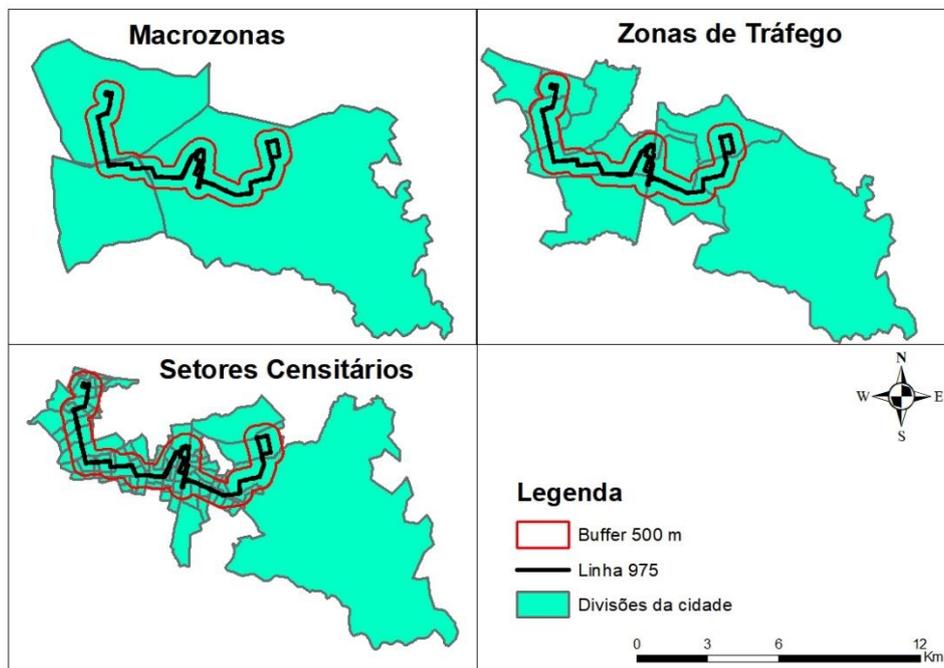


Fonte: Autor (2020).

Na Figura 4.4 é possível observar a predominância do intervalo de 251 a 500 m na distância entre os pontos de paradas (37 dos 48 pares). A saber, apenas três pontos de paradas tem uma distância de até 250 m entre si, enquanto que cinco pontos situam no intervalo de 501 a 750 m, um ponto de parada entre 751 a 1000 m e um entre 1750 a 2000 m. O gráfico também mostra a presença de dois *outliers*. Um detalhamento revelou que a discrepância entre tais pontos de paradas é situada na margem da BR153. O acesso aos pontos de parada na região fica limitado, por ser uma via de tráfego rápido.

Após a criação do *buffer* foram feitas as interseções da linha com as feições (representações geográficas) das divisões territoriais escolhidas, como ilustrado pela Figura 4.5. Como resultado, obtêm-se as subdivisões que são influenciadas diretamente pela linha 975. Este conjunto de subdivisões expressa a base para a construção dos pares de origem e destino dos passageiros nas respectivas divisões.

Figura 4.5: Raio de influência da linha 975 para as diferentes divisões territoriais na cidade de Aparecida de Goiânia.



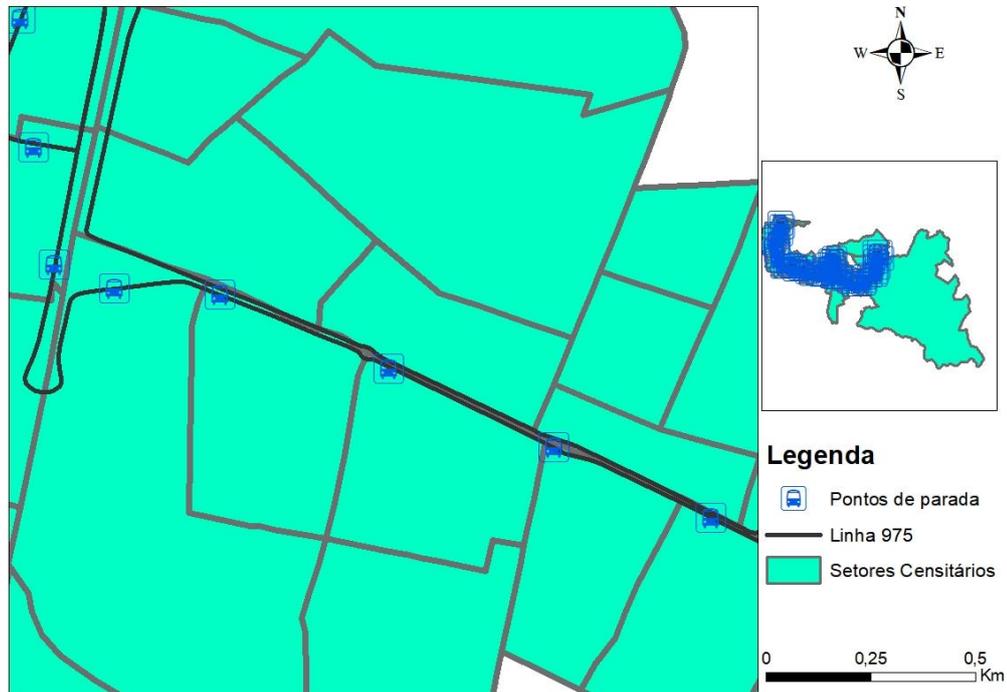
Fonte: Autor (2020).

Durante a execução desta etapa observou-se um fenômeno peculiar, uma parte significativa dos pontos de parada situou na borda das divisões. Tal situação é relevante, uma vez que se sabe que o sensor GPS de *smartphone* possui uma imprecisão, comumente entre 10 e 30 metros (BAUER, 2013). Desta forma, pode ocorrer da localização de um pacote Wi-Fi capturado pelo *sniffer* ser atrelado a uma divisão diferente daquela onde ele realmente ocorreu.

A Figura 4.6 ilustra esta situação. Observe que quando o ponto de parada fica na borda da divisão torna-se complicado a estimação da matriz O/D, visto que o embarque/desembarque

de um passageiro nesse ponto de parada na borda dificultará a identificação do local exato da ocorrência do evento. Tal incongruência dificulta a associação da origem ou destino do passageiro na respectiva área.

Figura 4.6: Exemplo dos pontos de parada na borda dos setores censitários.



Fonte: Autor (2020).

Na Tabela 4.1 são apresentados os valores da quantidade dos pontos de parada em cada uma das três macrozonas e quantos desses pontos estão na borda das macrozonas. Pode-se observar que este tipo de divisão não possui nenhum ponto de parada na borda, o que pode ser explicado pelo tamanho das macrozonas que são suficientemente espaçadas entre si.

Tabela 4.1: Número de pontos de parada para macrozonas.

Macrozonas	Número de pontos de parada	Na borda
1	9	0
2	6	0
3	33	0

Fonte: Autor (2020).

A Tabela 4.2 apresenta a relação de pontos de paradas com a divisão territorial em zonas de tráfego. Nesse caso já é possível identificar que alguns pontos de parada se situam na borda das zonas de tráfego. Além disso, observa-se que 18,75% (9 dos 48) do total de pontos de parada se situam na borda. Além da granularidade da área ser menor, uma possível explicação para tal fenômeno decorre de que as linhas de transportes percorrem as principais avenidas da cidade, que em muitos casos essas avenidas são usadas como limite para divisão das zonas de tráfego.

Tabela 4.2: Número de pontos de parada para zonas de tráfego.

Zonas de tráfego	Número de pontos de parada	Na borda
1	2	0
2	6	0
3	9	2
4	12	4
5	1	1
6	7	2
7	6	0
8	5	0

Fonte: Autor (2020).

Já na Tabela 4.3 são apresentados os valores para os vinte e oito setores censitários. Pode-se observar que nessa divisão mais pontos de parada ficaram na borda o que indica que esse tipo de divisão é mais granular que as demais o que pode dificultar as análises posteriores. Observar-se que 56,25% (27 dos 48) dos pontos de parada ficam na borda.

Tabela 4.3: Número de pontos de parada para setores censitários.

Setores Censitários	Número de ponto de parada	Na borda
1	2	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	2	2
8	1	0
9	1	0
10	2	0
11	1	1
12	2	1
13	1	1
14	1	1
15	2	2
16	2	1
17	2	1
18	1	1
19	3	0
20	2	2
21	1	0
22	1	1
23	1	1
24	1	1
25	1	1
26	2	2
27	6	0
28	5	2

Fonte: Autor (2020).

4.3 CAPTURA DAS CARACTERÍSTICAS DOS DADOS WI-FI

Em posse das divisões, a próxima etapa consiste em coletar os dados Wi-Fi emitidos pelos *smartphones* dos passageiros. Inicialmente, previa-se a execução desta etapa dentro de um ônibus da linha. Entretanto, por conta da pandemia do COVID-19, não foi possível realizar os experimentos em campo devido ao alto índice de propagação da doença (HU *et al.*, 2020; SOHRABI *et al.*, 2020).

Apesar da pandemia, esta etapa pode ser realizada sem tamanho prejuízo de forma local, uma vez que ela visa extrair o perfil de transmissão dos dados Wi-Fi, como a quantidade de pacotes e frequência. Desta maneira, optou-se por configurar um cenário onde os dispositivos

ficaram parados e com a tela desligada em um raio de 10 metros do *sniffer* durante um período de duas horas. A saber, foram utilizados os dispositivos apresentados na Tabela 4.4.

Neste experimento, cada *smartphone* representa um respectivo passageiro. O valor do raio é escolhido com base no comprimento dos ônibus, uma vez que os ônibus convencionais apresentam diâmetro inferior a 14 metros (DIÁRIO DO TRANSPORTE, 2014). Desta forma, considerando o *sniffer* no centro do veículo, o raio de 10 metros é suficiente para cobrir sua extensão.

Destaca-se que o período de duas horas não tem como intuito simular o tempo de viagem do passageiro, mas sim extrair as características de transmissão da tecnologia Wi-Fi dos dispositivos. Ao deixar o *smartphone* em repouso e com a tela desligada, pretende-se aproximar do cenário em que o aparelho se encontra no bolso do passageiro e sem receber mensagens. Sabe-se que este cenário é o pior caso de uso da tecnologia, isto é, é aquele que apresenta a menor taxa de transmissão de pacotes, uma vez que o sistema operacional entende que não é necessário buscar frequentemente redes Wi-Fi pois o dispositivo se encontra ocioso (FREUDIGER, 2015).

Tabela 4.4: *Smartphones* usados no trabalho.

Marca	Modelo	Versão Android	MAC Address
LG	K8+	9	A0:4F:85:D7:12:DC
LG	K8+	9	C8:F3:19:60:48:77
SAMSUNG	Galaxy Core 2 Duos (SM-6355M)	4.4.2	E4:58:E7:9D:CF:EF
SONY	Xperia L (XPERIA C2104)	4.2.2	68:76:4F:0A:60:DA
MULTILASER	M10A-LITE	7	40:45:DA:B6:82:BA
SAMSUNG	Galaxy Gran Neo Duos (GT-I9063T)	4.2.2	78:A8:73:95:37:A6

Fonte: Autor (2020).

Ao término do experimento obteve-se a frequência e quantitativo de pacotes enviados por cada dispositivo. A obtenção de tais dados foi feita utilizando o *software* Excel com base no histórico dos pacotes capturados. A Tabela 4.5 apresenta o tempo mínimo, médio e máximo para transmissão de pacotes para cada dispositivo.

Tabela 4.5: Dados sobre os *smartphones*.

Aparelho (MAC)	Mínimo (s)	Tempo médio (s)	Máximo (s)
A0:4F:85:D7:12:DC	20	630	1563
C8:F3:19:60:48:77	28	472	1512
E4:58:E7:9D:CF:EF	17	448	1657
68:76:4f:0a:60:da	15	330	600
40:45:DA:B6:82:BA	15	40	628
78:A8:73:95:37:A6	28	420	860

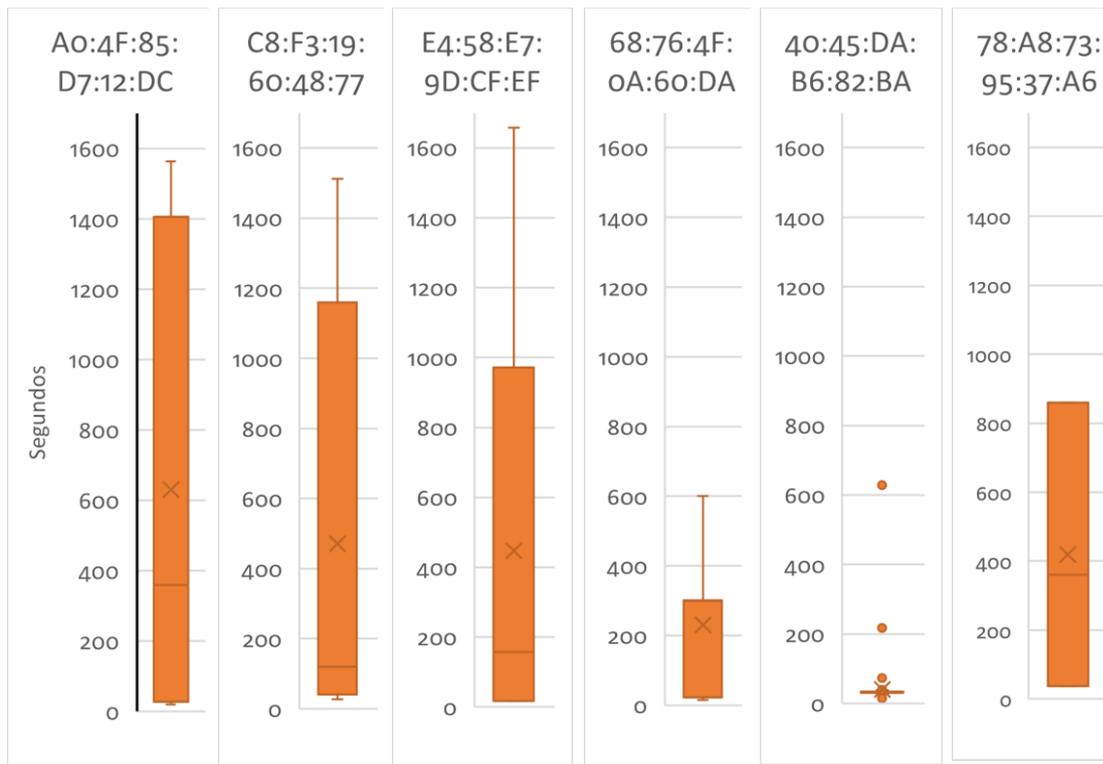
Fonte: Autor (2020).

Pode-se observar que os valores encontrados possuem uma diferença entre os diversos *smartphones*, o que confirma que a marca e o modelo influenciam diretamente nessa taxa de transmissão. Vale lembrar, que os dados encontrados representam o pior cenário da tecnologia, aquele em que o dispositivo se encontra em repouso e com a tela desligada.

O *smartphone* que possui a maior frequência média de transmissão é o identificado pelo MAC 40:45:DA:B6:82:BA, que é da marca Multilaser. Já o que possui a menor frequência média é o *smartphone* da LG representado pelo MAC A0:4F:85:D7:12:DC. A diferença entre esses valores reflete sobre a capacidade de rastrear o deslocamento do passageiro e, conseqüentemente, a própria geração da matriz O/D. Enquanto um *smartphone* leva 40 segundos para enviar um pacote, o outro demora 10 minutos e 30 segundos. Nesse último caso, tal frequência pode inviabilizar a detecção de viagens curtas, bem como aumentar a imprecisão a respeito dos locais de embarque e desembarque do passageiro.

Para exemplificar esse fenômeno, construiu-se o gráfico de caixas (*boxplot*) ilustrado na Figura 4.7. Nele, pode-se observar e comparar o comportamento de cada um dos *smartphones* do experimento. Observa-se uma discrepância na frequência de envio dos aparelhos.

Figura 4.7: Boxplot da frequência de transmissão dos *smartphones*.

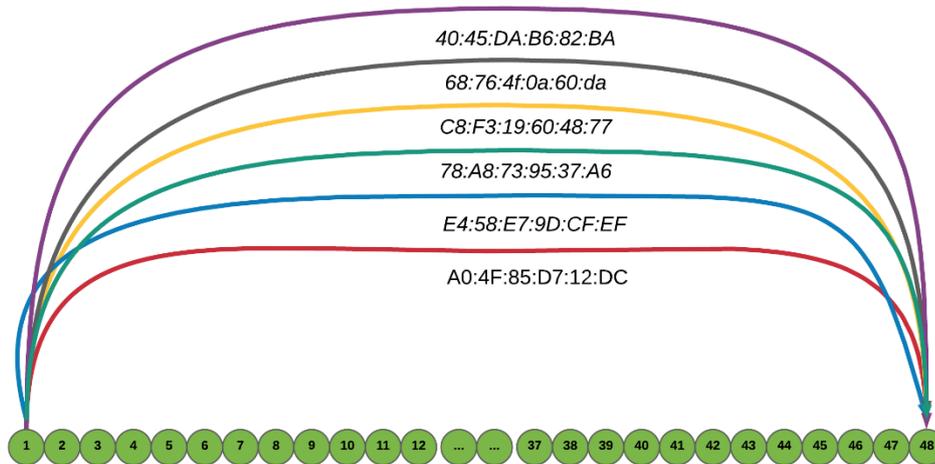


Fonte: Autor (2020).

4.4 MATRIZ O/D MANUAL

Após a obtenção dos dados anteriores inicia-se a etapa de construção da matriz O/D manual, isto é, aquela que servirá de gabarito para comparação. A saber, a mesma visa abordar as diferentes áreas das divisões escolhidas. Para compreender a capacidade de rastrear os passageiros, assume-se que todos os *smartphones* embarcam no primeiro ponto de parada e desembarcam no último, como ilustrado na Figura 4.8.

Figura 4.8: Matriz O/D manual ponto a ponto.



Fonte: Autor (2020).

Após a construção da matriz O/D manual ponto a ponto mostrado pela Figura 4.8 é possível estruturar as informações e montar a matriz O/D para cada divisão territorial. A Tabela 4.6 apresenta a matriz O/D manual para as macrozonas. As áreas das macrozonas foram numeradas de um a três a fim de facilitar a compreensão da tabela.

Tabela 4.6: Matriz O/D manual dos *smartphones* para as macrozonas.

ORIGEM/DESTINO	MC1	MC2	MC3	PRODUÇÃO
MC1	0	0	6	6
MC2	0	0	0	0
MC3	0	0	0	0
ATRAÇÃO	0	0	6	6

Fonte: Autor (2020).

A estruturação da matriz O/D para as macrozonas seguirão o mesmo que aconteceu para a matriz O/D ponto a ponto, a origem de todos os *smartphones* será na primeira macrozona e seus destinos serão na última macrozona.

A Tabela 4.7 apresenta como fica a distribuição dos dispositivos ao longo das zonas de tráfego. Assim como as macrozonas as zonas de tráfego foram numeradas de um a oito.

Tabela 4.7: Matriz O/D manual dos *smartphones* para as zonas de tráfego.

ORIGEM/DESTINO	ZT1	ZT2	ZT3	...	ZT7	ZT8	PRODUÇÃO
ZT1	0	0	0	0	0	6	6
ZT2	0	0	0	0	0	0	0
ZT3	0	0	0	0	0	0	0
...	0	0	0	0	0	0	0
ZT7	0	0	0	0	0	0	0
ZT8	0	0	0	0	0	0	0
ATRAÇÃO	0	0	0	0	0	6	6

Fonte: Autor (2020).

Com a Tabela 4.7 pode-se observar que todos os dispositivos têm o embarque na primeira zona de tráfego e o desembarque na oitava região.

Para os setores censitários também foi montado a matriz manual para esse tipo de divisão territorial. A Tabela 4.8 apresenta as origens e os destinos dos passageiros para os setores censitários. Nesse caso as áreas foram numeradas de um a vinte e oito para a melhor compreensão da tabela.

Tabela 4.8: Matriz O/D manual dos *smartphones* para os setores censitários.

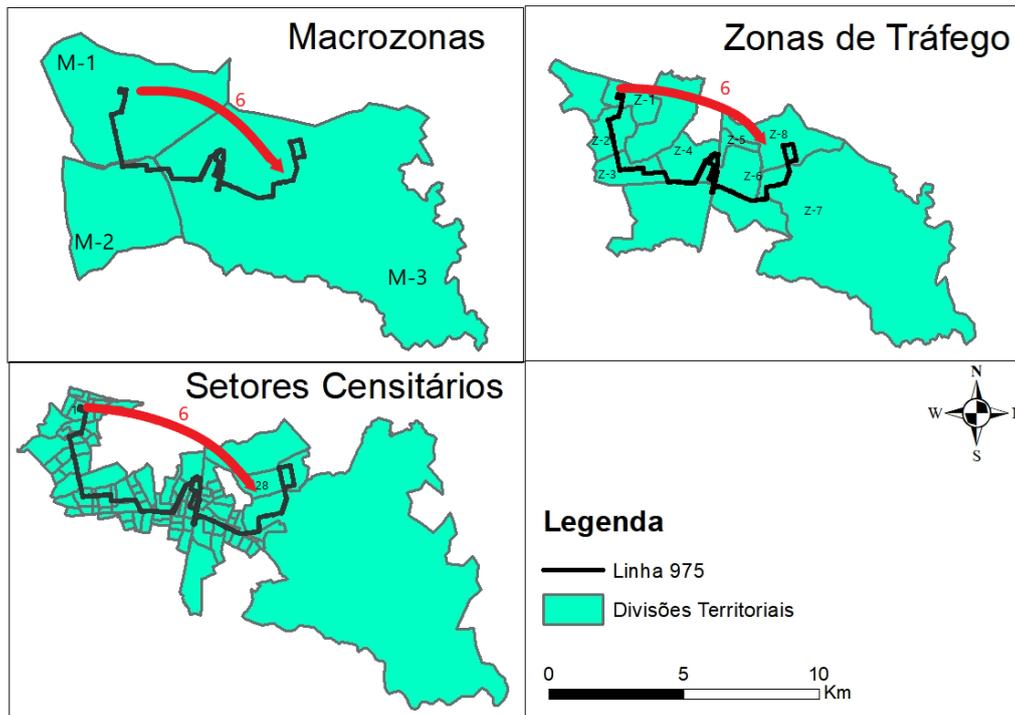
ORIGEM/DESTINO	SC1	...	SC6	SC7	...	SC10	...	SC27	SC28	PRODUÇÃO
SC1	0	0	0	0	0	0	0	0	6	6
...	0	0	0	0	0	0	0	0	0	0
SC6	0	0	0	0	0	0	0	0	0	0
SC7	0	0	0	0	0	0	0	0	0	0
...	0	0	0	0	0	0	0	0	0	0
SC10	0	0	0	0	0	0	0	0	0	0
...	0	0	0	0	0	0	0	0	0	0
SC27	0	0	0	0	0	0	0	0	0	0
SC28	0	0	0	0	0	0	0	0	0	0
ATRAÇÃO	0	0	0	0	0	0	0	0	6	6

Fonte: Autor (2020).

Em posse das matrizes O/D para cada divisão territorial pode-se montar a Figura 4.9, que ilustra os deslocamentos dos passageiros para diferentes divisões do espaço. Pode-se observar

que assim como citado anteriormente todos os passageiros terão origem na primeira área e destino na última área.

Figura 4.9: Matriz O/D manual para as diferentes divisões territoriais.



Fonte: Autor (2020).

4.5 RECONSTRUÇÃO DA MATRIZ O/D USANDO DADOS WI-FI

Nesta seção será realizada a projeção do deslocamento dos passageiros com base no perfil de transmissão dos dados Wi-Fi de seus dispositivos. Destaca-se que não foi possível testar em campo os deslocamentos previstos inicialmente, devido a pandemia da COVID-19. Para contornar tal problema, o experimento irá utilizar o perfil de transmissão dos *smartphones* para sintetizar a matriz.

Para o cálculo da projeção dos deslocamentos será utilizado a Equação 8, tal equação requer os dados de frequência do recebimento de pacotes Wi-Fi descritos na seção 4.3 e os valores de velocidade operacional da linha em estudo. A velocidade adotada nos cálculos foi de 28 km/h obtida diretamente no sítio eletrônico da RMTTC. Combinando esses dois dados, é possível projetar a trajetória de deslocamento dos passageiros e analisar suas possíveis posições de origem e destino.

A Tabela 4.9 mostra a distância média em que os dados do passageiro são capturados (deslocamento médio).

Tabela 4.9: Dados sobre os *smartphones*.

Aparelhos (MAC)	Frequência média de transmissão (s)	Velocidade média (28 km/h)	Deslocamento médio (m)
A0:4F:85:D7:12:DC	630	7,778	4900
C8:F3:19:60:48:77	472	7,778	3671
E4:58:E7:9D:CF:EF	448	7,778	3484
68:76:4f:0a:60:da	330	7,778	2567
40:45:DA:B6:82:BA	40	7,778	311
78:A8:73:95:37:A6	420	7,778	3267

Fonte: Autor (2020).

Considerando o valor máximo de transmissão de pacotes de dados que é 630 segundos (10.5 minutos), temos que o deslocamento médio do *smartphone* é 4,9 km, ou seja, o *sniffer* pode vir a capturar o pacote do dispositivo do passageiro após percorrer esta distância. Como dito, este dado é preocupante visto que em alguns casos o passageiro pode desembarcar antes do *sniffer* receber o primeiro pacote.

Com os valores médios de deslocamento foi possível realizar uma análise de quantos pacotes em média cada *smartphone* recebe durante o trajeto da linha 975. Este cálculo foi feito com base no comprimento total e no deslocamento médio, dessa forma pode-se observar na Tabela 4.10 os valores médios de envio de pacote para cada *smartphone*.

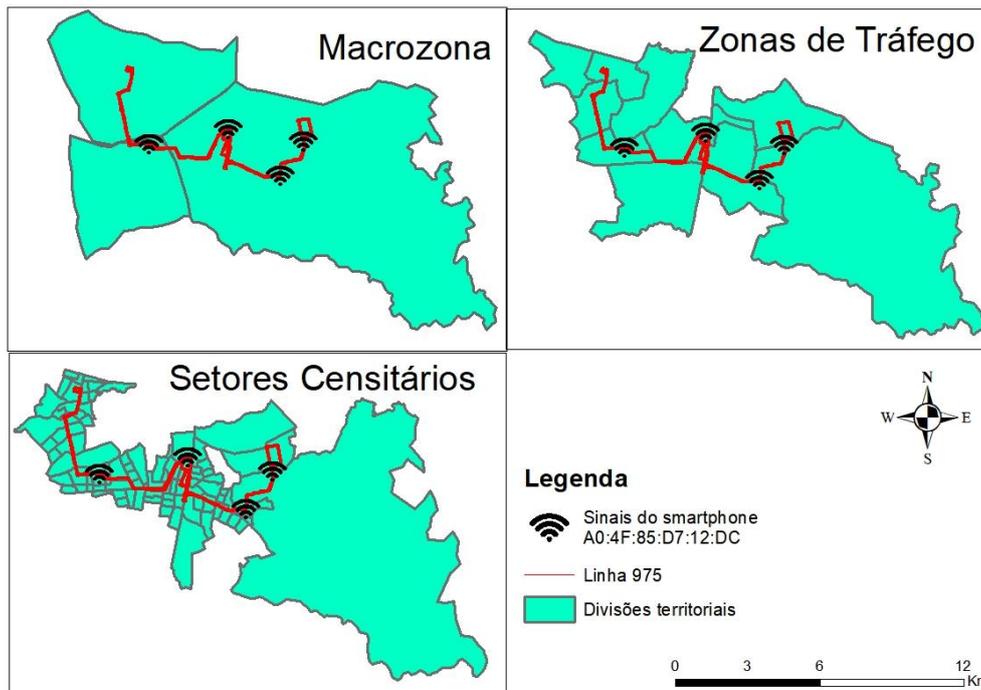
Tabela 4.10: Quantidade média de pacotes enviados por cada *smartphone*.

Aparelhos (MAC)	Quantidade média de pacotes enviado ao longo da viagem
A0:4F:85:D7:12:DC	4
C8:F3:19:60:48:77	5
E4:58:E7:9D:CF:EF	6
68:76:4f:0a:60:da	8
40:45:DA:B6:82:BA	70
78:A8:73:95:37:A6	6

Fonte: Autor (2020).

Após a obtenção do deslocamento médio conjuntamente com os valores da quantidade média de pacotes enviados por cada *smartphone* foi possível projetar espacialmente o deslocamento médio de cada passageiro. A Figura 4. 10 apresenta a posição de cada pacote de dados recebido do dispositivo A0:4F:85:D7:12:DC, e como é a espacialização desses dados para as diferentes divisões territoriais.

Figura 4.10: Espacialização dos sinais emitidos pelo *smartphone* A0:4F:85:D7:12:DC.



Fonte: Autor (2020).

Com a projeção da emissão dos sinais do *smartphone* A0:4F:85:D7:12:DC é possível observar que ele emitiria ao todo quatro sinais ao longo da linha, como ilustrado na Figura 4.10. Pode -se observar que esse *smartphone* teve uma baixa quantidade de emissão, o que pode tornar a matriz O/D do passageiro atrasada, pois muitas vezes o passageiro embarca em uma área, mas seu sinal é capturado em regiões posteriores.

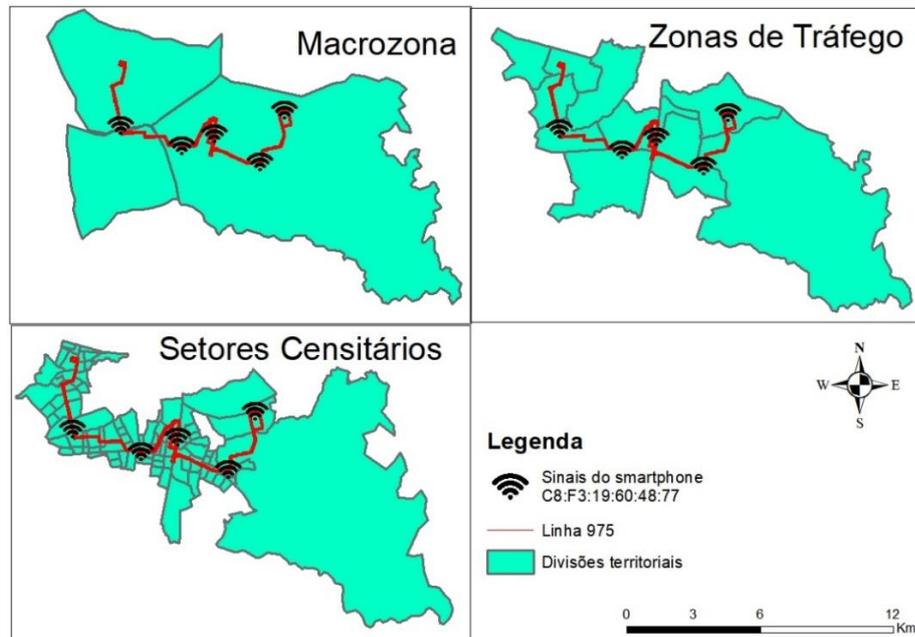
Note-se que apesar do embarque no cenário ocorrer no primeiro ponto de parada, a presença do passageiro só é identificada na segunda macrozona, na terceira zona de tráfego e no décimo setor censitário. Assim, pode-se observar que há uma defasagem entre o local de embarque e o obtido pela tecnologia. A mesma situação ocorre no desembarque, que apesar de identificar corretamente o aparelho na terceira macrozona, situa o aparelho na sétima zona de tráfego e no vigésimo sétimo setor censitário, sendo que o correto seria ser na oitava e vigésima oitava regiões respectivamente.

A Figura 4.11 apresenta a espacialização de cada pacote de sinal emitido pelo dispositivo C8:F3:19:60:48:77.

A projeção do *smartphone* C8:F3:19:60:48:77 indica que ao longo da linha seria emitido cinco pacotes de dados o que altera um pouco a configuração anterior, uma vez que quanto maior a quantidade de pacotes maior a facilidade de capturar o passageiro e, conseqüentemente, de se obter uma similaridade entre as origens e destinos das matrizes. Pode-se observar que o primeiro sinal foi recebido na primeira macrozona, mas para a zona de tráfego o sinal foi

captado na terceira área e já para os setores censitários o sinal foi captado na sétima área. Já os destinos apresentam uma maior compatibilidade do que os pares de origem. Os sinais são recebidos na terceira macrozona, na oitava zona de tráfego e na vigésima sétima área dos setores censitários.

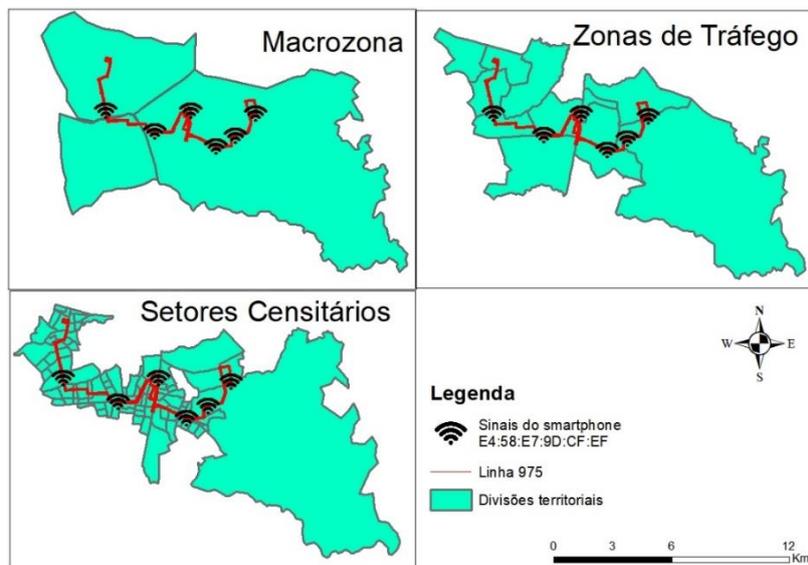
Figura 4.11: Espacialização dos sinais emitidos pelo *smartphone* C8:F3:19:60:48:77.



Fonte: Autor (2020).

A Figura 4.12 mostra como é a emissão dos pacotes de dados do *smartphone* E4:58:E7:9D:CF:EF ao longo das diferentes divisões territoriais.

Figura 4.12: Espacialização dos sinais emitidos pelo *smartphone* E4:58:E7:9D:CF:EF.



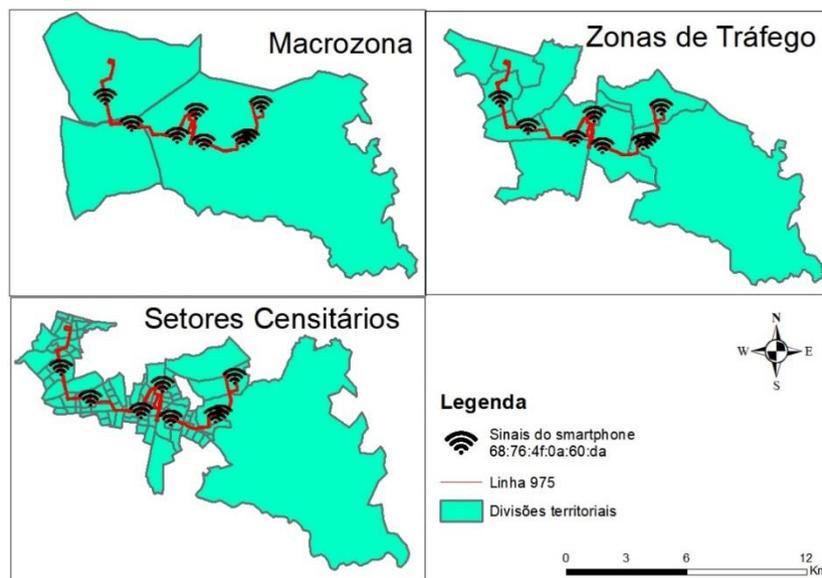
Fonte: Autor (2020).

O *smartphone* E4:58:E7:9D:CF:EF emitiu ao todo seis pacotes de dados ao longo da linha segundo os cálculos das projeções. Observa-se que a emissão do primeiro pacote de dados se

deu na primeira macrozona, na terceira zona de tráfego e no sétimo setores censitário. O último pacote de sinal foi recebido na terceira macrozona, na oitava zona de tráfego e vigésimo sétimo setor censitário. Destaca-se que os destinos são os valores esperados, mas a origem possui uma pequena discrepância para a zona de tráfego e para o setor censitário, onde estima-se que o embarque ocorreu respectivamente na terceira e sétima região.

A Figura 4.13 apresentada abaixo mostra como é a distribuição dos pacotes de dados para o *smartphone* 68:76:4F:0A:60:DA ao longo da linha 975.

Figura 4.13: Espacialização dos sinais emitidos pelo *smartphone* 68:76:4f:0a:60:da.



Fonte: Autor (2020).

O *smartphone* 68:76:4F:0A:60:DA emitiu oito pacotes de dados de acordo com a projeção o que pode influenciar diretamente na construção da matriz O/D. O primeiro sinal foi recebido na primeira macrozona, na segunda zona de tráfego e no sexto setor censitário. Já o último sinal foi captado na terceira macrozona, na oitava zona de tráfego e vigésimo sétimo setor censitário.

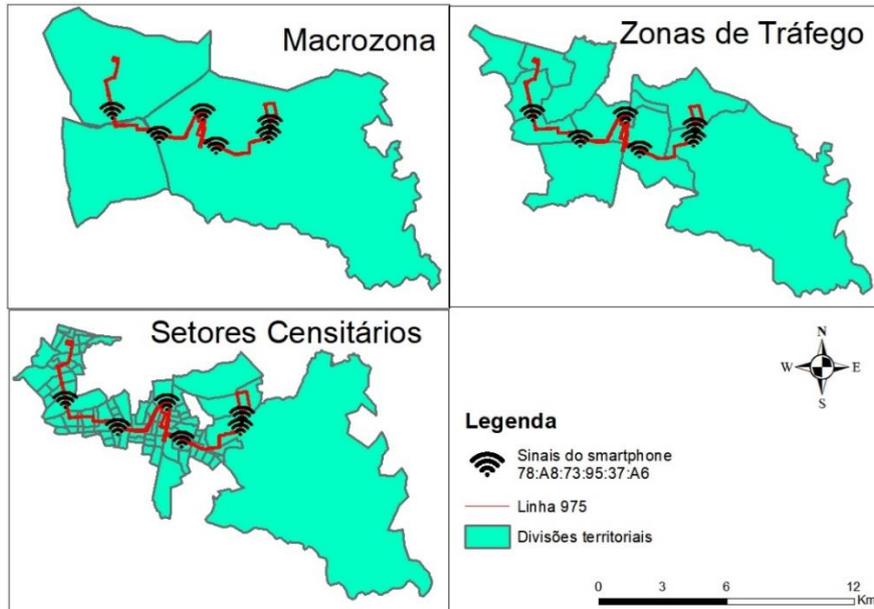
Já a Figura 4.14 apresentará os dados dos pacotes de dados recebido pelo dispositivo 78:A8:73:95:37:A6 para diferentes divisões do espaço.

O *smartphone* 78:A8:73:95:37:A6 de acordo com a projeção emitiria seis pacotes de dados ao longo da linha. O primeiro sinal emitido foi na primeira macrozona, na terceira zona de tráfego e sexto setor censitário. Já o último sinal foi captado na terceira macrozona, na sétima zona de tráfego e no vigésimo sétimo setor censitário.

Para o *smartphone* 40:45:DA:B6:82:BA não foi montado uma figura, devido a quantidade de pacotes de dados emitidos (70 pacotes de dados), o que tornaria a figura de difícil visualização, pois os dados ficariam muito próximo entre si. Desse modo é preferível descrever

que o primeiro sinal foi detectado nas primeiras áreas de todas as divisões territoriais e o último sinal captado foi nas últimas áreas de cada divisão. Tal fato pode ser constatado devido ao espaçamento médio entre os pacotes de dados recebidos serem de 311 metros.

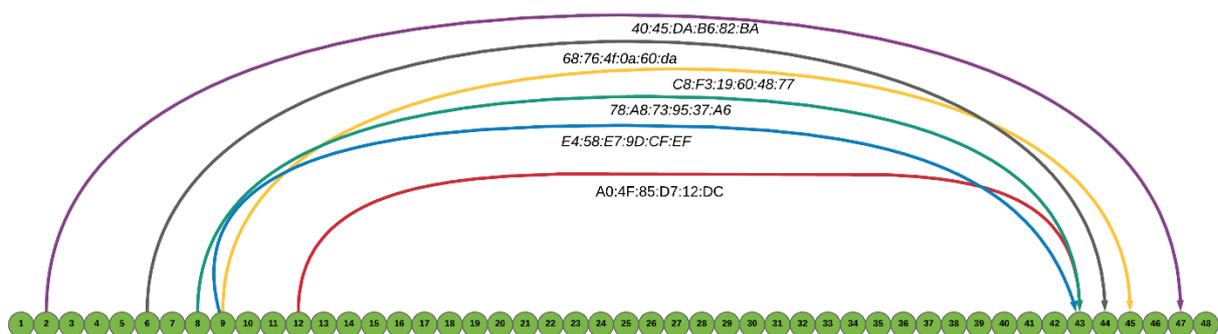
Figura 4.14: Espacialização dos sinais emitidos pelo *smartphone* 78:A8:73:95:37:A6.



Fonte: Autor (2020).

Após a espacialização dos pacotes de dados recebidos por cada *smartphone* pode-se construir a matriz ilustrada pela Figura 4.15. Observe-se que apesar do cenário contemplar que todos os passageiros embarcam no primeiro de ponto de parada, os dados obtidos pela projeção com a tecnologia mostram divergências entre os locais de embarque e desembarque. A saber, considere o dispositivo 68:76:4f:0a:60:da. O mesmo teve sua presença capturada apenas no sexto ponto de parada e o desembarque detectado na quadragésima quarta estação. Para compreender como esta diferença afeta a matriz obtida, posteriormente será realizado um detalhamento do deslocamento projetado para cada dispositivo.

Figura 4.15: Matriz O/D ponto a ponto obtida.



Fonte: Autor (2020).

Tomando como base a matriz O/D ponto a ponto obtida, ilustrada pela Figura 4.15 e relatada nos resultados acima, pode-se montar a matriz O/D dos *smartphones* para cada divisão territorial. A Tabela 4.11 apresenta a matriz O/D obtida com dados dos *smartphone* para a divisão em macrozonas. As macrozonas foram numeradas a fim de facilitar a compreensão e montagem da matriz O/D ao total são três macrozonas.

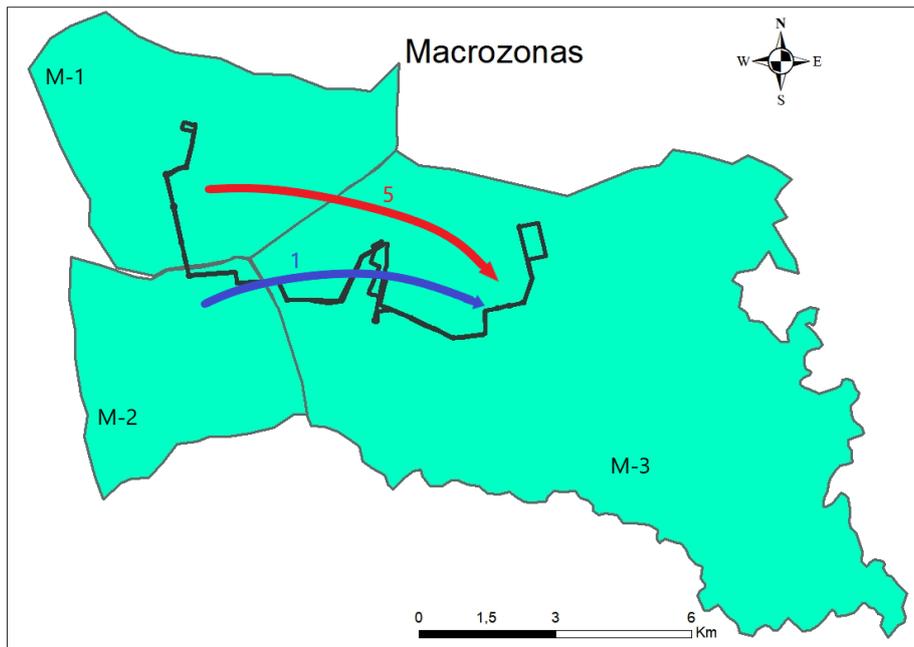
Tabela 4.11: Matriz O/D dos *smartphones* para as macrozonas.

ORIGEM/DESTINO	MC1	MC2	MC3	PRODUÇÃO
MC1	0	0	5	5
MC2	0	0	1	1
MC3	0	0	0	0
ATRAÇÃO	0	0	6	6

Fonte: Autor (2020).

Além da esquematização dos elementos da matriz O/D dos dispositivos podemos também visualizar com a Figura 4.16 como é a distribuição dos *smartphones* ao longo das áreas das macrozonas.

Figura 4.16: Matriz O/D obtida para as macrozonas.



Fonte: Autor (2020).

Pode-se observar que cinco dos seis *smartphones* saem da primeira macrozona e se deslocam até a terceira macrozona e apenas um dos seis *smartphones* saem da segunda macrozona e vão para a terceira macrozona.

A Tabela 4.12 mostra a matriz O/D dos dispositivos para a zonas de tráfego. No total são oito áreas que foram numeradas a fim de simplificar o entendimento da matriz.

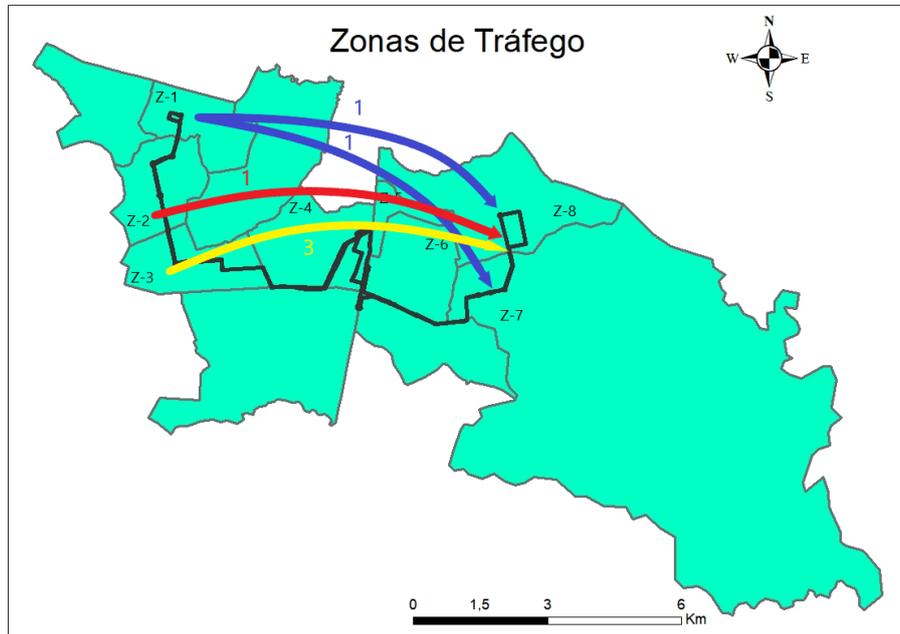
Tabela 4.12: Matriz O/D dos *smartphones* para as zonas de tráfego.

ORIGEM/DESTINO	ZT1	ZT2	ZT3	...	ZT7	ZT8	PRODUÇÃO
ZT1	0	0	0	0	1	1	2
ZT2	0	0	0	0	0	1	1
ZT3	0	0	0	0	0	3	3
...	0	0	0	0	0	0	0
ZT7	0	0	0	0	0	0	0
ZT8	0	0	0	0	0	0	0
ATRAÇÃO	0	0	0	0	1	5	6

Fonte: Autor (2020).

Assim como para as macrozonas também foi criado a Figura 4.17 para a visualização da distribuição dos pares de embarque e desembarque nas zonas de tráfego.

Figura 4.17: Matriz O/D obtida para as zonas de tráfego.



Fonte: Autor (2020).

Com a Figura 4.17 pode-se observar que devido a áreas serem menores comparada com as macrozonas, a distribuição dos pares de origem e destino nesse tipo de divisão já apresenta uma característica diferente da apresentada anteriormente. Desse modo é possível visualizar que um dos *smartphones* se deslocam da primeira área e vai para oitava região, outro dispositivo se desloca da primeira área para a sétima região, o terceiro aparelho se desloca da segunda área para a oitava região, e os demais dispositivos se deslocam da terceira área para a oitava região.

Assim como para as demais divisões territoriais também foi montada a matriz O/D dos aparelhos para os setores censitários. Ao todo são vinte e oito áreas que foram numeradas para facilitar a visualização da distribuição dos aparelhos ao longo das regiões como mostra a Tabela 4.13.

Os setores censitários apresentam áreas menores que demais divisões, isso implica que os aparelhos estarão mais espaçados entre as regiões. Pode-se observar que apenas que a vigésima sétima área atraiu cinco dispositivos e a vigésima oitava apenas um dispositivo.

A Figura 4.18 apresenta os deslocamentos dos *smartphones* ao longo das áreas das divisões dos setores censitários.

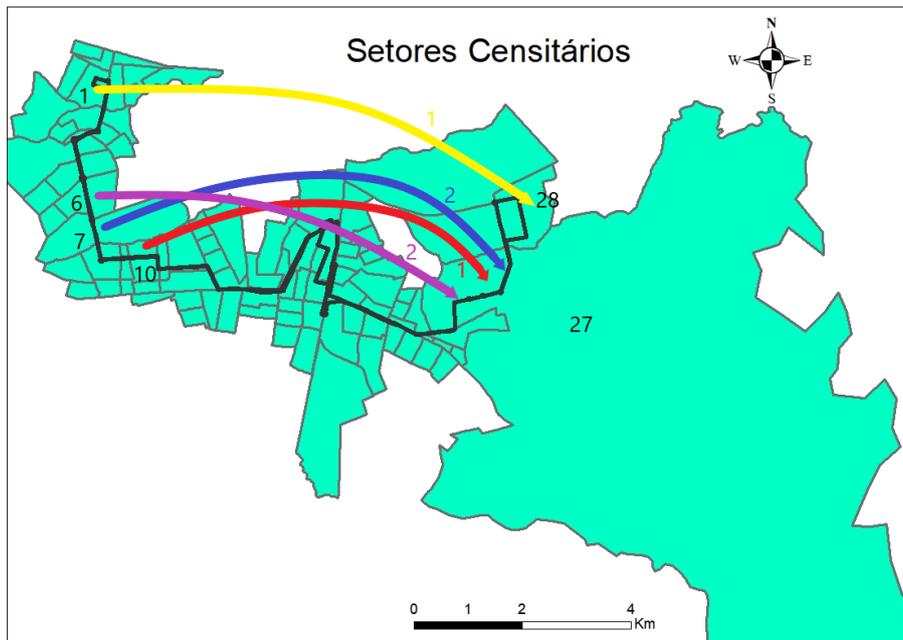
Pode-se observar na Figura 4.18 que nesse tipo de divisão a distribuição dos *smartphones* entre as áreas são diferentes das demais. Os dispositivos já apresentam diferença entre seus pares de origem e destino, uma vez que as áreas são menores logo as origens e destinos tendem a ser mais diferentes entre os aparelhos.

Tabela 4.13: Matriz O/D dos *smartphones* para as zonas de tráfego.

ORIGEM/DESTINO	SC1	...	SC6	SC7	...	SC10	...	SC27	SC28	PRODUÇÃO
SC1	0	0	0	0	0	0	0	0	1	1
...	0	0	0	0	0	0	0	0	0	0
SC6	0	0	0	0	0	0	0	2	0	2
SC7	0	0	0	0	0	0	0	2	0	2
...	0	0	0	0	0	0	0	0	0	0
SC10	0	0	0	0	0	0	0	1	0	1
...	0	0	0	0	0	0	0	0	0	0
SC27	0	0	0	0	0	0	0	0	0	0
SC28	0	0	0	0	0	0	0	0	0	0
ATRAÇÃO	0	0	0	0	0	0	0	5	1	6

Fonte: Autor (2020).

Figura 4.18: Matriz O/D obtida para os setores censitários.



Fonte: Autor (2020).

4.6 COMPARAÇÃO DAS MATRIZES O/D

A última etapa do método consiste em comparar a similaridade entre as matrizes. Apesar de realizar tal comparação indiretamente na última subseção, aqui almeja-se encontrar a similaridade conforme uma métrica. Para tal, decidiu-se utilizar a diferença média entre as matrizes O/D manual com a obtida com dados Wi-Fi, como relatado na Equação 7.

Observa-se que para a matriz manual os *smartphones* embarcaram no primeiro ponto de parada e desembarcaram na última estação. A matriz O/D usando a tecnologia Wi-Fi dos aparelhos é apresentada na seção 4.5. Dessa forma ao realizar o cálculo para as diferentes divisões territoriais temos a Tabela 4.14 que apresenta os valores de similaridade entre as duas matrizes construídas.

Tabela 4.14: Comparação entre as origens das matrizes O/D para diferentes granularidades.

	Similaridade entre as matrizes O/D
Macrozonas	0,167
Zonas de Tráfego	1,167
Setores Censitários	5,167
Ponto a Ponto	6,5

Fonte: Autor (2020).

Pode-se observar que os valores de similaridade entre as matrizes O/D são diferentes para cada resolução, quanto maiores são as áreas mais similares são as matrizes, uma vez que as mesmas agrupam mais pontos de parada, conseqüentemente, menor a discrepância entre as origens e destinos. Nota-se que o tempo de transmissão de frequência dos *smartphones* em áreas maiores fazem com que os pares de origem e destino se encontrem na mesma região dos pares reais. Por exemplo, ao considerar o agrupamento de pontos de parada conforme os limites das zonas de tráfego, observa-se uma defasagem média de 1,167 zona de tráfego da origem real. Já o cenário que utiliza setores censitários observa-se uma diferença média nas origens da matriz projetada de 5,167 setores censitários daquela prevista manualmente.

Os dados sugerem uma proximidade entre os valores de similaridade entre os setores censitários e ponto a ponto são muito altos comparados ao demais, isso mostra que a granularidade entre as duas divisões são similares entre si. Especificamente, no caso dos setores censitários observa-se uma discrepância entre as matrizes de 5,167 setores, enquanto que para a diferença entre as matrizes ponto a ponto é de 6,5 pontos de parada entre os locais de embarque.

Semelhantemente como para origem, no cálculo da similaridade entre os desembarques dos passageiros foi usada a Equação 7. Na seção 4.4 é apresentada a matriz O/D manual e na seção 4.5 é apresentado a matriz O/D usando a tecnologia Wi-Fi. Em posse desses valores pode-se realizar o cálculo da similaridade das matrizes para diferentes granularidades. Os resultados das similaridades das matrizes O/D para diferentes granularidades é apresentada na Tabela 4.15.

Tabela 4.15: Comparação entre os destinos das matrizes O/D para diferentes granularidades.

	Similaridade entre as matrizes O/D
Macrozonas	0
Zonas de Tráfego	0,333
Setores Censitários	0,333
Ponto a ponto	2,83

Fonte: Autor (2020).

Diferentemente da similaridade entre as origens os pares de destino apresentaram uma similaridade maior entre si. O que pode ser justificado pelo fato de que o tamanho das áreas finais das divisões territoriais ser maior, desse modo aumentando assim a similaridade entre as matrizes O/D. Nota-se que para as macrozonas o destino das matrizes O/D são idênticas. Para as zonas de tráfego e setores censitários vemos que a discrepância entre as matrizes é de 0,333 em relação ao desembarque real. Já para a ponto a ponto a discrepância entre as matrizes é de 2,83 pontos de parada.

5 CONSIDERAÇÕES FINAIS

O trabalho teve como objetivo principal verificar a viabilidade da reconstrução da matriz O/D a partir da utilização e captação de dados Wi-Fi dos *smartphones* dos passageiros do transporte público, considerando as diferentes divisões urbanas. A ideia de explorar diferentes agrupamentos de pontos de parada tem o intuito de flexibilizar a estrutura da matriz na tentativa de contornar o problema da frequência de transmissão de dados da tecnologia Wi-Fi, uma vez que o padrão IEEE 802.11 não especifica a periodicidade que os aparelhos devem transmitir os dados com essa tecnologia. Tal problema pode dificultar a captura e, consequentemente, degradar a matriz de origem e destino obtida, uma vez que o passageiro pode adentrar no veículo em um ponto de parada e sua presença só ser capturada em pontos subsequentes. Para tal, propôs-se comparar a semelhança entre as matrizes O/D geradas com os dados coletados dos *smartphones*, com diferentes granularidades, com aquela construída manualmente

Para analisar a aplicabilidade da tecnologia Wi-Fi para esta finalidade, o trabalho apresentou uma metodologia composta de seis etapas. A primeira consiste em analisar as políticas de divisões territoriais das cidades, com o intuito de identificar e apresentar quais são as principais divisões de uma cidade. A segunda etapa visa identificar e detalhar uma linha de transporte público escolhida. Já a terceira etapa visa construir a matriz O/D dos dados Wi-Fi. A quarta etapa determina o perfil de transmissão dos pacotes enviados pelos *smartphones*. A quinta etapa visa a realização da projeção dos deslocamentos dos passageiros com base na tecnologia Wi-Fi. A sexta e última etapa consiste em realizar a comparação entre as matrizes O/D em diferentes granularidades.

Para avaliação do método, previa-se a coleta de dados manualmente com pesquisadores embarcados no transporte público da cidade de Aparecida de Goiânia. Entretanto, devido a pandemia do COVID-19, não foi possível realizar os experimentos em campo. No entanto, modificou-se a metodologia para se adequar a esse cenário. A saber, decidiu-se fazer os experimentos de forma local com a finalidade de extrair o perfil de transmissão dos dados Wi-Fi dos dispositivos, como a quantidade e frequência de dados emitidos.

O experimento foi realizado com seis dispositivos durante um período de duas horas. De forma que cada aparelho representa um passageiro do transporte público. Para o experimento deixou-se os *smartphones* parados e com a tela desligada, como se estivessem dentro do bolso dos passageiros e sem receber mensagens. É importante destacar que tal cenário apresenta o pior caso de transmissão de dados Wi-Fi, uma vez que os aparelhos emitem uma

quantidade inferior de dados pois o sistema operacional considera que o dispositivo se encontra ocioso.

Ao considerar diferentes granularidades de uma cidade para análise de uma linha de transporte público conclui-se que alguns tipos de agrupamentos auxiliam na construção da matriz O/D usando os dados dos pacotes de sinais enviados por *smartphone*. Como esperado, os experimentos realizados sugerem que quanto maior são as áreas, menor é o erro na estimação das matrizes O/D. Entretanto, quanto maior são essas áreas menor será o detalhamento das informações, uma vez que nessas regiões possui muitos pontos de parada agrupados no mesmo local, dificultando análises pontuais do local.

Os valores encontrados na comparação das matrizes O/D manual e usando tecnologia Wi-Fi ficaram muito próximo daquele que é desejado (zero). Desse modo ressalta-se esses resultados, visto que as divisões das áreas das zonas de tráfego são menores que as macrozonas. Tais resultados são importantes para o planejamento do transporte público pois auxiliam no estudo mais amplo das linhas que operam na cidade comparado ponto a ponto. Uma questão que merece atenção é o fato de que ao dividir a cidade em áreas maiores pode ocorrer de vários pontos de parada se situarem nas bordas dessas divisões. Esses pontos de parada situados nas bordas das regiões podem alterar os pares de origem e destino dos passageiros, uma vez que o sensor GPS possui uma imprecisão de 10 a 30 metros, ou seja, quando um dispositivo emite um sinal próximo aquele ponto de parada na borda devido a imprecisão do GPS a origem ou destino de um passageiro mudará de acordo com GPS.

Embora seja possível estimar a matriz O/D usando a tecnologia Wi-Fi, tal abordagem deve apresentar cautela. Devido a diversidade de aparelhos encontrados no mercado e de que não possui uma padronização de emissão de pacotes de dados, existe uma discrepância na frequência de transmissão de dados. Com base nos resultados encontrados sabe-se que um dos aparelhos apresentou uma frequência alta, enviando dados a cada 40 segundos. Neste caso é possível inclusive construir uma matriz O/D ponto a ponto. Entretanto, outros aparelhos apresentaram uma frequência menor, com intervalos de até dez minutos para transmissão dos dados. Desse modo, é difícil a captura de dados em pequenas viagens e também a estimação do local exato de embarque ou desembarque do passageiro.

Mesmo que ocorra erros na estimação da matriz O/D dos passageiros para granularidade menores, a utilização da tecnologia Wi-Fi é um recurso importante para o planejamento e operação da rede de transporte. Uma vez que a estimação da matriz O/D de forma tradicional por métodos diretos é bastante onerosa é a utilização de tal tecnologia auxiliaria diversos órgãos responsáveis pelo planejamento e operação do transporte público, visto que em grande parte

das cidades brasileiras não se tem nenhuma informação sobre os deslocamentos dos passageiros na cidade.

Uma dificuldade encontrada no trabalho foi a disponibilidade de aparelhos para a realização dos experimentos já que para a análise dos dados era necessário a identificação do endereço MAC dos dispositivos.

Recomenda-se para trabalhos futuros um estudo mais detalhado com diferentes aparelhos para compreender o impacto da frequência de transmissão de dados nos vários modelos de dispositivos. Como também um estudo aprofundado das variações da taxa de transmissão de envio de pacotes de dados, uma vez que ao adotar a média esconde-se uma variação entre os dados. Dessa forma para testar os limites da tecnologia, deve optar por escolher o maior intervalo de transmissão de cada um dos smartphones. Além disso, recomenda-se um estudo considerando vários dias da semana, e vários horários, a fim de diminuir os problemas encontrados.

REFERÊNCIAS BIBLIOGRÁFICAS

- AFSHARI, H. H.; JALALI, S.; GHODS, A. H.; RAAHEMI, B. An Intelligent Traffic Management System Based on the Wi-Fi and Bluetooth Sensing and Data Clustering. *Advances in Intelligent Systems and Computing*. Cham: Springer International Publishing, 2019. v. 880. p. 298–312. Disponível em: <<http://link.springer.com/10.1007/978-3-030-02686-8>>. Acesso em: 15 de agosto de 2019.
- ANTP: ASSOCIAÇÃO NACIONAL DE TRANSPORTES PÚBLICOS. *Sistemas Inteligentes de Transportes*. Série Cadernos Técnicos, vol. 8. São Paulo, 2012, 163 p.
- APARECIDA DE GOIÂNIA, Lei Complementar Nº 5, de 30 de janeiro de 2002. *Zoneamento e ocupação do Solo, Aparecida de Goiânia*, 2002.
- BAUER, C. 2013. On the (In-)Accuracy of GPS Measures of Smartphones: A Study of Running Tracking Applications. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia (MoMM '13)*. Association for Computing Machinery, New York, NY, USA, 335–341. DOI:<https://doi.org/10.1145/2536853.2536893>.
- BERTONCINI, B. V. *Uma Proposta de Carregamento Incremental de Fluxos Veiculares Para Estimaco De Matriz O-D Sinttica*. Dissertao (Mestrado em Engenharia Civil) -Escola de Engenharia de So Carlos de So Paulo. Universidade de So Paulo, So Carlos, 2007.
- BEZERRA, J. A. *Como Definir O Bairro? Uma Breve Reviso*. 2011. *Temas GEO*. Pau dos Ferros, Rio Grande do Norte, Brasil, v. 1, n. 1, p. 21-31, jan./jun., 2011.
- BOISIER, S. *Em busca do esquivo desenvolvimento regional: entre a caixa-preta e o projeto poltico*. Planejamento e polticas pblicas, 2009.
- BRASIL. *Estatuto da Cidade: guia para implementaco pelos municpios e cidados*. 2 ed. Braslia: Cmara dos Deputados, Coordenao de Publicao, 2002.
- CACERES, N; WIDEBERG, J.P.; BENITEZ, F.G. Review of traffic data estimations extracted from cellular networks. *London: IET Intell. Transp. Syst.*, 2008, Vol. 2, No. 3, pp. 179–192 doi: 10.1049/iet-its:20080003, 2009. p. 2008–2010.
- CAMPOS, V. B. G. *Planejamento de Transporte. Conceitos e Modelos*. Editora Intercincias. Rio de Janeiro 2013. 188 p.
- CAZZOLATO, J. D. *Os bairros como instncia territorial local - contribuio metodolgica para o caso de So Paulo*. 2005. Dissertao (Mestrado em Geografia Humana) - Faculdade de Filosofia, Letras e Cincias Humanas, Universidade de So Paulo, So Paulo, 2005. doi: 10.11606/D.8.2005.tde-22022006-234813. Acesso em: 2020-09-23.

- CIGOLINI, A.A.; TERRITÓRIO E CRIAÇÃO DE MUNICÍPIOS NO BRASIL Uma abordagem histórico-geográfica sobre a compartimentação do espaço. 2009. Dissertação (Doutorado em Geografia) - Centro de Filosofia e Ciências Humanas da Universidade Federal de Santa Catarina. Santa Catarina, 2009.
- CROW, B. P.; WIDJAJA, I.; KIM, J. G.; SAKAI, P. T. (1997) IEEE 802.11 Wireless Local Area Networks. IEEE Communications Magazine, 35(9), p. 116–126.
- CUNCHE, M.; KAAFAR, M.; BORELI, R.; Linking wireless devices using information contained in Wi-Fi probe requests. 2014. Pervasive and Mobile Computing. Volume 11, p. 56-69.
- CUNHA, C. A. Relações Entre Características De Linhas De Transporte Coletivo Urbano E De Áreas Urbanas. Dissertação (Mestrado em Engenharia Civil e Ambiental) - Universidade de Brasília. Brasília, 2005.
- DIÁRIO DO TRANSPORTE. HP Transportes compra 70 chassis Mercedes-Benz para renovação em Goiânia. Disponível em :<
<https://diariodotransporte.com.br/2014/12/19/hp-transportes-compra-70-chassis-mercedes-benz-para-renovacao-em-goiania/> >. Acesso em: 28 de dezembro de 2020.
- DUNLAP, M.; LI, Z.; HENRICKSON, K; WANG, Y. Estimation of Origin and Destination Information from Bluetooth and Wi-Fi Sensing for Transit. Transportation Research Record: Journal of the Transportation Research Board, v. 2595, n. 1, p. 11–17, Jan 2016. Disponível em: <<http://journals.sagepub.com/doi/10.3141/2595-02>>. Acesso em: 21 de junho de 2019.
- EL-TAWAB, S.; ORAM, R.; GARCIA, M.; JOHNS, C.; PARK, B. B.; et al. Data analysis of transit systems using low-cost IoT technology. IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). Hawaii, USA. 2017. p. 497–502. Disponível em: <<https://ieeexplore.ieee.org/document/7917613/>>. Acesso em: 9 de julho de 2019.
- EL-TAWAB, S.; YORIO, Z.; SALMAN, A.; ORAM, R.; PARK, B. B. Origin-Destination Tracking Analysis of an Intelligent Transit Bus System using Internet of Things. IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). 2019. p. 139–144. Disponível em: <<https://ieeexplore.ieee.org/document/8730746/>>. Acesso em: 9 de julho de 2019.
- ESTATUTO DA CIDADE – Lei 10.257, de 10.07.2001 Comentários. Ed. Revista dos Tribunais, 2002.
- FERRAZ, A. C. C. P e TORRES, I. G. E. Transporte Público Urbano. 2 ed. Editora Rima, São

- Carlos, SP, 2004.
- FERRAZ, A. C. P.; TORRES, I. G. E. Transporte Público Urbano. São Carlos: Rima, 2004. P. 112.
- FERREIRA, E. A. Um método de utilização de dados de pesquisa embarque/desembarque na calibração de modelos de distribuição do tipo gravitacional. Dissertação (Mestrado em Engenharia Civil) -Escola de Engenharia de São Carlos de São Paulo. Universidade de São Paulo. São Carlos, p. 110. 1999.
- FERRONATTO, L. G. Análise de Capacidade de Terminais Urbanos de Ônibus: Metodologia Adaptada 2005. Congresso de Pesquisa e Ensino em Transportes, ANPET, 2005.
- FREUDIGER, J. How Talkative is your Mobile Device? An Experimental Study of Wi-Fi Probe Requests. WiSec '15 Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks, 2015. p. 1–6. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2766498.2766517>>. Acesso em: 21 de junho de 2019.
- GAST, M. S. (2005) 802.11 Wireless Networks: The Definitive Guide, Second Edition. O'Reilly Media, Inc.
- GUERRA, A. L. Determinação de Matriz Origem/Destino Utilizando Dados do Sistema de Bilhetagem Eletrônica. Dissertação (Mestrado em Geotecnia e Transportes) Universidade Federal de Minas Gerais. Belo Horizonte, 2011.
- GUIMARÃES, M. A.; FILHO, L. E. C. Análise Da Percepção Da Qualidade Do Transporte Público Na Região Metropolitana De Goiânia. Congresso de Pesquisa e Ensino em Transportes, ANPET. Gramado, RS, Brazil: 2018. p. 830–841.
- HAKEGARD, J. E.; MYRVOLL, T. A.; SKOGLUND, T. R. Statistical Modelling for Estimation of OD Matrices for Public Transport Using Wi-Fi and APC Data. IEEE International Conference on Pervasive Computing and Communications Workshops, 2018. p. 1005–1010. Disponível em: <<https://ieeexplore.ieee.org/document/8570009/>>. Acesso em: 21 de junho de 2019.
- HIDALGO, B. D. (2013). As divisões territoriais do Município de São Paulo: uma proposta de classificação por meio da análise dos Distritos. Monografia. Departamento de Geografia da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo.
- HIDAYAT, A.; TERABE, S.; YAGINUMA, H. Determine Non-Passenger Data from WiFi Scanner Data (MAC Address), A Case Study: Romango Bus, Obuse, Nagano Prefecture, Japan. International Review for Spatial Planning and Sustainable Development, v. 6, n. 3, p. 154–167, 15 Jul 2018. Disponível em:

- <https://www.jstage.jst.go.jp/article/irspsd/6/3/6_154/_article>. Acesso em: 21 de junho de 2019.
- HOLTERMANN, R. S. Um Estudo Sobre o Transporte Urbano De Passageiros Do Distrito Federal: Dinâmica da mobilidade em áreas centrais. Dissertação (Tese de Conclusão de Curso em Engenharia de Produção) -Faculdade de Tecnologia. Universidade De Brasília. Brasília, 2016.
- HU, B. ; GUO, H.; ZHOU, Z; SHI, Z. L. Characteristics of SARS-CoV-2and COVID-19, *Nature Reviews Microbiology* (December) (2020).doi:10.1038/s41579-020-00459-7.120.
- IBGE, 2013. Notas Metodológicas. Censo 2010. Disponível em: https://www.ibge.gov.br/apps/snig/v1/notas_metodologicas.html?loc=0. Acesso em: 20 de julho de 2019.
- JALALI, S. Estimating Bus Passengers' Origin-Destination of Travel Route Using Data Analytics on Wi-Fi and Bluetooth Signals. University of Ottawa. 2019.
- JAMIL, S. et al. Classifying smartphone screen ON/OFF State Based on Wi-Fi probe patterns. *UbiComp 2016. Adjunct – Proceedings of the 2016. ACM International Joint Conference on Pervasive and Ubiquitous Computing*. p. 301- 304, 2016.
- JI, Y.; ZHAO, J.; ZHANG, Z.; DU, Y. Estimating Bus Loads and OD Flows Using Location-Stamped Farebox and Wi-Fi Signal Data. *Journal of Advanced Transportation*, v. 2017, n. 8, p. 1–10, 2017. Disponível em: <<https://www.hindawi.com/journals/jat/2017/6374858/>>. Acesso em: 21 de junho de 2019.
- KNEIB, E. C; Subcentros Urbanos: Contribuição Conceitual e Metodológica á sua Definição e Identificação para Planejamento de Transportes. Dissertação (Doutorado em Transportes). Universidade de Brasília. Brasília, 2008.
- KURKCU, A.; OZBAY, K. Estimating pedestrian densities, wait times, and flows with wi-fi and bluetooth sensors. *Transportation Research Record*, v. 2644, n. 1, p. 72–82, 2017.
- LEFEBVRE, H. Barrio y vida de barrio. In: _____. *De lo rural a lo urbano*. 3. ed. Barcelona: Ediciones Península, 1975, p. 195-203.
- LOUREIRO, C. F. G.; MATO, A. R. P.; BERTONCINI, B. V. Abordagem Bayesiana na Estimção de Matrizes de Origem e Destino Sintéticas em Redes de Transportes. XXV Congresso de Pesquisa e Ensino em Transportes. ANPET, Belo Horizonte, 2011.
- MESQUIRA, H. C.; AMARAL, M.J. Matriz O/D com Base nos Dados do Sistema de Bilhetagem Eletrônica. Tese de Conclusão de Curso. Engenharia Civil. UFG.

- Goiânia.2016.
- MIKKELSEN, L., BUCHAKCHIEV, R., MADSEN, T., & SCHWEFEL, HP (2016). Estimativa de ocupação de transporte público usando sondagem WLAN. Em Proceedings of 2016 8th International Workshop on Resilient Networks Design and Modeling, RNDM 2016 (pp. 302–308).
- MISHALANI, R. G.; MCCORD, M. R.; REINHOLD, T. Use of Mobile Device Wireless Signals to Determine Transit Route-Level Passenger Origin–Destination Flows: Methodology and Empirical Evaluation. *Transportation Research Record: Journal of the Transportation Research Board*, v. 2544, n. 1, p. 123–130, Jan 2016. Disponível em: <<http://journals.sagepub.com/doi/10.3141/2544-14>>. Acesso em: 19 de julho de 2019.
- NIELSEN, O. A. Two New Methods for Estimating Trip Matrices from Traffic Counts. In: ORTÚZAR, J. D. D.; HENSHER, D. A.; JARA-DIAZ, S. *Travel Behaviour Research: Updating the State of Play*. Oxford: Elsevier Science Ltd., 1998. Cap. 13, p. 221- 250.
- ORANSIRIKUL, T., NISHIDE, R., PIUMARTA, I., E TAKADA, H. (2016) Feasibility of analyzing Wi-Fi activity to estimate transit passenger population. *Proc. Intl. Conf. on Adv. Info. Networking and Applications, AINA, 2016-May*.
- ORTÚZAR, J. D.; WILLUMSEN, L. G. *Modelling Transport*. 4th ed. Wiley, 2011. 606 p.
- PARADEDA BENITES, Diego e KRAUS JUNIOR, Werner e CARLSON, Rodrigo Castelan. Estimativa De Embarque E Desembarque De Passageiros Por Detecção De Sinal De Wi-Fi De Dispositivos Móveis. 2018, Gramado, RS, Brazil: ANPET, 2018. p. 988–
- PEIXOTO, A. G. Método para a Construção de uma Matriz de Origem e Destino Utilizando Algoritmo para Localização de Trechos para a Realização das Pesquisas. *Dissertação (Mestrado em Engenharia Civil) Universidade Federal de Santa Catarina, 2013*. Disponível em: <<https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/122815/323241.pdf?sequence=1&isAllowed=y>>. Acesso em 19 de agosto de 2019.
- PELLETIER, M.-P.; TRÉPANIER, M.; MORENCY, C. Smart card data use in public transit: A literature review. *Transp. Res. Part C Emerg. Technol.* 19 (2011) 557–568.
- PERA, C. K. L.; BUENO, L. M. M. (2016). Revendo o uso de dados do IBGE para pesquisa e planejamento territorial: reflexões quanto à classificação da situação urbana e rural. *Cad. Metropolitano, São Paulo*.
- PLANO DIRETOR DE CAMPINAS. 2006. Disponível em:<https://planodiretor.campinas.sp.gov.br/timeline/timeline/41_nova_versao_caderno_subsidios_janeiro_17/Caderno_com_anexos.pdf>. Acesso em: 24 de maio de 2020.

- PLANO DIRETOR DO RIO DE JANEIRO. Lei Complementar n.º 111 de 1º de fevereiro de 2011.
- RAGAZZO, C.; LIMA, J. (2013). Planejamento urbano e redução de trânsito: o caso da estação de metrô da Nossa Senhora da Paz, em Ipanema. *Urbe : Revista Brasileira de Gestão Urbana*. 5. 10.7213/urbe.05.002.AC02.
- RICHARDSON, A. J.; AMPT, E. S.; MEYBURG, A. H. *Survey Methods for Transport Planning*. 1a. ed. [S.l.]: Eucalyptus Press, 1995. Disponível em <<http://www.transporturveymethods.com.au>>. Acesso em: 08 de junho de 2019.
- RODRIGUES, M. A.; SORRATINE, J. A. *A Qualidade No Transporte Coletivo Urbano*. Faculdade de Engenharia Civil. Universidade Federal de Uberlândia, 2014. Disponível em: <https://www.researchgate.net/profile/Jose_Sorratini/publication/267387600_A_QUALIDADE_NO_TRANSPORTE_COLETIVO_URBANO/links/549807360cf2c5a7e342874f.pdf>. Acesso em: 21 de setembro de 2019.
- RORIZ JUNIOR, M. P.; MEDRANO, R. A.; UAI-FI: Um Método Baseado em Aprendizado de Máquina para Contagem Automática de Passageiros Utilizando Sinais Wi-Fi. 2018, Gramado, RS, Brazil: Congresso de Pesquisa e Ensino em Transportes, ANPET, 2018. p. 1048–1059.
- SILVA, J. A. B.; SILVA, S. Critérios De Qualidade Em Serviços De Transporte Público Urbano: Uma Contribuição Teórica. *Brazilian Journal of Production Engineering*, v. 4, p. 83–98, 2018.
- SOHRABI, C.; ALSAFI, Z.; O'NEILL, N.; KHAN, M.; KERWAN, A.; AL-JABIR, C.; Iosifidis, R. Agha, World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19), *International Journal of Surgery* 76 (February) (2020) 71–76. doi:10.1016/j.ijssu.2020.02.034.
- SOUSA JÚNIOR, P. T. G.; RORIZ JUNIOR, M. P.; RODRIGUES, K. C.; MEDRANO, R. M. A. Estimativa Da Viabilidade Da Reconstrução Da Matriz Origem Destino Através De Sinais Wi-Fi De Passageiros. 2019, Balneário Camboriú, SC, Brazil: Congresso de Pesquisa e Ensino em Transportes, ANPET, 2019.
- SOUSA, A. C. M. *Os parceiros do rio Bonito*. São Paulo: Duas Cidades, 1987.
- SOUZA, A. C. M. *Análise do Planejamento Urbano de um Espaço em Transformação: As cidades de Macapá e Santana na perspectiva do Desenvolvimento Local*. Dissertação (Mestrado Integrado em Desenvolvimento Regional). Universidade Federal do Amapá. Macapá, 2014.

- TACO, P. W. G., TEDESCO, G. M. T, GUERRA, H. O., et al. Reestruturação do Transporte Coletivo Urbano por Ônibus: Um Modelo Funcional.1997. Anais do XX Congresso de Pesquisa e Ensino em Transportes. Volume1, p. 457 – 468.
- TELECO. Informação em Telecomunicações. Número de celulares, 2018. Disponível em: <http://www.teleco.com.br/ncel.asp>. Acesso em: 11 de setembro de 2019.
- WASHINGTON, S., KARLAFTIS, M. G. & MANNERING, F. L. (2011) Statistical and econometric methods for transportation data analysis, Boca Raton, FL, CRC Press.
- WILHEIM, J. O bairro, unidade urbana. Projeto São Paulo: propostas para a melhoria da vida urbana. Rio de Janeiro: Paz e Terra, 1982.