

UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Leonardo Luís Röpke

**METODOLOGIA AUTOMATIZADA PARA DESCREVER A COBERTURA
DE SINAL MÓVEL EM RODOVIAS: ESTUDO DE CASO
CONSIDERANDO INCERTEZA DE DADOS**

Santa Maria, RS
2023

Leonardo Luís Röpke

**METODOLOGIA AUTOMATIZADA PARA DESCREVER A COBERTURA DE SINAL
MÓVEL EM RODOVIAS: ESTUDO DE CASO CONSIDERANDO INCERTEZA DE
DADOS**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação, Área de Concentração em Computação, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Ciência da Computação**. Defesa realizada por videoconferência.

Orientadora: Prof.^a Marcia Pasin

Santa Maria, RS
2023

Leonardo Luís Röpke

**METODOLOGIA AUTOMATIZADA PARA DESCREVER A COBERTURA DE SINAL
MÓVEL EM RODOVIAS: ESTUDO DE CASO CONSIDERANDO INCERTEZA DE
DADOS**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação, Área de Concentração em Computação, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Ciência da Computação**.

Aprovado em 24 de março de 2023:

**Marcia Pasin, Dra. (UFSM)
(Presidenta/Orientadora)**

João Luiz Dihl Comba, Dr. (UFRGS) (videoconferência)

João Carlos Damasceno Lima, Dr. (UFSM) (videoconferência)

Santa Maria, RS
2023

DEDICATÓRIA

À minha família. Sem eles, nada disso seria possível.

AGRADECIMENTOS

Aos meus pais, Arnildo e Eloide, por estarem sempre ao meu lado, por todo o incentivo, carinho e amor dados. Obrigado pela compreensão durante esse período. Obrigado por tudo!

Aos meus irmãos, Eduarda e Jean, pelo apoio, amizade, compreensão e momentos de descontração. Vocês são incríveis!

À Jaqueline Hübner, que como uma segunda mãe, uma amiga, acompanhou e me apoiou durante esta caminhada. Obrigado pelos conselhos, pelas conversas, pela paciência. Minha eterna gratidão!

À minha orientadora, professora Marcia Pasin, pelo apoio, confiança e ensinamentos transmitidos para o desenvolvimento deste trabalho. Minha gratidão e admiração!

Ao professor, Lucas Mello Schnorr, por todos os conhecimentos transmitidos e suporte dado ao longo deste trabalho. Muito obrigado!

Aos amigos, Carlos Dias, Jackson Mireski, Kelvin Bremm, Tiago Chiapinotto, pelos momentos de descontração, conselhos, bebedeiras e pela amizade.

À Universidade Federal de Santa Maria (UFSM) e ao Programa de Pós-Graduação em Ciência da Computação (PPGCC), pela oportunidade e pelos conhecimentos transmitidos. Meu muito obrigado!

Agradeço também a todos que de uma forma ou de outra contribuíram para a realização deste trabalho.

*O importante é não parar de questionar.
A curiosidade tem sua própria razão de
existir.*

(Albert Einstein)

RESUMO

METODOLOGIA AUTOMATIZADA PARA DESCREVER A COBERTURA DE SINAL MÓVEL EM RODOVIAS: ESTUDO DE CASO CONSIDERANDO INCERTEZA DE DADOS

AUTOR: Leonardo Luís Röpke
Orientadora: Marcia Pasin

Devido à grande distribuição geográfica ou à existência de ampla infraestrutura rodoviária, é desafiador garantir que operadoras de dados móveis forneçam um serviço adequado ao longo das rodovias. Uma metodologia automatizada para descrever e avaliar a cobertura de sinal de rede móvel em rodovias pode colaborar para que governos e operadoras planejem com eficiência ações para melhorar a qualidade do serviço provido pelas operadoras. Neste trabalho, é proposta uma metodologia automatizada para descrever e permitir a avaliação da cobertura de sinal de rede móvel em rodovias usando dados de rastreamento de veículos. A abordagem proposta agrega, usando um critério espaço-temporal, dados obtidos através de um sistema de rastreamento de veículos. A saída resultante deste processo de agregação é um mapa das rodovias com um esquema de cores gradiente, para indicar a disponibilidade do sinal e a confiança desta informação. Como o tráfego é dinâmico e os dados de tráfego não seguem uma distribuição uniforme, a medida de incerteza associada ao mapa indica a confiança da informação provida. Como estudo de caso, é usado um conjunto de dados relativos a viagens no estado do Rio Grande do Sul (RS), Brasil, coletados através de um serviço privado de rastreamento de veículos. A metodologia proposta permite analisar a disponibilidade do sinal nas rodovias, dentro da área geográfica delimitada pelos dados, e possibilita o uso de filtros por período ou nome de via, para análises mais específicas. Por último, com o uso da metodologia proposta e os dados disponíveis pôde-se concluir que a disponibilidade do sinal das rodovias do RS está acima do mínimo que estabelece o Projeto de Lei 1721/2019, que estabelece a obrigatoriedade de oferta de serviços de telefonia móvel e de Internet móvel ao longo de rodovias federais.

Palavras-chave: Disponibilidade do sinal. Internet das coisas. Mapa de cobertura. Rastreamento de veículos.

ABSTRACT

METHODOLOGY TO DESCRIBE THE MOBILE SIGNAL COVERAGE ON ROADS: CASE STUDY CONSIDERING DATA UNCERTAINTY

AUTHOR: Leonardo Luís Röpke

ADVISOR: Marcia Pasin

Due to the wide geographical distribution or the existence of extensive road infrastructure, ensuring that mobile operators provide adequate service along roads can be challenging. An automated methodology for describing and evaluating mobile network signal coverage on roads can help government and mobile operators effectively plan actions to improve network signal coverage. In this work, an automated methodology is proposed to describe and allow the evaluation of mobile network signal coverage on roads using vehicle tracking data. The proposed approach aggregates data samples obtained through an online vehicle tracking system using a spatio-temporal criterion. The resulting output of this aggregation process is a road map with a gradient color scheme to indicate the signal availability and reliability of this information. As the traffic is dynamic and the traffic data does not follow a uniform distribution, the uncertainty measure associated with the map indicates the reliability of the information provided. As a case study, we used a set of private travel data in southern Brazil collected through a vehicle tracking service. The proposed methodology makes it possible to analyze the availability of the signal on the roads, within the geographic area delimited by the data, and allows the use of filters by period or road name, for more specific analyses. Finally, with the use of the proposed methodology and the available data, it could be concluded that the availability of the signal on the roads in southern Brazil is above the minimum established by Draft Law 1721/2019, which establishes the obligation to offer telephone services mobile and mobile Internet services along federal roads.

Keywords: Coverage map. Internet of things. Signal availability. Vehicle tracking.

LISTA DE FIGURAS

FIGURA 1 – Etapas do processo de seleção de estudos.	20
FIGURA 2 – Número de Autores por País.	24
FIGURA 3 – Ano de publicação dos estudos.	25
FIGURA 4 – Frequência de palavras nos resumos dos artigos.	26
FIGURA 5 – Total de artigos por subárea.	27
FIGURA 6 – Esquema da metodologia automatizada proposta.	30
FIGURA 7 – Etapas do processo de filtragem dos dados.	33
FIGURA 8 – Plotagem com dados de Disponibilidade de sinal com 50 mil pontos (a) e 100 mil pontos (b).	36
FIGURA 9 – Exemplo de estrada traçada na abordagem de agregação de dados. ...	36
FIGURA 10 – Mapa da distribuição de dados no RS, com $\delta = 5$ km. Na legenda, N indica o número de dados por quadrante.	37
FIGURA 11 – Disponibilidade (Availability) e Incerteza (Uncertainty) do sinal das rodovias do estado do RS, utilizando todas as amostras no período total considerado.	39
FIGURA 12 – Aparência da ferramenta <i>web</i> em desenvolvimento.	40
FIGURA 13 – Disponibilidade e Incerteza do sinal 2G das rodovias do estado do RS.	42
FIGURA 14 – Comparação de mapas destacando Incerteza Baixa (a) e Incerteza Alta (b).	43
FIGURA 15 – Disponibilidade de sinal 2G no RS usando diferentes valores para δ , sendo $\delta = \{0.5, 5, 10 \text{ e } 15\}$ (km).	44
FIGURA 16 – Disponibilidade de sinal 2G na rodovia BR-386 usando diferentes valores para δ , sendo $\delta = \{0.5, 5, 10 \text{ e } 15\}$ (km).	44
FIGURA 17 – Análise geral da Incerteza no estado do RS.	46
FIGURA 18 – Disponibilidade ρ do sinal para a BR-386 com diferentes dispositivos (Dispositivo A, Dispositivo B e Dispositivo C) usando $\delta = 1$ km.	48
FIGURA 19 – Disponibilidade do sinal nas rodovias do estado do RS, em quatro anos, onde cada ponto em cada gráfico representa um mês.	50

LISTA DE TABELAS

TABELA 1 – Combinações usadas na busca.	19
TABELA 2 – Artigos selecionados e informações gerais.	21
TABELA 3 – Características do dataset.	34
TABELA 4 – Estudo do impacto da granularidade na visualização da Disponibilidade do sinal, variando o valor de δ	45
TABELA 5 – Estudo do impacto da granularidade na visualização da Disponibilidade do sinal, variando o valor de δ e considerando a Acurácia do experimento.	46
TABELA 6 – Quantidade de amostras de diferentes dispositivos - referente à BR-386.	47
TABELA 7 – Disponibilidade do sinal por rodovia do estado do RS no período 2017-2020.	51

LISTA DE ABREVIATURAS E SIGLAS

<i>2G</i>	Telefonia Móvel de Segunda Geração
<i>API</i>	<i>Application Programming Interface</i>
<i>CSV</i>	<i>Comma-separated Values</i>
<i>DBSCAN</i>	<i>Density-based Spatial Clustering of Applications With Noise</i>
<i>GPS</i>	<i>Global Positioning System</i>
<i>HCLUST</i>	<i>Hierarchical Clustering</i>
<i>HTTP</i>	<i>Hypertext Transfer Protocol</i>
<i>IoT</i>	<i>Internet of Things</i>
<i>JSON</i>	<i>JavaScript Object Notation</i>
<i>KM</i>	Quilômetros
<i>LAT</i>	Latitude
<i>LNG</i>	Longitude
<i>LTE</i>	Long Term Evolution
<i>MBB</i>	Banda Larga Móvel
<i>MONROE</i>	<i>Measuring Mobile Broadband Networks in Europe</i>
<i>QoE</i>	<i>Quality of Experience</i>
<i>QoS</i>	<i>Quality of Service</i>
<i>RS</i>	Estado do Rio Grande do Sul, Brasil
<i>RSL</i>	Revisão Sistemática da Literatura
<i>RSSI</i>	<i>Received Signal Strength Indication</i>
<i>TCP</i>	<i>Transmission Control Protocol</i>
<i>UDP</i>	<i>User Datagram Protocol</i>
<i>V2I</i>	<i>Vehicle-to-Infrastructure</i>

LISTA DE SÍMBOLOS

A	Amplitude
δ	Delta
ρ	Disponibilidade de Sinal
e	Erro percentual
u	Incerteza
y	Valor correto
V_{max}	Valor máximo
V_{min}	Valor mínimo
\hat{y}	Valor obtido

SUMÁRIO

1	INTRODUÇÃO	13
1.1	CONTEXTO E PROBLEMÁTICA	13
1.2	OBJETIVO	15
1.3	ESTRUTURA DO TEXTO	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	PALAVRAS-CHAVE	16
2.2	REVISÃO SISTEMÁTICA DA LITERATURA	18
2.2.1	Protocolo de Pesquisa	18
2.2.2	Seleção e Coleta de Dados	19
2.2.3	Trabalhos Selecionados	20
2.2.4	Revisão Meta-analítica	23
2.2.5	Limitações do Protocolo de Pesquisa da Revisão Sistemática	27
2.2.6	Resultados da Revisão Sistemática	27
3	METODOLOGIA	29
3.1	GERAÇÃO DE DADOS PARA O BANCO DE DADOS	30
3.2	GERAÇÃO DO <i>DATASET</i>	31
3.2.1	Características do Dataset	32
3.2.2	Configurações para os experimentos	34
3.3	MÓDULO DE PROCESSAMENTO	34
3.3.1	Agregação espaço-temporal	35
3.3.2	Tratamento de dados sobrepostos	35
3.3.3	Cálculo da Disponibilidade e Tratamento da Incerteza dos dados	37
3.4	FERRAMENTA DE VISUALIZAÇÃO DOS RESULTADOS	39
4	EXPERIMENTOS	41
4.1	DISPONIBILIDADE X INCERTEZA	41
4.2	ESTUDO DA GRANULARIDADE	43
4.3	ESTUDO DA INCERTEZA	46
4.4	ESTUDO COM FOCO NO DISPOSITIVO	47
4.5	ESTUDO COM FOCO NAS RODOVIAS	48
5	CONCLUSÃO	52
5.1	TRABALHOS FUTUROS	53
	REFERÊNCIAS BIBLIOGRÁFICAS	54

1 INTRODUÇÃO

O rastreamento de veículos permite maior confiabilidade, eficiência e segurança para os serviços de transporte rodoviário. Para operar corretamente, este serviço requer suporte adequado de infraestrutura de rede de comunicação ao longo das rodovias. Normalmente essa infraestrutura é fornecida pelas operadoras de telefonia móvel, cabendo ao governo verificar a qualidade do serviço prestado. A qualidade do sinal de telefonia móvel pode ser facilmente medida por um aparelho e avaliada por um humano. No entanto, em geral, a avaliação manual é passível de erros e de difícil aplicação em territórios com áreas abrangentes. Além disso, a avaliação manual está intimamente relacionada à visão do especialista em um curto período de tempo. Quando este serviço é aplicado a territórios de maior abrangência, como é o caso das medições ao longo de rodovias, pode envolver deslocamento de pessoal e equipamentos, o que gera custos adicionais. Acidentes geográficos, como montanhas, prédios e túneis, podem comprometer a eficiência do serviço de comunicação e, conseqüentemente, a resposta do serviço de avaliação da Disponibilidade do sinal. Assim, avaliar e descrever a Disponibilidade do serviço de comunicação em rodovias não é tarefa trivial.

1.1 CONTEXTO E PROBLEMÁTICA

Uma solução mais adequada seria ter uma metodologia automatizada para descrever e permitir avaliar o serviço da rede móvel em rodovias, onde a cobertura resultante pudesse ser observada em termos de espaço e tempo plotados em um mapa. Este trabalho contribui neste sentido. Como estudo de caso para demonstrar os resultados de nossa proposta, foi utilizado um conjunto de dados de propriedade privada relativo a viagens de rastreamento de veículos coletadas no sul do Brasil. Os dados consumidos neste trabalho foram gerados a partir de dispositivos que usam a tecnologia 2G. Entretanto, a metodologia aqui proposta pode ser aplicada a outras tecnologias.

A metodologia automatizada consome dados de sensores, que podem estar nos veículos/objetos em circulação nas rodovias, desde que possuam dispositivos embarcados com capacidade de comunicação e localização. Os dados coletados são armazenados em um banco de dados centralizado (em uma empresa) e são analisados e agregados a partir de uma perspectiva espaço-temporal. Como o esquema de agregação reduz a quantidade de dados de entrada, é definido o grau de Incerteza para descrever a qualidade dos dados resultantes que são plotados no mapa. Nesse contexto, as características do próprio conjunto de dados devem ser observadas. Em particular, são destacados os seguintes problemas que podem colaborar para adicionar ruído indesejado no mapa resultante:

- uma vez que uma característica intrínseca do tráfego é a falta de distribuição uniforme, existem estradas para as quais ocorre maior coleta de dados de comunicação e estradas com menos dados de comunicação disponíveis. Assim, a distribuição de dados no *dataset* carece de uniformidade geográfica e temporal, e esse ruído impacta no resultado do processo de agregação, principalmente nas áreas com quantidade reduzida de amostras. Então, deve ser calculado um grau de Incerteza no resultado. Assim, dada uma certa quantidade de dados com menor variação de valores (i.e., baixo desvio padrão), a chance de Incerteza dos dados é baixa; caso contrário, alta. Neste trabalho, representamos esses valores como um gradiente de cores no mapa, buscando prover informação com mais fidelidade;
- em especial, em algum dispositivo, o serviço de rastreamento pode ter sofrido falhas por motivos não especificados e os dados podem não ter sido obtidos adequadamente apesar da Disponibilidade do serviço de comunicação. Portanto, dispositivos que apresentem excesso de amostras negativas coletadas no intervalo de tempo e espaço considerados não devem ser incluídos no processo de agregação;
- nem todas as estradas ou segmentos rodoviários têm infraestrutura de comunicação disponível. Nesse caso, devem ser considerados que os veículos também podem coletar dados de forma *offline*, e que esses dados serão transferidos quando o equipamento do veículo reconectar a rede móvel;
- uma vez que pode haver muitos veículos em circulação nas rodovias e, potencialmente, muitos dados podem ser gerados, e esses dados precisam ser tratados adequadamente antes de serem plotados em um mapa. Assim, o mapa é dividido em quadrantes e dados coletados são agregados considerando os quadrantes. Como o esquema de agregação reduz a quantidade de dados, o grau de Incerteza descreve a qualidade dos quadrantes medidos no mapa resultante. Um esquema de cores gradiente é usado para representar a Incerteza. Este esquema de gradiente garante uma descrição dos dados sobrepostos.

Finalmente, nem todas as estradas têm veículos que passam por elas com recursos de rastreamento e comunicação. Esta é uma limitação deste trabalho: estradas que não possuem veículos rastreados passando por elas não são contabilizadas. Os dados devem estar disponíveis para que a cobertura do sinal seja estimada. De fato, é uma limitação intrínseca de qualquer metodologia automatizada baseada em dados obtidos pelo rastreamento de veículos.

1.2 OBJETIVO

O objetivo deste trabalho é propor uma metodologia automatizada para avaliar a qualidade da rede de sinal móvel em rodovias e outras vias pavimentadas. Para validar a metodologia proposta, é usado um conjunto de dados coletados por rastreadores disponibilizados por meio de um sistema de rastreamento usando suporte de comunicação veículo para infraestrutura – V2I. Aqui, a qualidade do sinal na rede móvel é avaliada em termos de espaço e tempo, com especial atenção à Incerteza dos dados.

Mais especificamente, a metodologia aqui proposta abrange:

- a construção de um índice de Disponibilidade de sinal nas rodovias estaduais e federais, que contemple a Incerteza dos dados;
- o mapeamento da Disponibilidade da rede de sinal móvel nas rodovias estaduais e federais do Sul do Brasil.

A metodologia automatizada permite avaliar a qualidade do serviço de rede móvel disponível nas rodovias estaduais e federais, e também permite que tanto o governo quanto as empresas responsáveis planejem ações adicionais para melhorar a qualidade da cobertura de rede móvel nas rodovias. Em especial, a metodologia automatizada pode ajudar a verificar se o Projeto de Lei 1721/2019¹ que estabelece a obrigatoriedade de oferta de serviços de telefonia móvel e de Internet móvel ao longo de rodovias federais está sendo obedecido.

1.3 ESTRUTURA DO TEXTO

O texto que segue está organizado em quatro capítulos. O capítulo 2 apresenta trabalhos relacionados e uma Revisão Sistemática da Literatura, apresentando um panorama geral das pesquisas na área. O capítulo 3 detalha e exemplifica a metodologia aqui proposta. O capítulo 4 detalha os experimentos realizados, usando dados relativos a viagens no estado do Rio Grande do Sul, Brasil. Por fim, o capítulo 5 apresenta as conclusões.

¹Câmara dos Deputados. Projeto de Lei PL 1721/2019 e seus apensados. Altera a Lei no. 9.472, de 16 de julho de 1997, que "dispõe sobre a organização dos serviços de telecomunicações, a criação e funcionamento de um órgão regulador e outros aspectos institucionais, nos termos da Emenda Constitucional no. 8, de 1995", para estabelecer a obrigatoriedade de oferta de serviços de telefonia móvel e de Internet móvel ao longo de rodovias federais. <https://www.camara.leg.br/propostas-legislativas/2195292>

2 FUNDAMENTAÇÃO TEÓRICA

As redes móveis vem crescendo cada vez mais em área de cobertura, número de usuários e em novos domínios da IoT que são capazes de apoiar. Estudar a cobertura das redes móveis vem se tornando um desafio cada vez mais explorado.

Este capítulo apresenta a Fundamentação Teórica, a partir de uma Revisão Sistemática da Literatura (RSL), mostrando um panorama geral tendo como foco a avaliação da cobertura de sinal móvel. A ideia neste capítulo é explorar a literatura – que apresenta diferentes métodos de avaliação da cobertura móvel – não somente em rodovias que é o assunto deste trabalho, mas explorando o assunto de forma mais abrangente incluindo métodos de avaliação da cobertura móvel em cidades, áreas rurais, campus universitários, em ferrovias e etc..

Essa revisão, permite a construção de um embasamento para a condução do trabalho aqui proposto, explorando técnicas e soluções existentes para a avaliação da cobertura móvel. A partir de palavras-chaves relacionadas com o escopo da pesquisa, os artigos foram selecionados e depois, foi realizada uma classificação dos artigos mais relevantes, buscando categorizá-los para contrastar com a pesquisa aqui proposta.

Inicialmente, para a RSL, foram escolhidas as seguintes palavras-chaves: Cobertura Móvel (*Mobile Coverage*), Banda Larga Móvel (*Mobile Network*), Mapas de Cobertura (*Coverage Maps*) e Qualidade de Serviço (*QoS*), que são brevemente comentadas no texto que segue. Estas, permitiram definir o escopo da busca de artigos, descrito na seção 2.2.1. Depois, os artigos foram selecionados, usando como base o atendimento aos critérios de inclusão descritos na seção 2.2.2, com foco em QoS e algoritmos de avaliação da qualidade do sinal. E por último é realizada uma análise meta-analítica dos estudos selecionados, analisando a filiação dos autores, quantidade de citações, ano de publicação, frequência de palavras em resumo e total de trabalhos por subárea.

2.1 PALAVRAS-CHAVE

Para esta RSL, foram selecionadas as seguintes palavras-chave:

- **Cobertura Móvel (Mobile Coverage):** reflete o quão bem uma rede abrange o espaço físico, estando diretamente relacionada com a qualidade de serviço (QoS) (LIU et al., 2010). Conforme Liu et al. (2012), a cobertura móvel é uma questão crítica para a implantação e desempenho de uma rede de sensores sem fio, representando a qualidade da conexão que a rede pode fornecer, e o quão bem uma região de interesse é monitorada por sensores e com que eficiência;

- **Banda Larga Móvel (Mobile Network):** permite uma distribuição mais rápida de informações e novas ideias. Geralmente, a banda larga móvel pode substituir a banda larga fixa, enquanto a banda larga fixa não substitui a banda larga móvel (STORK; CALANDRO; GAMAGE, 2014). Isto implica que qualquer pessoa com uma ligação de banda larga móvel pode obter informações onde quer que esteja, o que da origem a grandes aumentos na produtividade (EDQUIST et al., 2018). A dependência crescente da sociedade em redes de banda larga móvel tornou o fornecimento de cobertura móvel ubíqua a mais alta prioridade para as operadoras móveis, antes mesmo do desempenho e da qualidade da experiência do usuário - *QoE* (Fida et al., 2017);
- **Mapas de Cobertura (Coverage Maps):** as operadoras móveis contam com indicadores-chave de desempenho (KPIs) para entender o desempenho e a cobertura de sua rede, bem como de seus concorrentes. Os mapas de cobertura consistem em um grande número de medições de KPIs e são de importância crucial para as operadoras de celular, para gerenciamento de rede, manutenção, atualizações, operações e resolução de problemas. Porém são caros para obter e possivelmente limitados ou imprecisos em alguns locais (ALIMPERTIS et al., 2019). A precisão dos mapas de cobertura móvel depende da qualidade, densidade e padrão dos dados disponíveis (Fida et al., 2017). Embora os próprios provedores de internet móvel possam coletar medições na borda da rede, eles optam cada vez mais por terceirizar a coleta de dados para terceiros por vários motivos, incluindo: custo, responsabilidade relacionada a questões de privacidade da coleta de dados em dispositivos de usuário final e falta de acesso a redes concorrentes. Empresas de análise móvel coletam medições diretamente de dispositivos de usuários finais. Assim, a coleta de mapas de sinal em grande escala é alcançada, mas as medições podem ser esparsas no espaço e no tempo. De qualquer forma, a construção de mapas de intensidade de sinal é uma atividade custosa tanto para operadoras quanto para empresas de *crowdsourcing*. Além disso, é esperado que a implantação da rede 5G e o estabelecimento das cidades inteligentes e da IoT aumentarão a necessidade de medições precisas de desempenho em redes (ALIMPERTIS et al., 2019);
- **Qualidade de Serviço (QoS):** possibilita avaliar serviços de rede para que eles sejam melhores e mais previsíveis. A QoS em telefonia móvel é um assunto que tem recebido atenção substancial, principalmente por meio de métodos subjetivos de avaliação (ISABONA; SRIVASTAVA, 2017). Conforme Budiman e Wicaksono (2016), a QoS tem sido um dos principais tópicos de pesquisa e desenvolvimento em redes de pacotes por muitos anos em organizações bem conhecidas, como ITU-T e Cisco. A QoS pode ser descrita qualitativamente ou quantitativamente. A definição qualitativa de QoS está relacionada ao tratamento recebido por uma classe de

pacotes para alguma outra classe de pacotes, enquanto as definições quantitativas se baseiam em métricas como atraso ou perda de pacotes (ZHAO; OLSHEFSKI; SCHULZRINNE, 2000). Ainda, conforme Budiman e Wicaksono (2016), diferentes abordagens são utilizadas para medir a QoS, como analisar as expectativas e experiências dos usuários em relação aos serviços de telefonia móvel, a qualidade do atendimento ao cliente, ou comparando o desempenho de redes de diferentes operadoras móveis.

2.2 REVISÃO SISTEMÁTICA DA LITERATURA

No texto que segue, é apresentada uma revisão sistemática da literatura que aborda métodos de avaliação da qualidade do sinal móvel. No total, foram selecionados 23 estudos dos anos de 2010 a 2021, que atenderam aos critérios de inclusão desta revisão sistemática, apresentando técnicas e melhorias para a área.

Nesta revisão sistemática, são apresentados os resultados de estudos encontrados na literatura, a fim de explorar os métodos de avaliação da qualidade do sinal móvel. Desta forma, é apresentado o estado da arte por meio de uma abordagem meta-analítica, analisando autores, instituições e outros aspectos.

2.2.1 Protocolo de Pesquisa

Uma revisão sistemática da literatura é um processo baseado em informações necessárias para identificar, avaliar e selecionar estudos relevantes na área de interesse. É um processo controlado por um protocolo de pesquisa definido no início da revisão sistemática para fornecer consistência e robustez aos resultados (TRANFIELD; DENYER; SMART, 2003).

Como primeira etapa, foi elaborado um protocolo de pesquisa para seleção dos estudos desta revisão sistemática. Para serem selecionados (além dos critérios de inclusão), os estudos publicados:

- Devem ter relação com qualidade do sinal móvel ou
- Algoritmos de avaliação da qualidade do sinal móvel.

A busca foi realizada na biblioteca digital *Scopus*, por ser considerada uma biblioteca bastante completa em número de trabalhos disponibilizados e, complementada com a inserção de artigos manuais que se julgaram relevantes e não foram encontrados inicialmente. Primeiramente, foi escolhida uma palavra primária e, em seguida, combinada

com palavras relacionadas à área. As combinações de palavras usadas nas buscas são mostradas na Tabela 1.

Tabela 1 – Combinações usadas na busca.

("Broadband"OR "Quality of Service")
 AND ("Mobile Network"OR "Network Operators")
 AND ("Coverage map"OR "Geo"OR "Dataset"OR "GPS")

Fonte: Autor.

As buscas na base de dados foram configuradas de forma que os estudos resultantes apresentassem pelo menos uma das combinações de palavras no título, resumo ou palavras-chave. Para as buscas realizadas na base de dados Scopus, a busca avançada foi utilizada para selecionar apenas publicações de conferências, bem como estudos publicados em periódicos e revistas.

Para a seleção dos artigos, os estudos devem atender aos seguintes critérios de inclusão pré-estabelecidos:

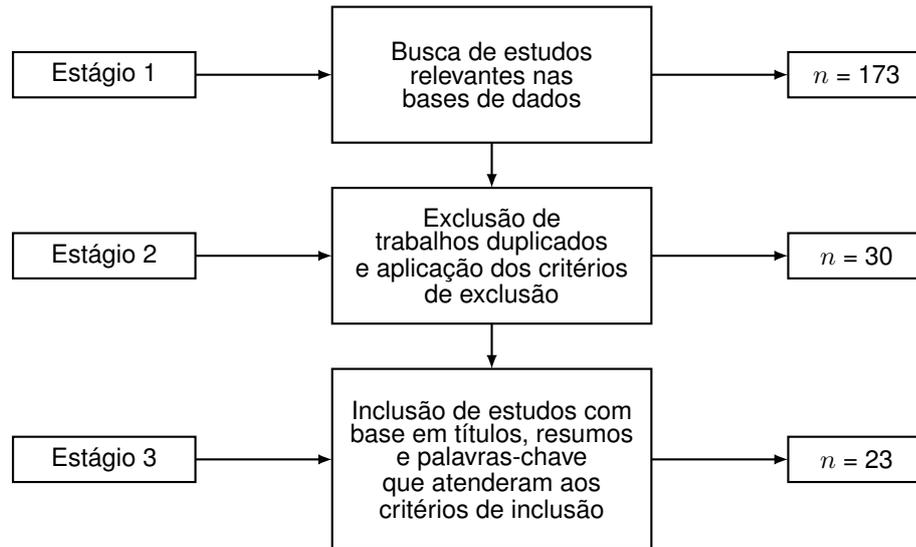
- Estudos que tratam da avaliação da qualidade do sinal móvel,
- Estudos que tratam da avaliação da qualidade do sinal sobre vias,
- Estudos que tratam de modelos de avaliação da qualidade do sinal,
- Estudos que tratam de redes móveis e geolocalização.

2.2.2 Seleção e Coleta de Dados

A primeira etapa da revisão sistemática foi a remoção de artigos duplicados encontrados e estudos que não atenderam aos critérios de inclusão iniciais. Posteriormente, foram lidos os títulos, palavras-chave e resumo de cada estudo. De acordo com os critérios de inclusão para esta revisão, os estudos foram incluídos quando atendiam a esses critérios no título, palavras-chave ou resumo. Após a seleção dos estudos, foram extraídas informações e características básicas de cada um, como título, autores, palavras-chave, resumo e afiliação dos autores.

Com as buscas realizadas, foram encontrados 173 artigos utilizando as combinações pré-estabelecidas. Destes, 2 artigos foram excluídos por estarem duplicados. Após foram removidos 141 artigos por não atenderem aos critérios de inclusão, restando 30 artigos. A partir de uma leitura mais aprofundada dos trabalhos, foram removidos mais 7 artigos. Desta forma, foram selecionados 23 artigos por atenderem aos critérios de inclusão desta revisão. A Figura 1 ilustra o protocolo de seleção dos estudos aplicado nos artigos.

Figura 1 – Etapas do processo de seleção de estudos.



Fonte: Autor.

2.2.3 Trabalhos Selecionados

Por meio da metodologia aplicada para a revisão sistemática, foram selecionados 23 artigos indicados na Tabela 2. Nos parágrafos seguintes, é apresentada uma síntese com o foco nesses trabalhos.

Os desafios em transformar dados em informação no contexto da mobilidade humana foram descritos anteriormente (ROSA; SILVA; ANALIDE, 2021). Devido a acidentes de terreno, clima, mudanças de rede, prédios, túneis, intensidade do sinal, falhas de dispositivos, etc., o processo de aquisição de dados possivelmente está sujeito a ruídos. Os dados podem estar ausentes, atrasados, distribuídos de forma heterogênea, o que impacta na qualidade do processo de aquisição dos dados. Além disso, a própria agregação de dados é um gargalo, em termos de qualidade da solução e tempo de processamento, e pode introduzir restrições importantes no resultado, pois um grupo de amostras precisa ser reduzido a um único valor. Assim, a agregação de dados aplicada para construir os mapas de cobertura precisa considerar a existência de ruído de dados.

O mapeamento da Disponibilidade e da qualidade do sinal móvel em rodovias continua sendo um problema em aberto. Em geral, trabalhos anteriores se concentram em descrever a existência ou a força do sinal em mapas de cobertura, considerando regiões em mapas. No texto que segue, trabalhos relacionados são brevemente comentados.

Com o foco em cobertura do sinal da rede celular, Lauridsen et al. (2017) compararam diferentes tecnologias IoT para ver qual oferece melhor cobertura de sinal. Experimentos foram realizados através da leitura da quantidade de sinal móvel em um misto de áreas

Tabela 2 – Artigos selecionados e informações gerais.

Artigo	Método de Inserção	Publicado em	Fator de Impacto	Citações
Location, localization, and localizability (LIU et al., 2010)	Manual	Journal of computer science and technology	2.58 (2020)	496
Coverage Comparison of GPRS, NB-IoT, LoRa, and SigFox in a 7800 km ² Area (LAURIDSEN et al., 2017)	Manual	IEEE 85th Vehicular Technology Conference (VTC Spring)	1.23	170
Speedtest-Like Measurements in 3G/4G Networks: The MONROE Experience (KHATOUNI et al., 2017)	Scopus	2017 29th International Teletraffic Congress (ITC 29)	1.07	39
Geospatial data aggregation and reduction in vehicular sensing applications: The case of road surface monitoring (Freschi et al., 2014)	Manual	2014 International Conference on Connected Vehicles and Expo (ICCVE)	-	37
Optimizing Network Connectivity for Mobile Health Technologies in sub-Saharan Africa (SIEDNER et al., 2012)	Manual	Public Library of Science San Francisco, USA	-	29
Impact of indoor-outdoor context on crowdsourcing based mobile coverage analysis (MARINA; RADU; BALAMPEKOS, 2015)	Scopus	Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges	-	22
Comprehensive mobile bandwidth traces from vehicular networks (BOKANI et al., 2016)	Scopus	Proceedings of the 7th International Conference on Multimedia Systems	0.00	19
City-Wide Signal Strength Maps: Prediction with Random Forests (ALIMPERTIS et al., 2019)	Manual	The World Wide Web Conference	6.91 (2020)	15
ZipWeave: Towards efficient and reliable measurement based mobile coverage maps (Fida et al., 2017)	Manual	IEEE INFOCOM 2017-IEEE Conference on Computer Communications	0.00	13
The good, the bad and the implications of profiling mobile broadband coverage (LUTU et al., 2016)	Manual	Computer Networks	4.47	11
Big Data-Driven Cellular Information Detection and Coverage Identification (WANG; XIE; LI K.AND AHMAD, 2019)	Manual	Sensors	4.14	8
The same, only different: Contrasting mobile operator behavior from crowdsourced dataset (KOUSIAS et al., 2017)	Scopus	2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)	1.12 (2020)	8
Impact of Device Diversity on Crowdsourced Mobile Coverage Maps (FIDA; MARINA, 2018)	Manual	14th International Conference on Network and Service Management (CNSM)	1.76	7
Unveiling radio resource utilization dynamics of mobile traffic through unsupervised learning (RAGO et al., 2019)	Scopus	2019 Network Traffic Measurement and Analysis Conference (TMA)	1.69	4
Connected Vehicles in Cellular Networks: Multi-Access Versus Single-Access Performance (ABRAHAMSSON et al., 2018)	Scopus	2018 Network Traffic Measurement and Analysis Conference (TMA)	2.39	4
Visualizing and analyzing georeferenced workloads of mobile networks (XAVIER; XAVIER; MARQUES-NETO, 2017)	Scopus	2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)	2.45 (2018)	4
Comparison of Signal Strengths of 2G/3G/4G services on a University Campus (ENGIZ; KURNAZ, 2016)	Manual	International Journal of Applied Mathematics Electronics and Computers	-	4
Crowdsourced network measurements in Germany: Mobile internet experience from end user perspective (SCHWIND et al., 2020)	Scopus	Broadband Coverage in Germany; 14. ITG Symposium	-	3
Multi-Criteria Coverage Map Construction Based on Adaptive Triangulation-Induced Interpolation for Cellular Networks (LIU et al., 2019)	Scopus	IEEE Access	5.56	3
Visualizing Mobile Coverage from Repetitive Measurements on Defined Trajectories (JARVIS et al., 2018)	Scopus	2018 Network Traffic Measurement and Analysis Conference (TMA)	2.39	2
Coverage is Not Binary: Quantifying Mobile Broadband Quality in Urban, Rural, and Tribal Contexts (ADARSH et al., 2021)	Manual	2021 International Conference on Computer Communications and Networks (ICCCN)	1.34 (2020)	1
Improving signal-strength aggregation for mobile crowdsourcing scenarios (MADARIAGA et al., 2021)	Manual	Sensors	3.98 (2020)	1
An Open Dataset of Operational Mobile Networks (KHATOUNI et al., 2020)	Scopus	Proceedings of the 18th ACM Symposium on Mobility Management and Wireless Access	-	0

Fonte: Autor.

urbanas e rurais da Dinamarca. Rago et al. (2019) buscaram identificar padrões de utilização de tráfego móvel através aprendizado não supervisionado. A análise apresentada é baseada em conjuntos de dados com informações do nível de sinal disponível. Adarsh et al. (2021) utilizaram métricas de Qualidade de Serviço (QoS) e Qualidade da Experiência (QoE) para medir o desempenho de redes *Long Term Evolution* (LTE) em regiões tribais, rurais e urbanas.

Com foco em avaliação de sinal de banda larga móvel, Khatouni et al. (2017) utilizaram a plataforma aberta MONROE para realizar testes de velocidade em redes 3G/4G. Os autores destacam os desafios da realização de medições em Redes de Banda Larga Móvel (MBB) devido a fatores como mobilidade dos usuários, deficiências físicas, diversidade de tecnologia, configurações de operadoras e outros. Siedner et al. (2012) analisou o sinal da rede móvel na zona rural de Uganda para calcular a frequência de falhas e compor mapas de cobertura. Marina, Radu e Balampekos (2015) estudaram o impacto do contexto ambiente interno/externo do usuário na medição da qualidade do sinal, utilizando a força do sinal como métrica e usando um grande conjunto de dados obtidos através de *crowdsourcing* os autores demonstram que o contexto interior/exterior tem um impacto significativo nos resultados das medições. Wang, Xie e Li K. and Ahmad (2019) utilizaram a intensidade do sinal recebido para descobrir a localização de antenas instaladas em uma cidade, utilizando um mapa dividido na forma de grade utilizando quadrantes de 50m por 50m, onde em cada grade o método de centroide ponderado modificado é empregado para realizar a agregação de dados e obter a localização.

Com foco em mapas de cobertura, Liu et al. (2010) apresentam em sua pesquisa uma visão geral sobre os problemas de localização utilizando redes sem fio e de sensores. Os autores fornecem uma visão geral das abordagens de localização existentes e discutem os desafios de pesquisa implícitos. Freschi et al. (2014) apresentam uma abordagem de redução e agregação de dados para uso em uma aplicação de *crowdsensing* de sensoriamento veicular destinada a monitorar a rugosidade de superfícies de estradas. A metodologia proposta pelos autores permitiu reduzir o tamanho dos dados em aplicações geoespaciais de forma bastante eficiente.

Bokani et al. (2016) avaliaram um conjunto de dados com repetidas medições realizadas em redes 3G e 4G sob condições de condução veicular sobre rotas pré-definidas. Em sua análise, os autores encontram implicações importantes sobre as flutuações da largura de banda que podem ocorrer em um ambiente urbano real. Alimpertis et al. (2019) desenvolveram uma estrutura de previsão baseada em florestas aleatórias para a criação de mapas de cobertura com dados limitados. Em Fida et al. (2017), apresentaram a construção de mapas de cobertura com a Indicação de Intensidade do Sinal Recebido (RSSI) usando um conjunto de dados com tamanho reduzido, para isso são utilizados algoritmos de clusterização como DBSCAN e HCLUST.

Em Lutu et al. (2016), a cobertura móvel nas ferrovias na Noruega foi avaliada em

relação à perspectiva do usuário, incluindo desempenho de download HTTP e taxa de perda de pacotes, foi construída uma grade de cobertura dividindo as rotas em quadrantes e analisando a cobertura de duas operadoras distintas em cada segmento. Fida e Marina (2018) realizaram um estudo sobre o impacto do uso de dispositivos heterogêneos no resultado de mapas de Disponibilidade do sinal. Os experimentos indicam que os mapas de cobertura móvel baseados em medições de diversos dispositivos são bastante confiáveis. Abrahamsson et al. (2018) usaram dados de banda larga móvel da plataforma MONROE para realizar um estudo de medição multiacesso em veículos. Uma requisição analisa a Disponibilidade de três operadoras e, para cada transação, escolhe a operadora com melhor tecnologia de acesso e melhor qualidade de sinal. Os experimentos do trabalho indicam que pode-se melhorar significativamente a Disponibilidade do serviço e os tempos de transação em comparação com o uso de serviços de operadoras individuais. Xavier, Xavier e Marques-Neto (2017) apresentaram uma abordagem para visualização e análise de dados georreferenciados de uso de redes móveis. Engiz e Kurnaz (2016) compararam a qualidade do sinal móvel 2G, 3G e 4G de três operadoras dentro de um campus universitário da Turquia. Os resultados do trabalho indicam que as operadoras não entregam o que exigem os órgãos regulatórios. Liu et al. (2019) propuseram um método de triangulação adaptativa para dividir as regiões de interesse em triângulos. Os resultados dos experimentos evidenciam que o método apresentou um bom desempenho na construção de mapas de cobertura e tem um custo significativamente menor e maior eficiência do que algoritmos tradicionais. Jarvis et al. (2018) desenvolvem um método de avaliação da qualidade do sinal em estradas através da divisão do mapa em segmentos de rota, tornando a visualização dos dados mais eficiente. Madariaga et al. (2021) propuseram a melhoria da agregação da intensidade de sinal em cenários de *crowdsourcing*, evitando problemas comuns relacionados ao manuseio incorreto de valores de sinal em escala logarítmica e pela proposta de um novo método de agregação baseado em interpolação. Khatouni et al. (2020) utilizaram dados de rede móvel coletados pela plataforma MONROE para analisar fluxos TCP e UDP e demonstrar o desempenho de serviços da Web.

Kousias et al. (2017) e Schwind et al. (2020) utilizaram dados provenientes de *crowdsourcing* para avaliar a experiência do usuário utilizando internet móvel, utilizando-se de métricas de desempenho, como taxa de *download/upload*, latência e intensidade do sinal.

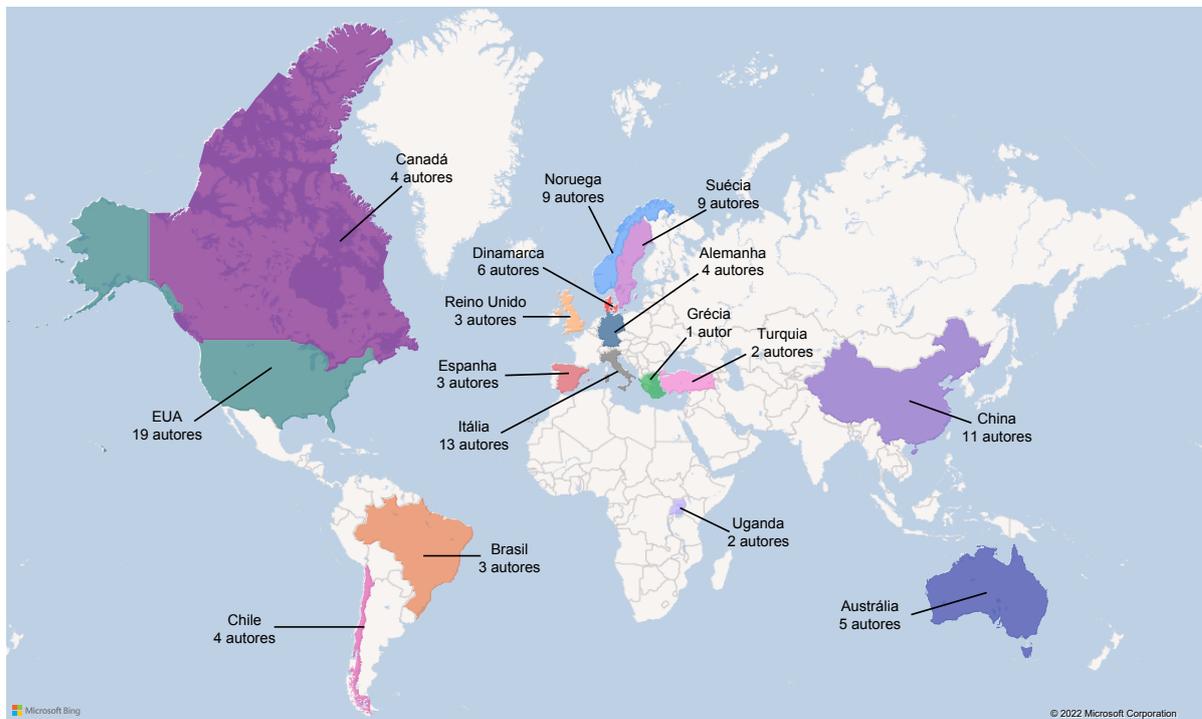
2.2.4 Revisão Meta-analítica

Após a aplicação do protocolo de pesquisa e dos critérios de inclusão/exclusão nos artigos, foi realizada uma abordagem meta-analítica nos estudos resultantes. A metanálise levou em consideração vários aspectos, que são apresentados nas subseções a seguir.

Filiações dos Autores

Primeiramente, foi analisado o local de filiação de todos os autores, ilustrado na Figura 2. Todas as filiações foram analisadas, uma vez que diversos estudos contam com autores de diferentes instituições e países. Por esse motivo, a soma das filiações na Figura 2 supera o total de artigos selecionados na revisão. A partir dos dados obtidos, observa-se a grande distribuição de países que realizam pesquisas relacionadas à área, com atenção especial aos países da Europa, o que pode ser explicado pela Disponibilidade de dados relacionados à banda larga móvel obtidos através da plataforma MONROE (MONROE, 2020). Percebe-se também uma alta colaboração nos trabalhos desenvolvidos, com 23 trabalhos selecionados obteve-se um total de 98 autores diferentes, e uma média de 5,13 autores por trabalho desenvolvido.

Figura 2 – Número de Autores por País.



Fonte: Autor.

Citações e Locais de Publicação

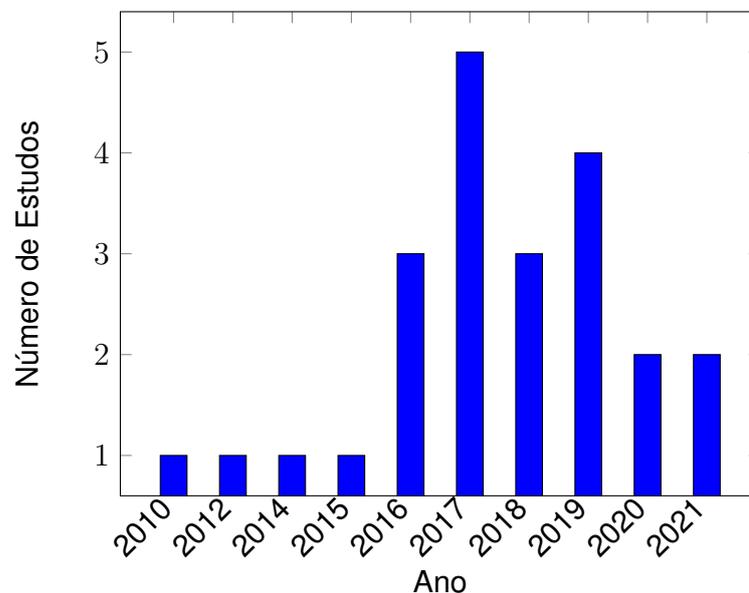
A Tabela 2 apresenta o número total de citações que cada estudo obteve. Pode-se observar que embora os estudos sejam relativamente recentes, apresentam um total de citações relativamente alto, o que representa uma área de grande abrangência e crescimento.

Também pode ser observado na Tabela 2, informações que foram coletadas das publicações. Dos 23 artigos selecionados, 7 estudos foram publicados em periódicos e há uma grande diversidade de veículos de publicação. É importante destacar que todos os periódicos possuem um fator de impacto de acordo com seus respectivos sites e, considerando a abrangência dos periódicos publicados, essa área abre diversas possibilidades de pesquisa. Dentre os periódicos, é possível observar que a maioria possui escopo voltado para a computação, enquanto outros são de escopos mais específicos, como o *Vehicular Technology Conference*, *International Teletraffic Congress*, *The World Wide Web Conference* e a *Sensors*. Percebe-se que boa parte dos estudos foi publicada em periódicos com alto fator de impacto, indicando a relevância e qualidade dos estudos e da área.

Ano de Publicação

A Figura 3 ilustra o número de estudos selecionados por data de publicação. Apesar das redes móveis existirem a anos, a preocupação em avaliar a qualidade do sinal móvel sob a perspectiva do usuário é relativamente recente. Os estudos foram publicados em no máximo 12 anos, mostrando que é uma área de pesquisa relativamente nova e ainda pouco explorada.

Figura 3 – Ano de publicação dos estudos.



Fonte: Autor.

Frequência de Palavras em Resumos

A Figura 4 mostra a frequência de palavras nos resumos de todos os estudos selecionados. Palavras como *network* e *mobile* aparecem em maior destaque pois são basicamente o objeto deste estudo. Em um destaque um pouco menor observam-se palavras como *measurement*, *coverage* e *performance* indicando os objetivos mais buscados nos trabalhos selecionados. Embora em menor quantidade, percebe-se também o uso de termos como *vehicles*, *routes*, *locations*, *areas*, mostrando algumas das necessidades que os autores buscam suprir com sua pesquisa na área.

Figura 4 – Frequência de palavras nos resumos dos artigos.



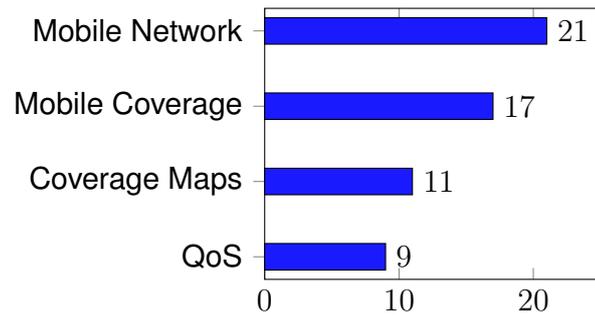
Fonte: Autor.

Em suma há um grande uso de palavras relacionadas às diferentes formas de localização, representando um dos grandes objetivos das pesquisas, com muitas questões a serem respondidas, principalmente no que diz respeito à avaliação da qualidade do sinal móvel.

Total de Trabalhos por Subárea

Como esta área está em expansão e possui focos de estudos menores dentro da grande área, foi realizada a classificação dos estudos dentro de subáreas de acordo com as palavras-chave deste artigo. A Figura 5 ilustra cada uma das subáreas com a respectiva quantidade de trabalhos onde elas são mencionadas.

Figura 5 – Total de artigos por subárea.



Fonte: Autor.

2.2.5 Limitações do Protocolo de Pesquisa da Revisão Sistemática

O protocolo de pesquisa definido foi elaborado para abranger o maior número possível de estudos relacionados à área. A base de dados Scopus foi definida por seu vasto conjunto de trabalhos, desta forma a pesquisa foi limitada a somente uma base de dados e complementada com alguns artigos que eram de conhecimento prévio.

Devido ao fato da análise da qualidade do sinal móvel não ser uma área claramente definida, mas sim um objeto de estudo que envolve uma composição de áreas, teve que ser realizada uma busca que fosse bastante ampla no sentido de obter diversos estudos e que tivesse também, critérios de inclusão de artigos bastante restritos, para que somente aqueles trabalhos relacionados a localização, mapas de cobertura, qualidade do serviço, métricas e/ou modelos de visualização dos dados fossem selecionados. Com isso, dos 173 artigos retornados na busca, apenas 23 foram selecionados para compor esta revisão sistemática, indicando um número baixo de pesquisas relevantes que abrangessem efetivamente as áreas de interesse e os critérios pré-definidos.

2.2.6 Resultados da Revisão Sistemática

A análise da qualidade do sinal móvel na atualidade é de extrema importância para os usuários e empresas que utilizam destas redes e consomem cada vez mais em seu dia a dia os benefícios propostos por elas. Além disso, com o desenvolvimento da IoT vê-se uma dependência cada vez maior da tecnologia móvel utilizada por indústria, comércio, logística, agrícola e de outros ramos da economia, desta forma medir e garantir que os serviços estejam sendo entregues de maneira satisfatória em diferentes áreas é muito importante para o desenvolvimento destas tecnologias e da economia de forma geral.

Nesse contexto, foi apresentada uma abordagem meta-analítica, onde após as seleções primária e secundária, 23 estudos completos foram selecionados. Com a análise do título, palavras-chave e resumo dos estudos selecionados, foram destacados os principais estudos na área, locais de publicação, países e os termos mais recorrentes.

Esta RSL sobre métodos de avaliação da qualidade do sinal móvel apresentou uma visão geral sobre os estudos na área, abrangendo não só a avaliação da qualidade do sinal móvel como um todo, mas também buscando explorar um pouco mais a fundo as diferentes áreas que são englobadas neste campo de estudo.

Sendo assim, a RSL contribui com esta pesquisa apresentando um panorama geral das pesquisas na área de avaliação da cobertura do sinal móvel. Através dos trabalhos analisados, percebe-se as diferentes formas que podem ser utilizadas para as análises, levando em conta os diferentes problemas, e as diferentes respostas que se deseja obter. Como em (JARVIS et al., 2018), onde os autores buscaram traçar o perfil de cobertura móvel ao longo de rotas de trens, através da segmentação das rotas em pequenos trechos. Ou em (Fida et al., 2017), em que os autores agregam dados por regiões de similaridade, constituindo assim análises mais regionais da disponibilidade do sinal. Além disso, destacam-se os trabalhos relacionados a plataforma MONROE, pelas contribuições com os diferentes métodos de análises propostos e pelas análises baseadas no conjunto de dados disponível.

Após a análise dos trabalhos encontrados na Revisão Sistemática, percebe-se que a metodologia proposta neste trabalho, diferencia-se das demais por propor uma metodologia automatizada que descreve a existência da cobertura da rede móvel em estradas e a Incerteza (ruído) desta informação. Os trabalhos analisados não descrevem a Incerteza dos resultados, além de utilizarem outros tipos de dados (alguns mais complexos) para as análises de Disponibilidade. Além disso, para avaliar a metodologia proposta, é utilizado neste trabalho um conjunto de dados real, obtido de dispositivos IoT de usuários finais.

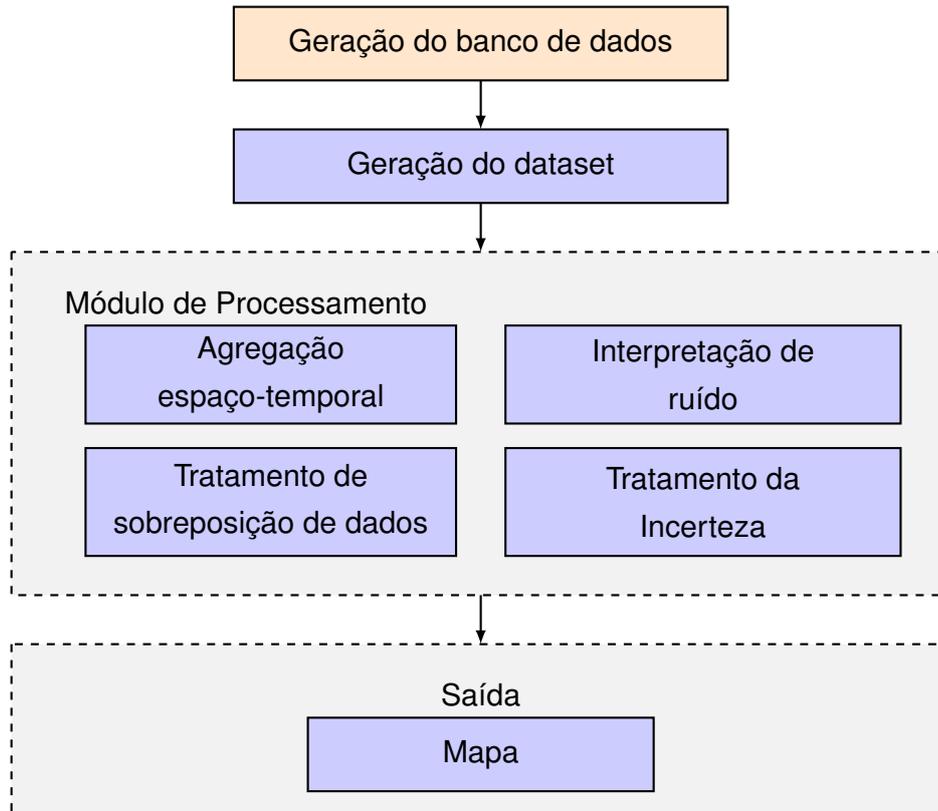
3 METODOLOGIA

Esta seção descreve com mais detalhes a metodologia automatizada desenvolvida para avaliar a qualidade do sinal da infraestrutura de rede móvel ao longo das rodovias, com informações sobre o grau de Incerteza. A metodologia consome dados obtidos da infraestrutura de comunicação móvel, que são coletados usando um conjunto de veículos. Durante a viagem de cada veículo, o dispositivo IoT de rastreamento instalado coleta dados que são enviados e armazenados em uma central. Portanto, os veículos participantes possuem dispositivos de rastreamento e recursos de comunicação disponíveis.

Em cada veículo, um dispositivo rastreador registra o veículo *id*, a existência ou não do sinal, e as coordenadas de latitude (LAT) e longitude (LNG) usando um suporte de GPS. Os dados coletados são enviados através de comunicação V2I para um serviço central, que armazena os dados em um banco de dados. Esses dados são então filtrados e usados como *dataset* para gerar um mapa da região rastreada com a indicação da Disponibilidade do sinal de rede móvel.

Basicamente, o serviço da unidade central filtra os dados armazenados, completa os dados faltantes usando o suporte de serviços abertos de terceiros, agrega e apresenta a informação final de forma adequada em um mapa, no contexto do espaço e do tempo. Para permitir a construção do mapa, uma série de filtros é aplicada, onde cada filtro é implementado por um módulo de processamento diferente. Um esquema com a arquitetura da metodologia proposta, e seus módulos, é apresentado na Figura 6. A seguir, são apresentados mais detalhes sobre os módulos que fazem parte da arquitetura.

Figura 6 – Esquema da metodologia automatizada proposta.



Fonte: Autor.

3.1 GERAÇÃO DE DADOS PARA O BANCO DE DADOS

Na metodologia proposta, dados são coletados por dispositivos IoT de rastreamento, instalados nos veículos participantes. Esses dados são enviados para o banco de dados centralizado por meio de comunicação V2I seguindo intervalos Δt durante a comunicação *online*. Se a conexão entre o dispositivo e o banco de dados centralizado não pode ser estabelecida, e a mensagem não pode ser enviada imediatamente após ser gerada, o dispositivo armazena os dados em uma memória local e os transmite, quando a conexão retorna. Neste caso, o sistema considera que os dados foram gerados de forma *offline*.

Ao receber as mensagens emitidas pelos dispositivos que estavam *online* ou *offline*, a unidade central consolida os dados em uma tabela com as seguintes colunas:

- *id* do veículo;
- valores de LAT e LNG coletados via GPS;

- *online*, que é um atributo binário para descrever se os dados foram transmitidos no mesmo momento em que foram gerados (*online*, sem atraso) ou se os dados foram transmitidos *offline*, ou seja, a mensagem tem um atraso de envio e foi enviada quando a conexão dispositivo–operadora móvel foi restabelecida. Valores possíveis são 1, que é usado em caso de comunicação *online* e 0 caso contrário;
- *date + time*, que é um timestamp atribuído pelo dispositivo emissor da mensagem no momento da sua criação.

Desde que os dados possam ser coletados ao longo dos anos, o banco de dados resultante pode ser usado como entrada para a metodologia avaliar a Disponibilidade do sinal da rede móvel nas rodovias. Aqui, o banco de dados foi gerado por uma empresa privada de rastreamento, e foi utilizado $\Delta t = 60s$, totalizando (inicialmente) 225 milhões de registros. Alternativamente, os dados poderiam ter sido coletados por meio de um aplicativo de *crowdsourcing*.

3.2 GERAÇÃO DO DATASET

A próxima etapa consiste em filtrar e organizar os dados armazenados no banco de dados para formar um *dataset*.

Inicialmente, os dados são filtrados, em uma abordagem de duas etapas:

- definição do período de dados: nesta fase, é definido um período de tempo durante o qual a Disponibilidade do sinal é avaliada; os dados que não estão nesse período são removidos;
- anonimização dos dados de identificação: para manter o anonimato dos veículos rastreados, o campo do banco de dados *id_object* é mascarado.

A seguir, os registros armazenados no banco de dados são completados com informações adicionais e um período de tempo (ou seja, um conjunto de anos) é selecionado para avaliar a Disponibilidade do sinal da rede móvel nas rodovias.

Dados ausentes no banco de dados são incluídos usando o suporte de serviços abertos de terceiros. Para cada registro armazenado no banco de dados, os dados de estradas ausentes são inseridos usando dados de latitude e longitude. Dados rodoviários são obtidos de acordo com as informações e classificação disponíveis via API Overpass¹, e incluem:

- *type*: tipo da via, que pode ser *motorway*, *motorway_link*, *primary*, *primary_link*, *secondary*, *secondary_link*, *tertiary*, *tertiary_link*, *trunk* e *trunk_link*;

¹<http://overpass-api.de>

- *name*: nome da via, como apresentado pela API.

Para completar os dados ausentes, os seguintes passos são executados para cada item no banco de dados, considerando os valores de LAT e LNG:

- aplicar a API Overpass com a seguinte consulta: "[out: json]; way (around: 25, LAT, LNG); out tags;"; que busca a partir da latitude/longitude informados e um raio de 25 metros, todas as vias disponíveis e grava em um arquivo com formato JSON;
- usando o arquivo JSON, extrair os elementos *name* e/ou *ref* para obter o nome da via e o elemento *highway* para obter o tipo da via. Se o processamento no arquivo JSON retornar mais de uma instância, o nível mais alto de via será utilizado.

Após este tratamento, são aplicados uma série de métodos de limpeza nos dados armazenados no banco de dados, nomeadamente:

- seleção de dados que pertencem apenas a uma determinada região – uma região em um mapa é selecionada para avaliar a qualidade do sinal. Os dados que não estão relacionados a esta região são removidos. Neste trabalho, foram utilizados dados relativos ao estado do Rio Grande do Sul (RS) – Brasil;
- seleção de dados que pertencem apenas às vias principais, descartando ruas e avenidas, dado que o estudo aqui apresentado foca em rodovias;
- retirada de dados de veículos que se comunicaram apenas em modo *offline* – este comportamento indica que o dispositivo instalado no veículo estava com algum defeito em seu módulo transmissor durante o período de tempo avaliado;
- remoção de pontos emissores que ficaram *offline* por mais de 48 horas. Quando um dispositivo para de funcionar, em até 48 horas um técnico é avisado. O equipamento é então retirado do veículo e levado ao laboratório para investigação das causas do problema e seus dados são armazenados no banco de dados como *offline*. Nesse caso, os dados *offline* relacionados a esse dispositivo, neste período de tempo *t*, são selecionados e removidos do banco de dados.

Por fim, os dados são armazenados como arquivo CSV, o que resulta em um conjunto de dados a ser utilizado para avaliar a Disponibilidade do sinal rede móvel nas rodovias.

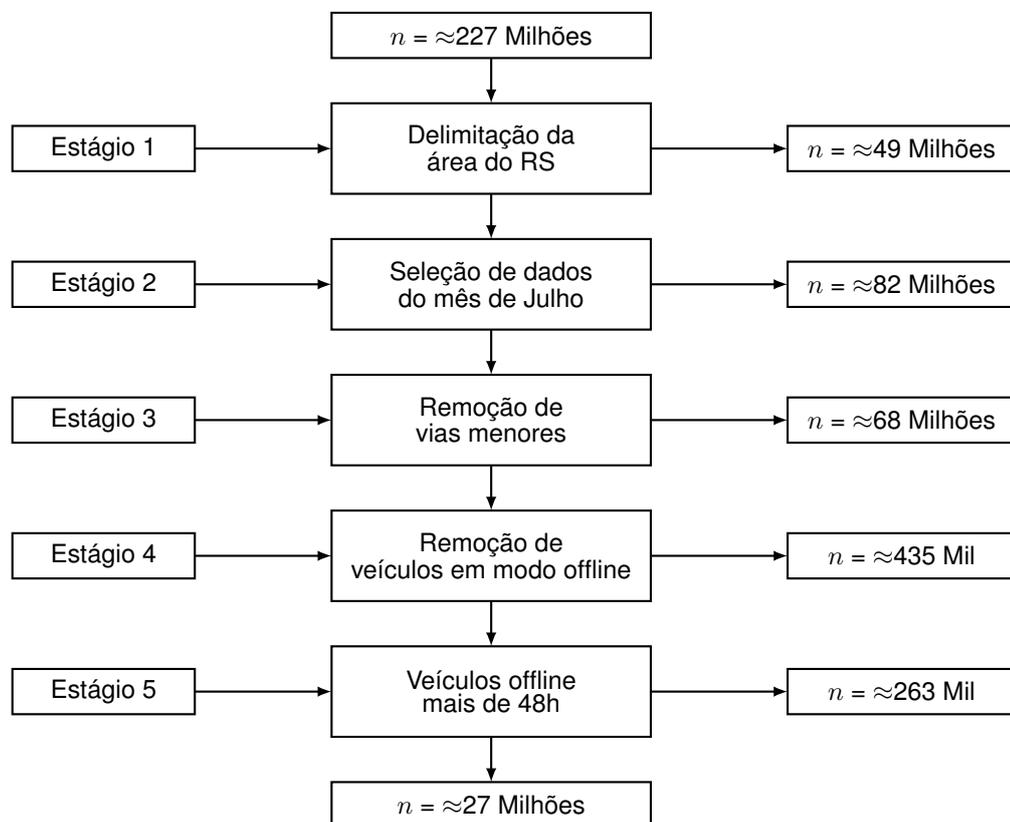
3.2.1 Características do Dataset

Os registros do banco de dados foram coletados durante o período de janeiro de 2017 a dezembro de 2020. No total, são 225 milhões de amostras armazenadas (antes

da filtragem), usando $\Delta t = 60s$ como intervalo de tempo entre as amostras coletadas. Este período temporal permite evitar o processamento dispendioso de dados, ao mesmo tempo que oferece uma quantidade volumosa de dados para validar a metodologia. O banco de dados original, antes da filtragem dos dados possuía ≈ 40 milhões de registros armazenados. Já a entrada resultante para os experimentos consiste em um conjunto de dados no formato CSV com ≈ 27 milhões de registros armazenados (≈ 3 GBytes).

A Figura 7 descreve todas as etapas do processo de filtragem dos dados, bem como a quantidade de dados removidos após cada uma das etapas. Um resumo das características do conjunto de dados é mostrado na Tabela 3. As colunas na tabela indicam o *Ano*, o *Total de Dispositivos* usados para coletar os dados e o número *Total de Amostras* coletadas naquele ano.

Figura 7 – Etapas do processo de filtragem dos dados.



Fonte: Autor.

No total, a empresa provedora dos dados disponibilizou dados de 861 dispositivos de rastreamento.

Tabela 3 – Características do dataset.

Ano	Total de Dispositivos	Total de Amostras
2017	381	4.998.386
2018	495	6.761.073
2019	564	7.375.570
2020	455	7.978.125

Fonte: Autor.

3.2.2 Configurações para os experimentos

Nesta seção, são descritos os experimentos realizados no contexto deste trabalho. Para validar a metodologia, foi implementada uma aplicação em nível de servidor para executar um conjunto de experimentos usando o *dataset* privado como entrada. A aplicação utiliza da linguagem R v.3.6.3, e da IDE RStudio² v.1.2.5033.

Os principais pacotes R utilizados neste trabalho são *geosphere* para o tratamento de dados envolvendo geolocalização, *ggmap* para plotagem de dados em mapas e *shiny* para geração de páginas *Web* que permitem a visualização das informações. O servidor de aplicação está atualmente hospedado em um computador comum. Na borda, os modelos de dispositivos de rastreamento são ST340, ST300HD e ST310U do fabricante Suntech. Os dispositivos de rastreamento utilizam o chip MT2503 da MediaTek como padrão para transmissão de dados, que possui um modem 2G integrado.

A aplicação desenvolvida plota dados em mapas gerados por diferentes dispositivos de rastreamento de uma empresa privada de rastreamento. A empresa de rastreamento tem clientes concentrados principalmente na região noroeste do estado do RS, no Sul do Brasil. Os clientes da empresa atuam nas mais diversas áreas: transportadores de produtos perecíveis, grãos, máquinas, equipes comerciais, assistência técnica, fornecimento de energia, setor público, máquinas pesadas, coleta de resíduos, etc.

3.3 MÓDULO DE PROCESSAMENTO

Após filtrar e organizar o conjunto de dados, o próximo passo é realizado pelo módulo de processamento, que gera o mapa de cobertura. Para isso, é usado o suporte da linguagem R³. O módulo de processamento é dividido em quatro partes que são descritas no texto a seguir.

²<https://rstudio.com>

³<https://www.r-project.org>

3.3.1 Agregação espaço-temporal

A agregação de dados considera que os dados do *dataset* pertencem a diferentes veículos, diferentes estradas e foram coletados em diferentes períodos de tempo. Para resolver isso, os dados são agregados no espaço e no tempo. Para agregação espacial, basicamente o mapa é dividido em quadrantes e, então, os pontos são agregados nos quadrantes correspondentes. Este procedimento é melhor descrito neste texto juntamente com o tratamento de pontos sobrepostos, no texto que segue. Para a agregação temporal, é avaliada a Disponibilidade do sinal ao longo dos anos, para verificar se houve melhora no provimento deste serviço. Para verificar se ocorreu a melhora do serviço ao longo dos anos, a informação é mostrada no mapa por ano.

3.3.2 Tratamento de dados sobrepostos

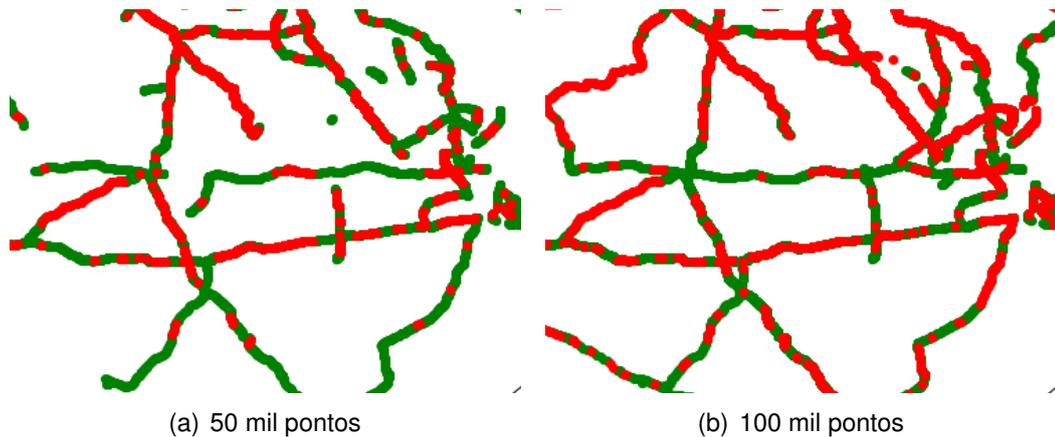
Dado um conjunto de pontos a ser plotado em um mapa, o ponto d_i pode sobrepor o ponto d_{i-1} anteriormente plotado, simplesmente porque d_i foi o último ponto a ser plotado no mapa. Isso pode dar a impressão de que o sinal está frequentemente disponível quando na verdade não está, ou vice-versa. O último ponto a ser plotado pode ocultar os valores previamente plotados.

A Figura 8 apresenta dois exemplos de mapas resultantes para o *dataset*, considerando pontos *online* e *offline*, em um recorte do estado do RS, sem usar esquema de agregação. Ambos os mapas descrevem dados para o mesmo período. Para a plotagem dos mapas, foram usadas diferentes quantidades de pontos. O gráfico representado na Figura 8 (a) usa 50 mil pontos enquanto que na Figura 8 (b) o número de pontos é o dobro. Uma observação nos mapas apresentados nesta figura pode resultar em uma interpretação incorreta dos dados. As imagens da Figura 8 não são idênticas, apesar dos dados em ambos os mapas representarem exatamente o mesmo período de tempo.

Para corrigir esse problema de sobreposição, dados devem ser agregados antes de serem plotados no mapa, e um grau de Incerteza deve ser calculado para medir o ruído gerado pela agregação de dados. Por exemplo, em um conjunto de dados com n amostras, um grupo de n_i amostras, com $n_i < n$, pode ser usado para calcular a média de amostras *online*, que pode ser plotada para representar a Disponibilidade do sinal. Além disso, dado que a quantidade de amostras agregadas pode ter um desvio padrão variável, por exemplo, muitos pontos em vermelho e poucos em verde resultando em um ponto vermelho e todos os pontos em vermelho também resultando em um ponto vermelho, um grau de Incerteza deve ser definido para esses dados plotados.

Neste trabalho, como o mapa possui duas dimensões, é realizado um esquema de discretização de coordenadas em quadrantes com área configurável e então, com base nessas coordenadas, é realizada uma agregação espacial para unir todos os pontos que

Figura 8 – Plotagem com dados de Disponibilidade de sinal com 50 mil pontos (a) e 100 mil pontos (b).

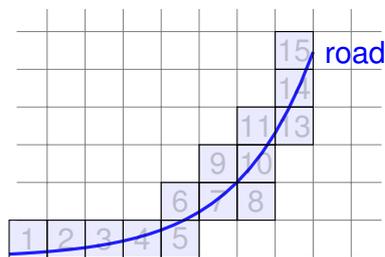


Fonte: Autor.

cabem em determinado um quadrante, delimitado por um par de coordenadas latitude e longitude. Uma grade foi gerada usando uma granularidade pré-definida δ e, assim, a agregação foi aplicada em termos de quadrantes do mapa.

A Figura 9 mostra um exemplo genérico para a aplicação do esquema de divisão de uma estrada em quadrantes de mesmo tamanho, dado um valor δ . Mais especificamente, neste exemplo, é realizada a identificação (1-15) e associação dos pontos amostrais à grade, partindo do quadrante do canto sudoeste (1) até o último quadrante (15). Os pontos são associados um a um ao quadrante correspondente. A seguir, os pontos são plotados na grade inicialmente gerada, reunindo os pontos nos quadrantes correspondentes e calculando a Disponibilidade do sinal, dada pela média das amostras *online* em relação ao número de amostras pertencentes ao quadrante.

Figura 9 – Exemplo de estrada traçada na abordagem de agregação de dados.

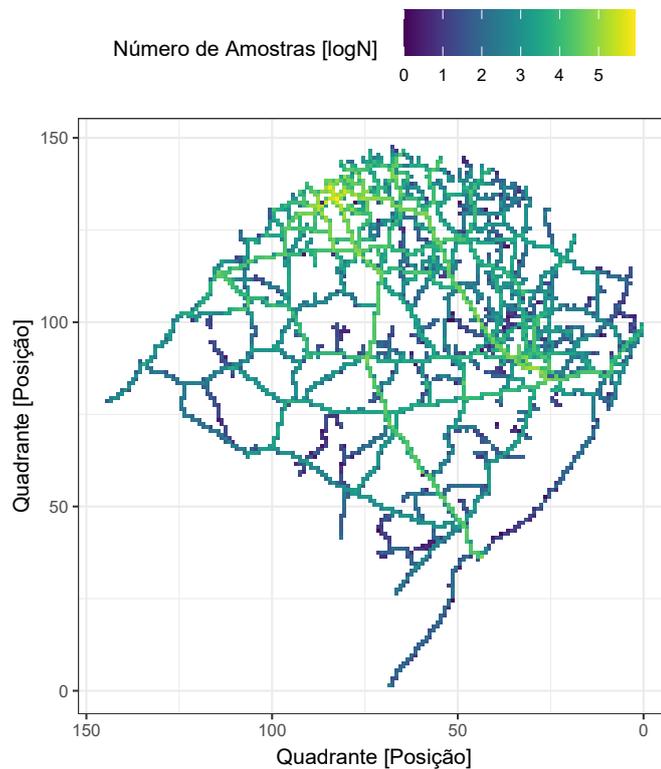


Fonte: Autor.

A Figura 10 ilustra o mapa resultante do processo de agregação de dados aplicado a dados reais coletados de janeiro de 2017 a dezembro de 2020, para o estado do RS. Em particular, o mapa apresenta dados plotados em escala logarítmica para destacar o impacto da quantidade de amostras usadas para gerar cada quadrante.

De fato, a distribuição das amostras coletadas pelos veículos é muito irregular na área geográfica considerada. Não há uma quantidade regular de caminhos percorrendo

Figura 10 – Mapa da distribuição de dados no RS, com $\delta = 5\text{km}$. Na legenda, N indica o número de dados por quadrante.



Fonte: Autor.

o RS no intervalo de tempo em que a coleta de dados foi aplicada. Neste exemplo, ≈ 27 milhões de amostras resultaram em $\approx 3,5$ mil pontos plotados no mapa. A abordagem de agregação de dados foi aplicada em um mapa de grade usando $\delta = 5\text{km}$. A agregação reduz o número de amostras e organiza os pontos a serem plotados no mapa.

No mapa da Figura 10, a cor azul significa que poucos pontos foram usados como entrada para o processo de agregação e amarelo significa a existência de muitos pontos nos dados de entrada originais. Há duas áreas em destaque no mapa, a leste e a noroeste, com predominância da cor amarela indicando o grande número de amostras nos dados de entrada. Isso significa que, nessas áreas, o grau de confiança nas informações fornecidas é maior.

3.3.3 Cálculo da Disponibilidade e Tratamento da Incerteza dos dados

Como efeito colateral do processo de agregação para o cálculo da Disponibilidade, dados podem ser perdidos, dado que um conjunto de amostras é transformado em um único valor resultante. Assim, é necessário também o tratamento da Incerteza dos dados. Para representar esta limitação, no escopo deste trabalho, foi computada a Disponibilidade

do sinal e o grau de Incerteza associado a essa informação.

Para calcular a Disponibilidade do sinal, é selecionado o número de pontos *online* n em um vetor $o = \{o_1, o_2, \dots, o_n\}$, em um determinado quadrante, limitado por LAT (x_1, x_2) e LNG (y_1, y_2). Assim, a Disponibilidade do sinal ρ é a fração da soma dos pontos *online* em relação ao número total de pontos em determinado quadrante, e pode ser dada por:

$$\rho = \frac{\sum_{i=1}^n o_i}{n_i}, \quad (3.1)$$

onde o valor resultante é um número entre 1 e 0. No extremo, 0 significa que a Disponibilidade do sinal é ruim, e 1 a Disponibilidade é boa. Então, para descrever a Disponibilidade do sinal ρ no mapa, foi utilizado um esquema de gradiente de cores, entre azul e vermelho:

- a cor azul, representada por 1, indica a Disponibilidade do sinal (ou seja, a maioria das amostras está *online*);
- a cor vermelha representada por 0, indica a Indisponibilidade do sinal (ou seja, a maioria das amostras está *offline*).

Com relação à Certeza/Incerteza, foi estimada a Incerteza u , calculada pelo desvio padrão sd da existência de pontos *online/offline* no mesmo quadrante, que é dado por:

$$u = 2 \cdot \frac{sd(o_i)}{\sqrt{n_i}}, \quad (3.2)$$

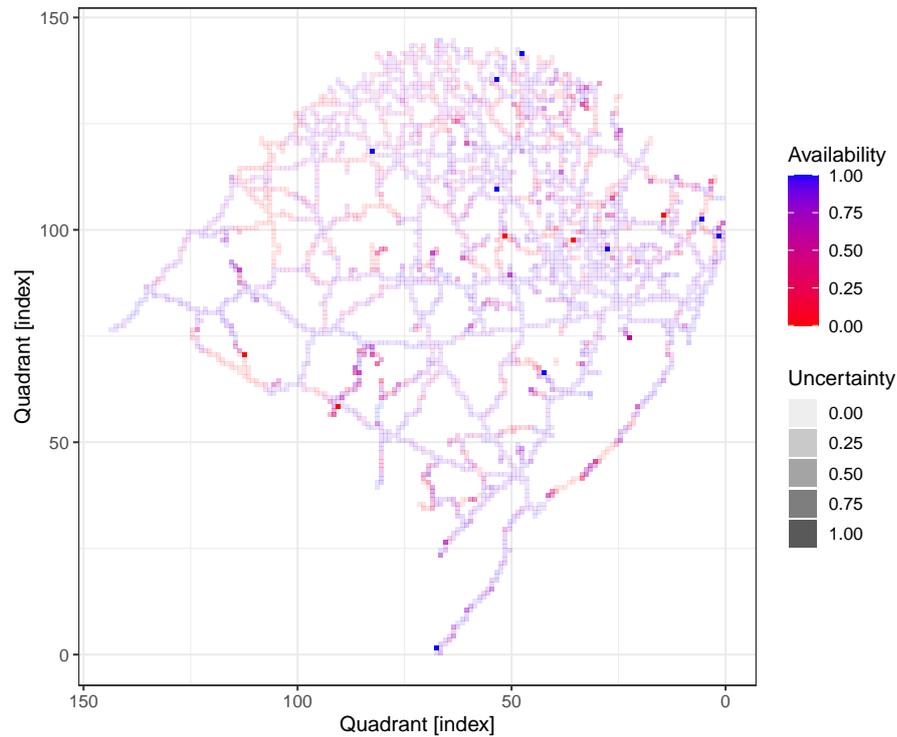
onde a constante $\sigma = 2$ na Eq. 3.2 significa que foi usado um intervalo de confiança de 95%, segundo a regra 68-95-99,7.

Como um ponto no mapa de Disponibilidade do sinal está associado a várias amostras de dados, ou seja, as amostras foram agregadas, também é gerado outro mapa gradiente com o grau de certeza/incerteza dos dados. O esquema de cores usado neste mapa é descrito por:

- a cor mais escura indica Incerteza u , e é representada por 1;
- a cor mais clara indica Certeza, e é representada por 0.

A Figura 11 descreve a aplicação de abordagem de agregação de dados usando as duas métricas, Disponibilidade de sinal ρ , calculada pela Eq. 3.1, onde 1 indica sinal *online* e 0 indica sinal *offline*, e a Incerteza u , dada pela Eq. 3.2, onde os resultantes 0 significam forte confiança de que os dados estão corretos, 1 a informação pode estar incorreta (já que poucas amostras foram usadas no processo de agregação). Na figura, o esquema de cores usado para descrever a Incerteza no mapa destaca o que está ruim. Pontos em azul e vermelho forte têm Incerteza muito alta. Já os pontos claros indicam baixa Incerteza.

Figura 11 – Disponibilidade (Availability) e Incerteza (Uncertainty) do sinal das rodovias do estado do RS, utilizando todas as amostras no período total considerado.



Fonte: Autor.

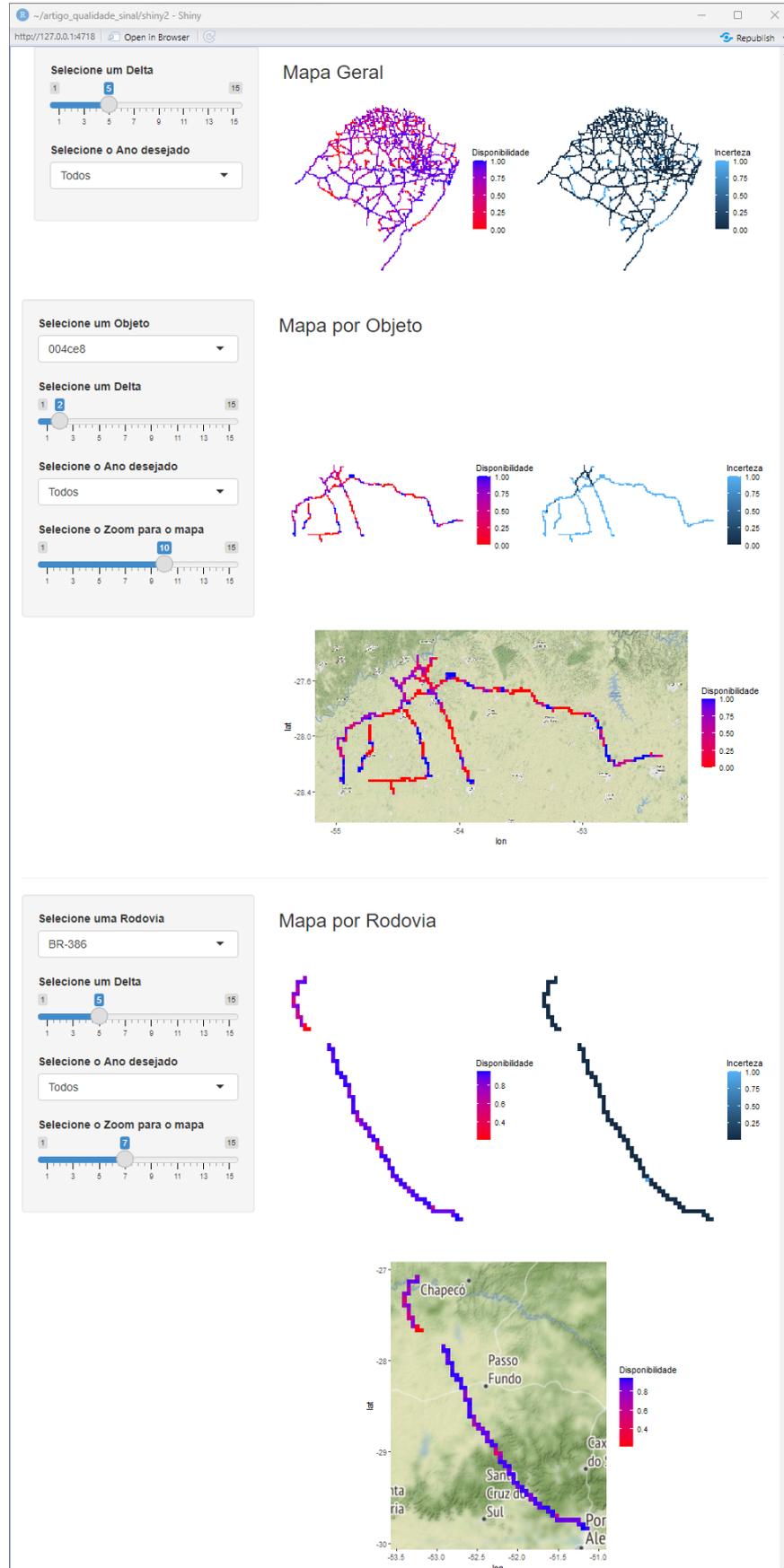
Na metodologia proposta, é imposta uma alta Incerteza sempre que o número de amostras da entrada do processo de agregação é abaixo de um limite. Em particular, os experimentos conduzidos no escopo deste trabalho indicam que abaixo de 30 amostras⁴ em um determinado quadrante com δ de 1km refletem uma medição de Disponibilidade do sinal de má qualidade.

3.4 FERRAMENTA DE VISUALIZAÇÃO DOS RESULTADOS

Para que os resultados deste estudo possam ser analisados por diferentes perspectivas, a metodologia está sendo adaptada para ser executada em formato de página *web*, usando o suporte do pacote *Shiny*⁵. A Figura 12 apresenta um esboço da página *web* que se divide em três seções: a primeira com o Mapa Geral, a segunda com a análise do Mapa por Objeto e a terceira para análise por rodovia, ou Mapa por Rodovia. Quando a página *web* ficar totalmente disponibilizada, o usuário poderá selecionar diferentes deltas, e terá a possibilidade verificar dados de todos os anos ou de um ano específico.

⁴Esta é uma escolha comumente abordada em trabalhos estatísticos (HOGG ELLIOT A. TANIS, 2014), apesar de existir controvérsia na comunidade (COHEN, 1990)

⁵<https://shiny.rstudio.com>

Figura 12 – Aparência da ferramenta *web* em desenvolvimento.

Fonte: Autor.

4 EXPERIMENTOS

Este capítulo apresenta os experimentos realizados com a metodologia descrita no capítulo anterior, e também apresenta uma breve discussão sobre os resultados obtidos.

Para realizar os experimentos aqui descritos, foram reunidas todas as amostras de dados, que compreendem o mês de Julho, do ano de 2017 até o ano de 2020. O conjunto de dados passou pela etapa de Geração do Dataset (Seção 3.2), totalizando ao final ≈ 27 milhões de amostras com ≈ 3 GBytes de tamanho em disco.

4.1 DISPONIBILIDADE X INCERTEZA

Utilizando a metodologia proposta, foram realizados experimentos com o objetivo de avaliar a Disponibilidade e a Incerteza geral nas rodovias do estado do RS, usando as Eqs. 3.1 e 3.2, respectivamente. Para isso, foram reunidas todas as amostras de dados (≈ 27 milhões) e plotadas no mapa com granularidade de 1 km ($\delta=1\text{km}$).

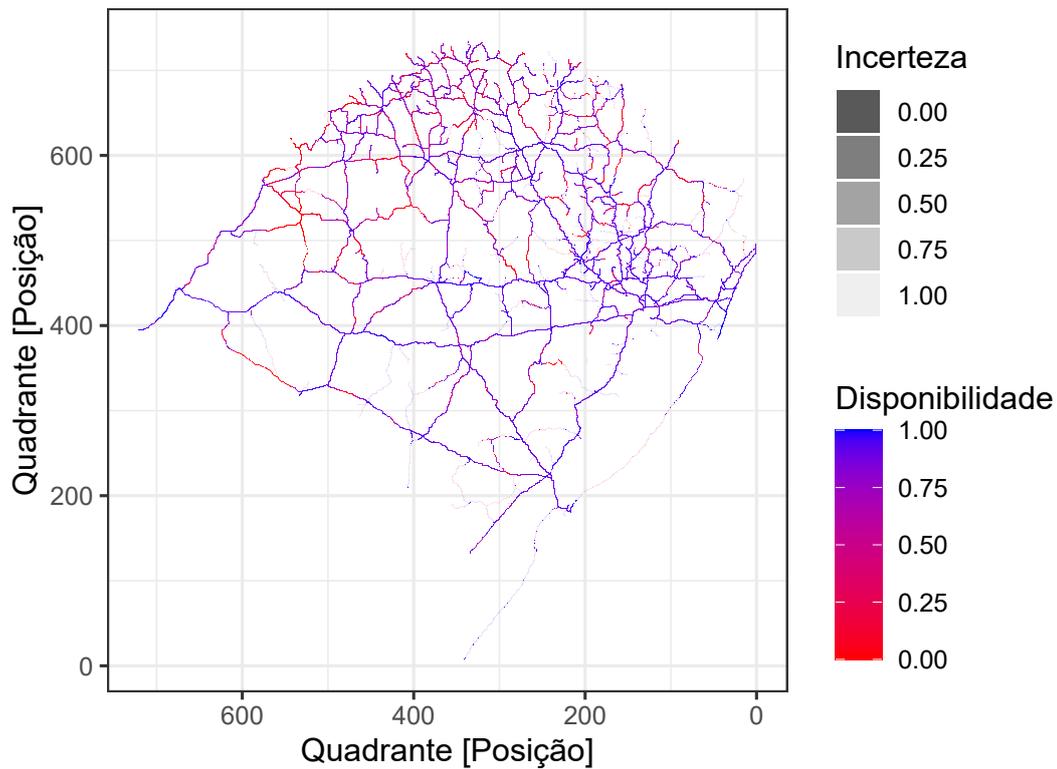
A Figura 13 ilustra o primeiro resultado obtido. Na figura, a escala da Incerteza é representada por um degradê de cores, onde pontos mais escuros, indicam uma baixa Incerteza (0), e os pontos mais claros, indicam uma alta Incerteza (1), sobre os resultados. De fato, a alta Incerteza está associada a uma pequena quantidade de amostras por quadrante, enquanto que a baixa Incerteza está associada a uma grande quantidade de amostras por quadrante. O degradê de cores azul até vermelho indicam, respectivamente, uma boa (1) e uma má (0) Disponibilidade de sinal.

De acordo com os experimentos, a Disponibilidade média (dada pela soma da Disponibilidade de cada um dos quadrantes e dividida pelo número total de quadrantes) de sinal nas rodovias do estado do RS é de 0,70. Este valor indica que a Disponibilidade geral de sinal em rodovias ficou em 70%, e está de acordo com o Projeto de Lei 1721/2019, que prevê "[...] meta de cobertura de no mínimo 60% dos trajetos das rodovias federais pavimentadas existentes na área da futura prestação do serviço [...]" Brasil (2019).

Já a Incerteza geral (dada pela soma da Incerteza de cada um dos quadrantes e dividida pelo número total de quadrantes) calculada foi de 0,29, mostrando uma boa quantidade de amostras na grande maioria das rodovias, trazendo assim maior Certeza sobre os resultados de Disponibilidade de sinal obtidos.

Avaliando a Disponibilidade de Sinal nas Regiões Geográficas Intermediárias do estado do RS, conforme IBGE (2017), pode-se dar mais atenção às rodovias das regiões de Porto Alegre, Caxias do Sul, Santa Cruz Do Sul - Lajeado e Santa Maria, onde a Disponibilidade atingiu níveis bastante satisfatórios, ficando na maioria dos quadrantes próxima da Disponibilidade máxima (1). Para as regiões intermediárias de Ijuí e Passo Fundo, de

Figura 13 – Disponibilidade e Incerteza do sinal 2G das rodovias do estado do RS.

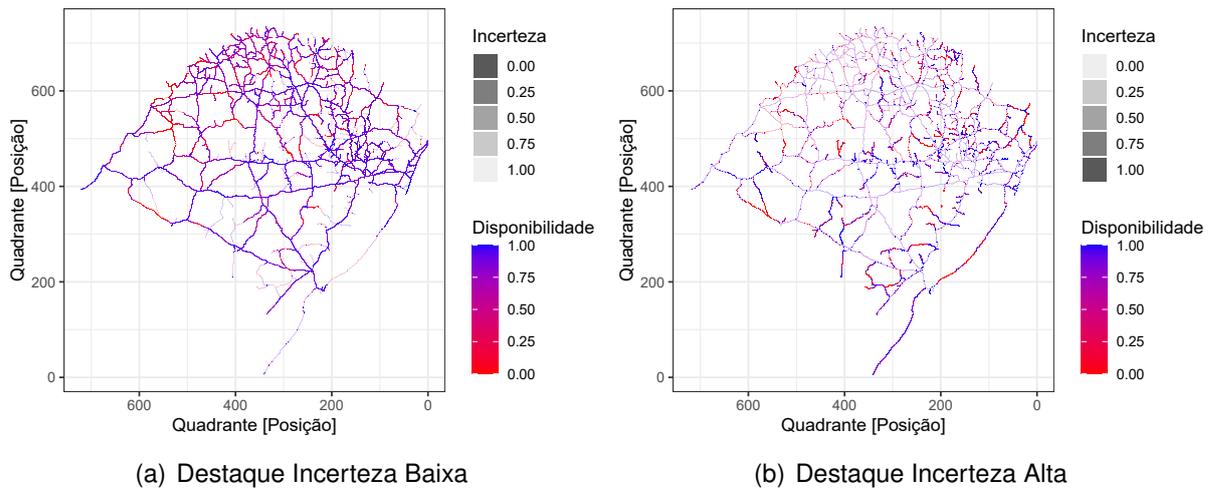


Fonte: Autor.

forma geral, a Disponibilidade oscila bastante entre uma média e uma má Disponibilidade de sinal, com alguns trechos sem Disponibilidade de sinal. E por último, as regiões intermediárias de Uruguaiana e Pelotas, que apesar de serem regiões essencialmente agrícolas e pouco povoadas, possuem na maioria dos trechos das rodovias uma boa Disponibilidade de sinal, com exceção de algumas áreas onde a Disponibilidade é realmente nula.

Para entender melhor a relação entre a Disponibilidade e a Incerteza do sinal, foi realizado na Figura 14 um estudo comparativo. Para o estudo, foram utilizados dois mapas, sendo o primeiro (a) igual ao da Figura 13, e o segundo (b) inverte-se a escala da Incerteza. Desta forma, em (a) são deixadas mais escuras - destacadas - as áreas com baixa Incerteza da Disponibilidade do sinal (0) e em seguida, em (b) são destacadas as áreas com alta Incerteza (1) sobre a Disponibilidade do sinal.

Figura 14 – Comparação de mapas destacando Incerteza Baixa (a) e Incerteza Alta (b).



Fonte: Autor.

Desta forma, na Figura 14 (a), observa-se que a maior parte das rodovias do estado do RS possuem baixa taxa de Incerteza, com todo o seu trajeto, ou grande parte dele, possuindo Incerteza 0 ou muito próxima disso.

Na Figura 14 (b), percebe-se que normalmente os trechos de rodovias com alta Incerteza pertencem a rodovias de acesso a regiões menos populosas, geralmente de cidades do interior do estado.

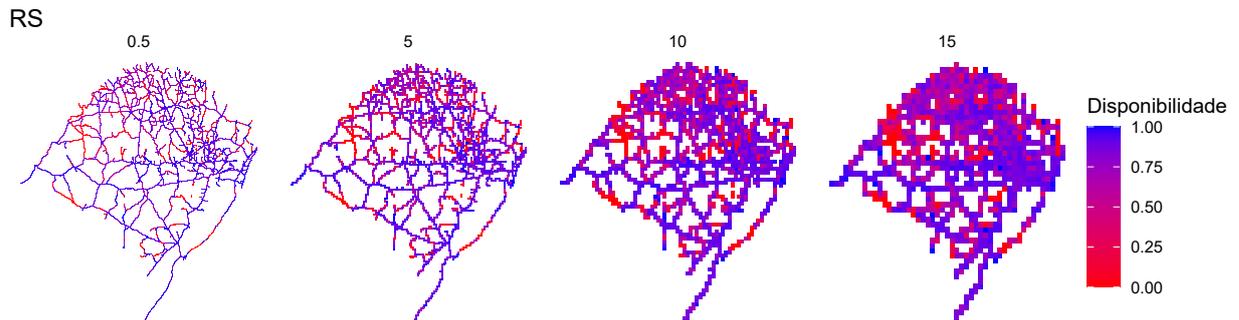
Através da comparação dos gráficos da Figura 14, observa-se também que não existe correlação entre as escalas de Incerteza e Disponibilidade, ou seja, um local com baixa Incerteza não corresponde a um local com baixa Disponibilidade do sinal e/ou vice-versa.

4.2 ESTUDO DA GRANULARIDADE

Neste estudo, foi avaliado o impacto de diferentes granularidades na obtenção do valor da Disponibilidade para o experimento, ou seja, a influência do uso de diferentes tamanhos de quadrados no mapa, tanto na percepção visual, quanto nos resultados quantitativos. O tamanho dos quadrados no mapa é dado por δ no processo de agregação de dados.

Para realizar este estudo, foi calculada a Disponibilidade do sinal 2G usando diferentes valores para δ , com $\delta = 0.5, 5, 10$ e 15 km. Os mapas obtidos são representados na Figura 15 com dados de todo o estado do RS.

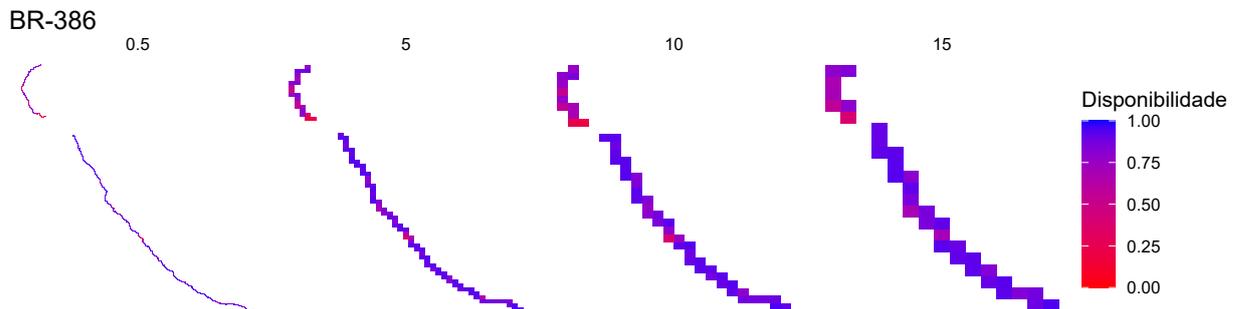
Figura 15 – Disponibilidade de sinal 2G no RS usando diferentes valores para δ , sendo $\delta=\{0.5, 5, 10 \text{ e } 15\}$ (km).



Fonte: Autor.

A Figura 16 mostra em destaque apenas a BR-386 por ser a rodovia com maior quantidade de dados disponíveis e, possivelmente, ofereça baixa Incerteza sobre os resultados obtidos no experimento.

Figura 16 – Disponibilidade de sinal 2G na rodovia BR-386 usando diferentes valores para δ , sendo $\delta=\{0.5, 5, 10 \text{ e } 15\}$ (km).



Fonte: Autor.

É possível visualizar nos mapas apresentados nas Figuras 15 e 16 que a granularidade influencia claramente na aparência do mapa. Como não existe um gabarito para a Disponibilidade, é difícil escolher qual granularidade é a mais correta, ficando a critério do usuário definir a melhor granularidade de acordo com os objetivos da sua análise.

Entretanto, como forma de quantificar o impacto dos diferentes valores de δ no resultado (mapa visual), foram calculados os valores quantitativos da Disponibilidade do sinal para os diferentes valores de δ . Também foi computado o valor da **Amplitude** (A) que é definida pela equação:

$$A = V_{max} - V_{min} \quad (4.1)$$

onde, V_{max} representa o maior valor obtido, e V_{min} representa o menor valor

obtido. Para encontrar o valor de A utiliza-se o maior e o menor valor computado para a Disponibilidade do sinal de uma área com o uso dos diferentes δ .

Os resultados obtidos são apresentados na Tabela 4.

Tabela 4 – Estudo do impacto da granularidade na visualização da Disponibilidade do sinal, variando o valor de δ .

Área	δ 0,5	δ 5	δ 10	δ 15	Amplitude A
RS	0,703	0,688	0,685	0,696	0,018
BR-386	0,828	0,834	0,822	0,841	0,019

Fonte: Autor.

Como indicam os valores da Tabela 4, pode-se observar que apesar da variação dos valores de δ , ou seja, aumentar o tamanho dos quadrantes (de 0,5 km para 15 km), os resultados para a Disponibilidade do sinal 2G praticamente são mantidos. De fato, pode ser observada uma pequena variação nos valores da tabela, com uma **Amplitude A** máxima de 0,018 no cálculo da Disponibilidade de todas as rodovias do RS, e de 0,019 no caso da rodovia BR-386. Essa variação pode ser considerada muito baixa, se comparada aos valores da variação de δ .

Uma ideia para tentar identificar a melhor granularidade para o mapa é considerar a **Acurácia** dos valores obtidos no experimento, e escolher a granularidade que oferece a melhor Acurácia. A Acurácia é a quantificação de quão perto as amostras estão do alvo, isto é, o valor aceito como correto (TAYLOR, 1997). Para determinar se o resultado de um experimento tem Acurácia, o valor obtido precisa ser comparado com o valor correto. Desde que não há um gabarito (valor correto) para este estudo, um conceito chamado erro percentual e pode ser usado. Seja y o valor correto, \hat{y} o valor obtido, o erro percentual e é dado por:

$$e = \left| \frac{(y - \hat{y})}{y} \right| \cdot 100 \quad (4.2)$$

que indica o percentual de diferença entre o valor correto e o valor obtido experimentalmente. Para a computação de y , novamente, desde que não há gabarito para o problema em questão, a média dos valores obtidos pode ser usada. Aplicando a Eq. 4.2 para os valores (obtidos) na Tabela 4, obtêm-se os valores apresentados na Tabela 5.

De acordo com os cálculos para este experimento, para o estado do RS, o melhor valor de δ é 15 (valor 0,43), e para a BR-386, o melhor valor de δ é 5 (valor 0,33). Novamente, é difícil escolher qual granularidade é a mais correta para o problema em questão. Ressalta-se que fica a critério do usuário definir a melhor granularidade de acordo com os objetivos do estudo.

Tabela 5 – Estudo do impacto da granularidade na visualização da Disponibilidade do sinal, variando o valor de δ e considerando a Acurácia do experimento.

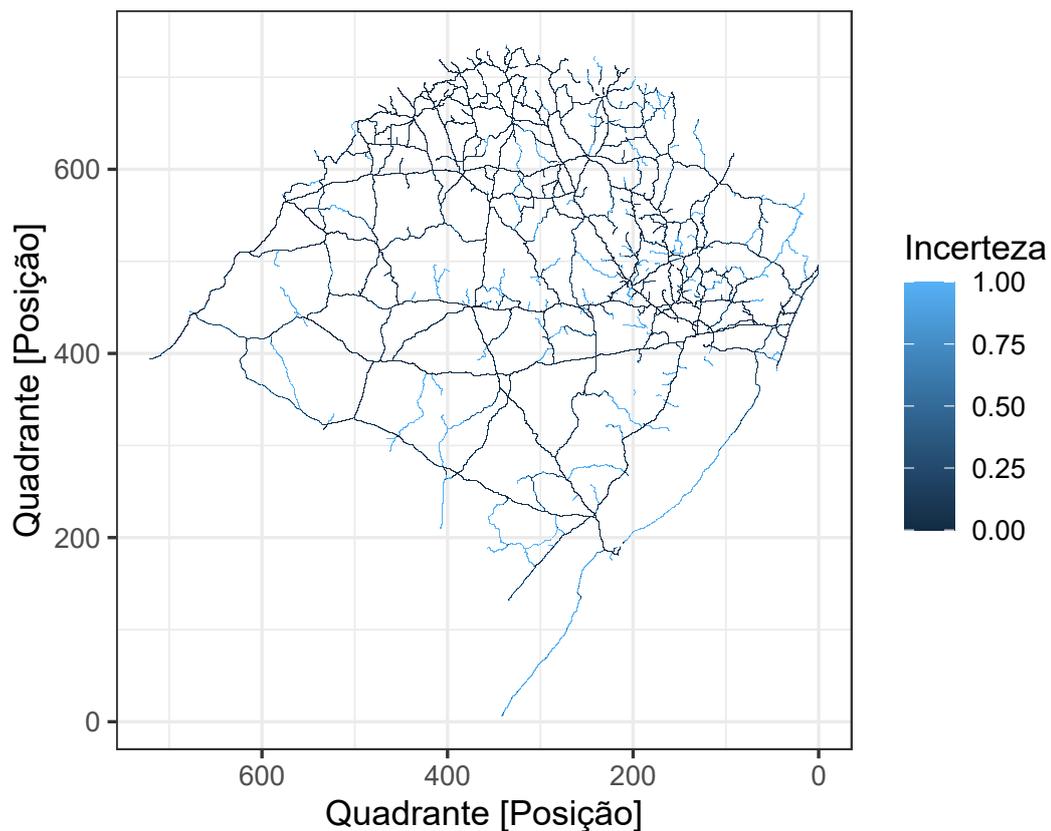
Área	δ 0,5	δ 5	δ 10	δ 15
RS	1,44	0,72	1,15	0,43
BR-386	0,39	0,33	1,11	1,17

Fonte: Autor.

4.3 ESTUDO DA INCERTEZA

Neste trabalho, para entender melhor a distribuição dos dados utilizados e para avaliar o impacto da distribuição dos dados na qualidade do resultado, foram mapeadas as áreas de maior Incerteza. A Figura 17 apresenta a plotagem dos dados considerando $\delta=1\text{km}$ e o número mínimo de amostras em 30 por quadrante – desta forma, quadrantes com menos de 30 amostras são considerados incertos. Na figura, os quadrantes na cor azul claro indicam trechos com alta Incerteza.

Figura 17 – Análise geral da Incerteza no estado do RS.



Fonte: Autor.

Observa-se que as rodovias com alta Incerteza (1), normalmente são as utilizadas para acesso às cidades menores de áreas interioranas. Além destes, o trecho da BR-101

paralelo a Lagoa dos Patos, e o trecho da BR-471 situado logo abaixo da Lagoa dos Patos, também possuem alta Incerteza.

4.4 ESTUDO COM FOCO NO DISPOSITIVO

Também foi realizado um estudo com foco em dispositivos, com o objetivo de avaliar se haveriam fatores externos que poderiam impactar no resultado do cálculo da Disponibilidade de sinal.

Para isso, foi escolhida a rodovia BR-386, novamente, por ser a rodovia com o maior número de amostras. Além disso, três dispositivos foram selecionados utilizando dois critérios:

- possuir o maior número de registros,
- possuir dados da rodovia BR-386 em todas as semanas dos quatro anos analisados.

Os dois critérios são necessários para que se tenha uma baixa Incerteza sobre o resultado e também para que se possa analisar todo o período selecionado, já que alguns veículos utilizam bastante uma rodovia, porém em trechos específicos ou em períodos de tempo curtos, não sendo possível uma análise temporal completa. Em relação aos três dispositivos considerados neste estudo, existem diferentes quantidades de amostras ao longo dos anos, conforme ilustrado na Tabela 6.

Tabela 6 – Quantidade de amostras de diferentes dispositivos - referente à BR-386.

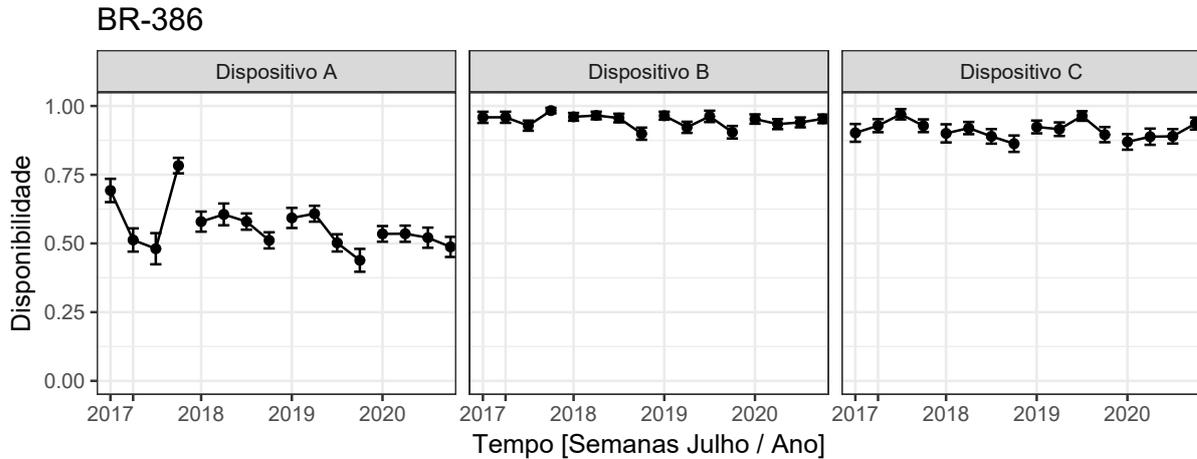
Identificação do Dispositivo	2017	2018	2019	2020	Total
Dispositivo A	17.553	49.199	52.423	47.264	166.439
Dispositivo B	21.769	31.453	32.539	34.149	119.910
Dispositivo C	21.480	27.670	23.978	23.723	96.851

Fonte: Autor.

No total, a quantidade de amostras dos três dispositivos é bem considerável (\approx 380k amostras). Em particular, o Dispositivo A é o com menos amostras no ano de 2017, com 17.553 amostras, porém é o que possui mais amostras nos anos seguintes, totalizando no somatório final 166.439 amostras. Os resultados relativos à Disponibilidade de sinal com relação aos diferentes dispositivos são plotados na Figura 18.

Para a obtenção da figura, as amostras foram organizadas em semanas (referentes ao mês de Julho – mês das amostras), dos anos de 2017 à 2020. No gráfico, o ponto indica a Disponibilidade média calculada para aquela semana e as barras indicam o valor

Figura 18 – Disponibilidade ρ do sinal para a BR-386 com diferentes dispositivos (Dispositivo A, Dispositivo B e Dispositivo C) usando $\delta = 1$ km.



Fonte: Autor.

máximo e o valor mínimo encontrado nos quadrantes. Para este estudo, foi mantido o δ em 1 km.

Os *Dispositivos B* e *C* apresentaram resultados excelentes de Disponibilidade de sinal, mantendo a mesma média de Disponibilidade ao longo dos anos. Com relação ao *Dispositivo A*, a Disponibilidade média de sinal ficou pouco acima da média, indicando que, de fato, este dispositivo foi instalado como um dispositivo de segurança, isto é, em um local mais oculto. Assim, o dispositivo foi instalado mais abaixo na cabine do caminhão e em um local que dificulte a comunicação com a antena externa. Mais especificamente, a melhora na Disponibilidade de sinal do *Dispositivo A*, na primeira e última semana analisadas em 2017, se deve ao trecho que o veículo percorreu, que ficou limitado a um trecho da BR-386 onde a Disponibilidade de sinal era melhor.

4.5 ESTUDO COM FOCO NAS RODOVIAS

Por fim, foi realizado um estudo para avaliar a Disponibilidade do sinal 2G em rodovias, com base em vias específicas. O foco deste estudo se concentra em realizar a análise histórica a partir da média de Disponibilidade do sinal de cada uma das rodovias selecionadas. Sendo assim, foi calculada a média geral de Disponibilidade do sinal para cada um dos anos, considerando o período de 2017 à 2020 com um δ de 1km.

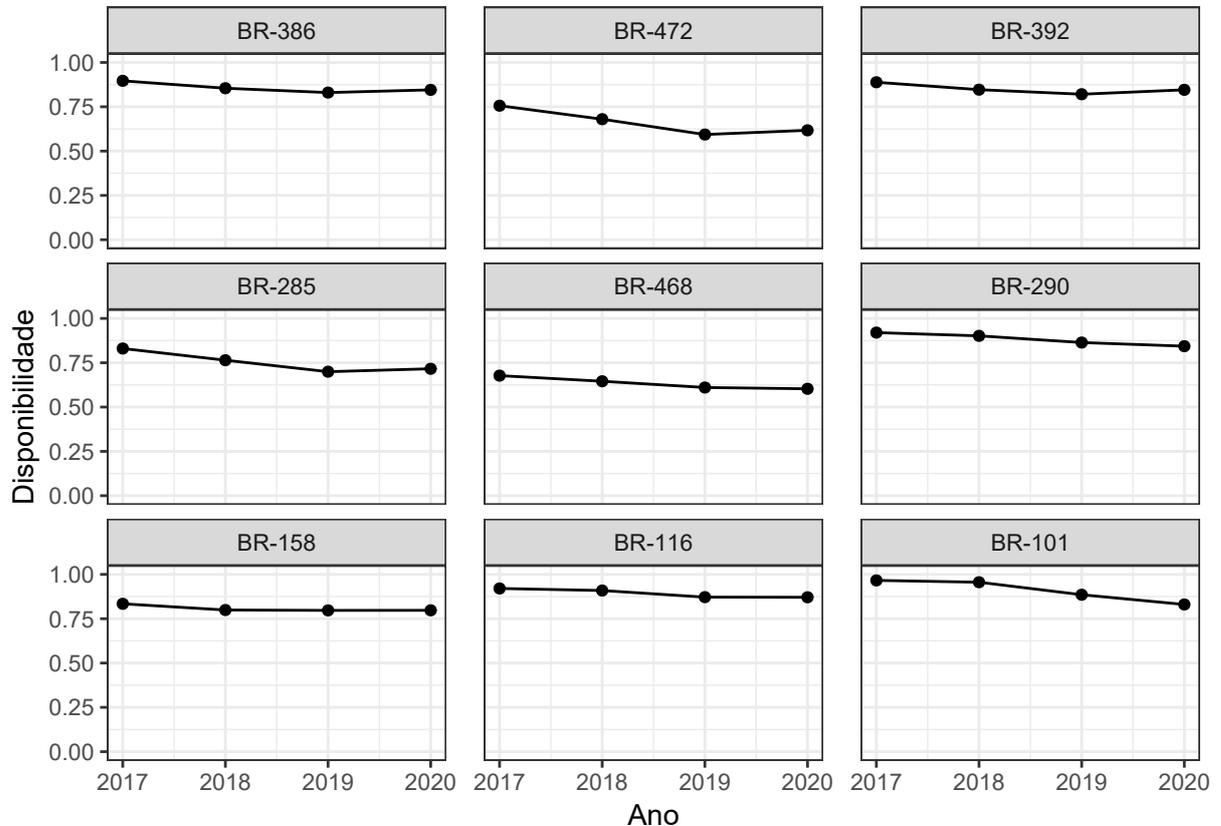
A abordagem utilizada neste estudo é bastante semelhante a dos estudos anteriores, com exceção da função *loop*:

- Ler os dados contidos no arquivo CSV;

- Usando um δ predefinido, calcular a grade de quadrados para toda a área do mapa;
- Execução da função abaixo em *loop*, passando o nome da rodovia como parâmetro:
 - Filtrar os dados pelo atributo *name*;
 - Agrupar os dados por ano, do ano de 2017 a 2020;
 - Calcular um resumo com os dados pertencentes a cada um dos quadrados da grade;
 - Selecionar os quadrados e respectivos dados que pertencem à rodovia selecionada;
 - Calcular a Disponibilidade do sinal ρ usando a Eq. 3.1;
 - Plotar em gráfico a Disponibilidade ρ de cada rodovia de maneira individual;
- Por último, a análise dos resultados obtidos.

Para obter resultados confiáveis, foram selecionadas para o estudo as 9 rodovias com mais dados disponíveis, sendo elas: BR-386, BR-472, BR-392, BR-285, BR-468, BR-290, BR-158, BR-116 e BR-101. Os resultados são apresentados na Figura 19, para os quatro anos considerados no estudo.

Figura 19 – Disponibilidade do sinal nas rodovias do estado do RS, em quatro anos, onde cada ponto em cada gráfico representa um mês.



Fonte: Autor.

A primeira análise que pode ser feita sobre o conjunto de rodovias plotado na Figura 19, é que todas as rodovias analisadas ficaram com a média de Disponibilidade de sinal mínima acima de 0,5 ao longo de todo o período analisado.

A rodovia BR-472 obteve o pior desempenho e a maior queda na Disponibilidade do sinal. Caindo em 0,14 a Disponibilidade do sinal, se comparados os anos de 2017 e 2020 e, com uma Amplitude de 0,16. Após a BR-472, as BRs 101 e 285 obtiveram a maior queda na Disponibilidade do sinal respectivamente. Sendo a BR-101 a rodovia que inicialmente possuía a melhor Disponibilidade do sinal dentre todas as rodovias analisadas. As rodovias BR-158 e BR-116 possuem a maior estabilidade na Disponibilidade do sinal, que manteve-se praticamente inalterado, variando somente 0,04 e 0,05 respectivamente.

E para concluir, de maneira geral, observa-se a decadência da Disponibilidade do sinal 2G em rodovias, sinalizando assim, a migração por parte das operadoras móveis para redes mais avançadas, como 3G, 4G ou até mesmo a 5G. Quando comparado o ano de 2017 ao ano de 2020, todas as rodovias analisadas sofreram uma diminuição na Disponibilidade do sinal. As rodovias BR-472, BR-285 e BR-101 foram as que mais sofreram com a diminuição da Disponibilidade do sinal, apesar de ainda terem uma Disponibilidade geral

acima de 0,6.

Para uma análise mais aprofundada, a Tabela 7 possui os dados analisados e plotados na Figura 19, além do valor da Amplitude, dado pela Eq. 4.1. Para a obtenção da tabela, primeiramente os dados são separados por rodovia e por ano, e após é calculada a Disponibilidade média de cada uma das rodovias – dada pela soma da Disponibilidade dos quadrantes pertencentes à rodovia, dividido pelo número de quadrantes.

Tabela 7 – Disponibilidade do sinal por rodovia do estado do RS no período 2017-2020.

Rodovia	2017	2018	2019	2020	Amplitude <i>A</i>
BR-386	0,89	0,85	0,83	0,84	0,07
BR-472	0,76	0,68	0,59	0,62	0,16
BR-392	0,89	0,84	0,82	0,84	0,07
BR-285	0,83	0,76	0,70	0,71	0,13
BR-468	0,67	0,64	0,60	0,59	0,08
BR-290	0,92	0,90	0,86	0,84	0,08
BR-158	0,83	0,80	0,80	0,80	0,04
BR-116	0,91	0,90	0,86	0,86	0,05
BR-101	0,97	0,96	0,89	0,83	0,14

Fonte: Autor.

A Tabela 7 indica que, com exceção da rodovia BR-468, as demais rodovias atendem ao que foi proposto no Projeto de Lei 1721/2019 (BRASIL, 2019). A rodovia BR-116 se destaca por apresentar ao final do período analisado, o melhor serviço e a menor variação (Amplitude) ao longo dos anos. Já as rodovias BR-101 e BR-472 se destacam pelo alto declínio na Disponibilidade do sinal.

5 CONCLUSÃO

Neste trabalho, foi proposta uma metodologia para descrever a cobertura do sinal de rede móvel em rodovias, utilizando dados de dispositivos IoT de rastreamento veicular. Os dispositivos, por sua vez, possuem embarcado em seu *hardware* um módulo de transmissão de dados que conta com a tecnologia 2G, utilizando *simcards* da operadora Vivo. Como estudo de caso, foi usado um conjunto de dados relativo às viagens de veículos rastreados no estado do Rio Grande do Sul, Brasil.

Ao contrário de trabalhos anteriores, a metodologia aqui apresentada considera a Incerteza dos dados, medida pela quantidade de dados agregados por quadrante e com uma medida mínima de amostras por quadrante. Uma vez que uma característica natural do tráfego é a falta de distribuição uniforme, podendo até mesmo existirem regiões e/ou vias com poucas medições realizadas, considera-se de suma importância avaliar, além da métrica de Disponibilidade do sinal, a métrica da Incerteza.

O *dataset* utilizado neste estudo se mostrou bastante completo por compreender geograficamente rodovias de quase todo o estado do RS com dados de sinal, indicando assim, uma boa certeza sobre os resultados de Disponibilidade obtidos ao longo das rodovias. Uma limitação do conjunto de dados é o uso da tecnologia 2G da operadora Vivo, porém esta não é uma limitação da Metodologia, que pode utilizar dados de qualquer tecnologia (3G, 4G, 5G e outras) ou operadora móvel, podendo inclusive utilizar além de dados binários, dados em escala, como por exemplo o *RSSI*.

A partir da análise dos resultados obtidos, pode-se constatar de forma geral, que a Disponibilidade do sinal 2G no estado do RS ainda não é a mais adequada, mas predomina a Disponibilidade do sinal, com uma média geral de Disponibilidade do sinal em rodovias, acima de 0,59. O destaque fica para a metade leste do estado, que possui de maneira geral as rodovias com melhor Disponibilidade de sinal.

Já o estudo da Disponibilidade X Incerteza, mostrou a importância do uso da métrica da Incerteza aliada a métrica da Disponibilidade. Ficando de forma clara quais os trechos de rodovias que possuem baixa Incerteza sobre o resultado da Disponibilidade.

Por fim, o estudo da granularidade trouxe os desafios da definição de um δ ideal para a análise visual dos mapas. O estudo mostrou que o principal impacto no aumento ou diminuição do δ se dá na percepção visual de quem está analisando, não ocorrendo grandes alterações nos resultados quantitativos obtidos, com uma Amplitude y máxima de 0,018 na análise da Disponibilidade do sinal de todo o RS, com um δ variando entre 0,5 km e 15 km.

Outra importante contribuição deste trabalho se deve a análise da Disponibilidade do sinal sob a perspectiva de dispositivos IoT no mundo real, trazendo a análise da Disponibilidade como ela é, com todos os obstáculos e intempéries que passam os dispositivos

IoT no seu dia a dia, mostrando uma visão realista da atual conectividade em rodovias. E claro, fornecendo resultados detalhados da análise da Disponibilidade e Incerteza do sinal sob diversas perspectivas.

Conclui-se que a metodologia proposta pode ser um ponto de partida para permitir que governos e empresas responsáveis planejem de forma mais eficientes ações para melhorar a cobertura do sinal da rede móvel em rodovias.

5.1 TRABALHOS FUTUROS

Como trabalhos futuros, pode ser explorada a divisão das rodovias não por quadrantes de mesmo tamanho, mas utilizando segmentos de estrada, como em Jarvis et al. (2018), onde os autores realizam a segmentação de rotas de trens. Isso permitirá descrever melhor a rodovia, sem considerar entroncamentos.

Outro trabalho futuro é agregar dados por regiões de similaridade como em Fida et al. (2017), o que permite ter uma visão macro do serviço de cobertura. Essas duas formas alternativas de agregação permitirão uma comparação direta de resultados com os referidos trabalhos relacionados. Alternativamente, os dados do conjunto de dados podem ser associados a cidades, ou regiões, possibilitando observar a Disponibilidade do sinal de uma perspectiva diferente.

REFERÊNCIAS

ABRAHAMSSON, H. et al. Connected vehicles in cellular networks: Multi-access versus single-access performance. In: **2018 Network Traffic Measurement and Analysis Conference (TMA)**. [S.l.: s.n.], 2018. p. 1–6.

ADARSH, V. et al. Coverage is not binary: Quantifying mobile broadband quality in urban, rural, and tribal contexts. In: **2021 International Conference on Computer Communications and Networks (ICCCN)**. [S.l.: s.n.], 2021. p. 1–9.

ALIMPERTIS, E. et al. City-wide signal strength maps: Prediction with random forests. In: **The World Wide Web Conference**. New York, NY, USA: Association for Computing Machinery, 2019. p. 2536–2542. ISBN 9781450366748.

BOKANI, A. et al. Comprehensive mobile bandwidth traces from vehicular networks. In: **Proceedings of the 7th International Conference on Multimedia Systems**. New York, NY, USA: Association for Computing Machinery, 2016. (MMSys '16). ISBN 9781450342971.

BRASIL. Projeto de lei n 1721, de 2019. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 2019. Disponível em: <https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1723720&filename=Tramitacao-PL\%201721/2019>. Acesso em: 08/12/2022.

BUDIMAN, E.; WICAKSONO, O. Measuring quality of service for mobile internet services. In: IEEE. **2016 2nd International Conference on Science in Information Technology (ICSITech)**. [S.l.], 2016. p. 300–305.

COHEN, J. Things i have learned (so far). **American Psychologist**, v. 45, n. 12, p. 1304–1312, 1990.

EDQUIST, H. et al. How important are mobile broadband networks for the global economic development? **Information Economics and Policy**, Elsevier, v. 45, p. 16–29, 2018.

ENGIZ, B.; KURNAZ, C. Comparison of signal strengths of 2G/3G/4G services on a university campus. **International Journal of Applied Mathematics, Electronics and Computers (IJAMEC)**, p. 37–42, December 2016.

Fida, M. et al. Zipweave: Towards efficient and reliable measurement based mobile coverage maps. In: **IEEE INFOCOM 2017 - IEEE Conference on Computer Communications**. [S.l.: s.n.], 2017. p. 1–9.

FIDA, M.-R.; MARINA, M. K. Impact of device diversity on crowdsourced mobile coverage maps. In: IEEE. **2018 14th International Conference on Network and Service Management (CNSM)**. [S.l.], 2018. p. 348–352.

Freschi, V. et al. Geospatial data aggregation and reduction in vehicular sensing applications: The case of road surface monitoring. In: **2014 International Conference on Connected Vehicles and Expo (ICCVE)**. [S.l.: s.n.], 2014. p. 711–716.

HOGG ELLIOT A. TANIS, D. Z. R. V. **Probability and Statistical Inference**. [S.l.]: Pearson Education, 2014. 554 p.

IBGE. **Regiões Geográficas - Estado do Rio Grande do Sul**. 2017. Disponível em: <https://geoftp.ibge.gov.br/organizacao_do_territorio/divisao_regional/divisao_regional_do_brasil/divisao_regional_do_brasil_em_regioes_geograficas_2017/mapas/43_regioes_geograficas_rio_grande_do_sul.pdf>. Acesso em: 10/01/2023.

ISABONA, J.; SRIVASTAVA, V. User-centric methodology for objective assessment of service quality in established wireless mobile communication networks. **International Journal on Communications Antenna and Propagation (IRECAP)**, v. 7, p. 26, 02 2017.

JARVIS, C. et al. Visualizing mobile coverage from repetitive measurements on defined trajectories. In: **2018 Network Traffic Measurement and Analysis Conference (TMA)**. [S.l.: s.n.], 2018. p. 1–6.

KHATOONI, A. S. et al. Speedtest-like measurements in 3g/4g networks: The monroe experience. In: **2017 29th International Teletraffic Congress (ITC 29)**. [S.l.: s.n.], 2017. v. 1, p. 169–177.

_____. An open dataset of operational mobile networks. In: **Proceedings of the 18th ACM Symposium on Mobility Management and Wireless Access**. New York, NY, USA: Association for Computing Machinery, 2020. (MobiWac '20), p. 8390. ISBN 9781450381192. Disponível em: <<https://doi.org/10.1145/3416012.3424619>>.

KOUSIAS, K. et al. The same, only different: Contrasting mobile operator behavior from crowdsourced dataset. In: **2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)**. [S.l.: s.n.], 2017. p. 1–6.

LAURIDSEN, M. et al. Coverage comparison of GPRS, NB-IoT, LoRa, and SigFox in a 7800 km² area. In: **2017 IEEE 85th Vehicular Technology Conference (VTC Spring)**. [S.l.: s.n.], 2017. p. 1–5.

LIU, B. et al. Dynamic coverage of mobile sensor networks. **IEEE Transactions on Parallel and Distributed Systems**, IEEE, v. 24, n. 2, p. 301–311, 2012.

LIU, Y. et al. Multi-criteria coverage map construction based on adaptive triangulation-induced interpolation for cellular networks. **IEEE Access**, v. 7, p. 80767–80777, 2019.

_____. Location, localization, and localizability. **Journal of Computer Science and Technology**, v. 25, n. 2, p. 274–297, 2010.

LUTU, A. et al. The good, the bad and the implications of profiling mobile broadband coverage. **Computer Networks**, v. 107, p. 76–93, 2016. ISSN 1389-1286.

MADARIAGA, D. et al. Improving signal-strength aggregation for mobile crowdsourcing scenarios. **Sensors**, v. 21, p. 1084, 02 2021.

MARINA, M. K.; RADU, V.; BALAMPEKOS, K. Impact of indoor-outdoor context on crowdsourcing based mobile coverage analysis. In: **Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges**. [S.l.: s.n.], 2015. p. 45–50.

MONROE. **About Monroe**. 2020. Disponível em: <<https://www.monroe-project.eu/themonroeproject/>>.

RAGO, A. et al. Unveiling radio resource utilization dynamics of mobile traffic through unsupervised learning. In: **2019 Network Traffic Measurement and Analysis Conference (TMA)**. [S.l.: s.n.], 2019. p. 209–214.

ROSA, L.; SILVA, F.; ANALIDE, C. Mobile networks and internet of things infrastructures to characterize smart human mobility. **Smart Cities**, v. 4, p. 894918, 2021.

SCHWIND, A. et al. Crowdsourced network measurements in germany: mobile internet experience from end user perspective. In: **Broadband Coverage in Germany; 14. ITG Symposium**. [S.l.: s.n.], 2020. p. 1–7.

SIEDNER, M. J. et al. Optimizing network connectivity for mobile health technologies in sub-Saharan Africa. **PLoS ONE 7(9): e45643**, September 2012.

STORK, C.; CALANDRO, E.; GAMAGE, R. The future of broadband in africa. **info**, Emerald Group Publishing Limited, 2014.

TAYLOR, J. R. **An introduction to Error Analysis: the study of Uncertainties in Physical Measurements**. [S.l.: s.n.], 1997.

TRANFIELD, D.; DENYER, D.; SMART, P. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. **British journal of management**, Wiley Online Library, v. 14, n. 3, p. 207–222, 2003.

WANG, H.; XIE, S.; LI K.AND AHMAD, M. Big data-driven cellular information detection and coverage identification. **Sensors**, v. 19, n. 4, p. 937–, 2019.

XAVIER, W. Z.; XAVIER, F. H. Z.; MARQUES-NETO, H. T. Visualizing and analyzing georeferenced workloads of mobile networks. In: **2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)**. [S.l.: s.n.], 2017. p. 306–310.

ZHAO, W.; OLSHEFSKI, D.; SCHULZRINNE, H. G. Internet quality of service: An overview. 2000.