

Estudo da Cobertura da Rede de Fibra Óptica nos Municípios Brasileiros sob a luz de Algoritmos Exato e Heurístico

Fausto Luiz Jorge Padua e Edson Luiz França Senne

Resumo— Dos 5570 municípios brasileiros, ainda há 1309 municípios sem conexão à Internet por fibra óptica, o que levanta a seguinte questão: qual a melhor forma de investir nessa cobertura para os municípios?

Este estudo buscou, através do cálculo da Árvore Geradora Mínima com algoritmos exato e heurístico, trazer alguma luz à questão. Foi realizado o cálculo do custo mínimo para a cobertura, com pelo menos um enlace, entre todos os municípios brasileiros, bem como em cenários menores, como cada estado do Brasil.

Por fim, foi realizada uma comparação da precisão e eficiência entre os diferentes algoritmos escolhidos e apresentados os resultados.

Palavras-Chave— Telecomunicações, fibra óptica, municípios, rotas, método exato e heurístico, Brasil.

Abstract— Off the 5570 Brazilian municipalities, there are still 1309 of them without an optical fiber Internet connection, which raises the following question: what is the best way to invest in this type of coverage?

This study sought to bring some light to the issue, through the calculation of the Minimum Spanning Tree with exact and heuristic algorithms. The calculation of the minimum cost for coverage, with at least one link, was performed between all Brazilian municipalities, as well as in smaller scenarios, such as each state in Brazil.

Finally, a comparison of precision and efficiency between the different chosen algorithms was performed and the results presented.

Keywords— Telecommunications, optical fiber, municipalities, exact and heuristics algorithm, Brazil

I. INTRODUÇÃO

Conforme o diagnóstico feito recentemente pela Anatel [1], temos hoje 4261 municípios atendidos com acesso à Internet por fibra óptica. No entanto, apesar do evidente avanço dos últimos anos (Figura 1), ainda há 1309 municípios sem esse acesso, que dependem de tecnologias ultrapassadas (ou com menor capacidade) para suas conexões à rede, como o cabo coaxial, o satélite e a transmissão por sinal de rádio.

Assim, falta ainda o investimento em fibra óptica no *backbone/backhaul* nacionais, para que todos os municípios tenham acesso de banda larga de qualidade à Internet.

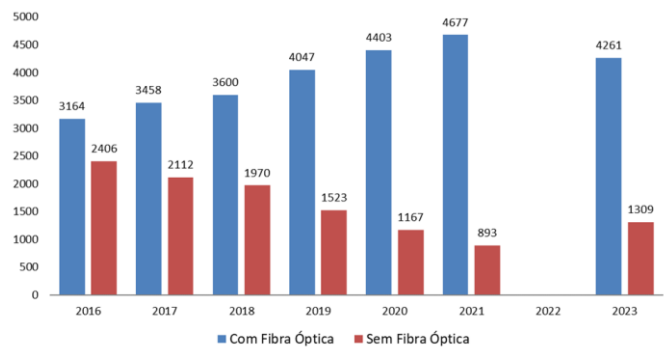


Fig. 1. Presença de fibra nos municípios
Fonte: Anatel [1]

No entanto, o tema é complexo. O Brasil tem grande extensão territorial, diversos municípios são pouco populosos e/ou com baixo IDH (Índice de Desenvolvimento Humano) e não são em territórios contíguos. Soma-se a isso o fato que os investimentos em infraestrutura de rede dependem das empresas privadas/públicas que detêm as concessões de serviços, temos então que os custos desses investimentos têm um papel fundamental na decisão de fazer ou não esses esforços.

O Plano Estrutural de Redes de Telecomunicações da Anatel [1], em sua avaliação anual, observa que há correspondência entre a presença das rotas de fibra óptica nos municípios e a velocidade e quantidade de acessos desses (Figura 2).

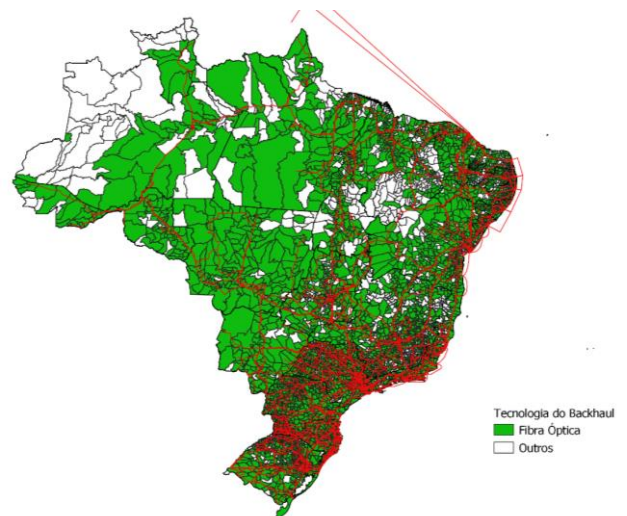


Fig. 2. Rede de backbone/Backhaul e municípios atendidos
Fonte: Anatel [1]

Portanto, dadas a extensão territorial, a sensibilidade de custos para os investimentos, as múltiplas possibilidades de conexões entre os municípios e a demanda crescente e constante por mais e melhores conexões à Internet, há a necessidade por soluções que busquem atender à demanda, porém sem descuidar dos custos necessários para tais investimentos. E é nesse ponto que se encaixa este estudo.

A proposta desse estudo é buscar, através da aplicação do problema da Árvore Geradora Mínima (AGM) a essa questão, qual seria a malha que conectasse todos os municípios do Brasil com o menor custo possível, bem como comparar os resultados da aplicação dos métodos exato e heurístico.

Os resultados poderão ser aplicados para a decisão de novas rotas, de manutenção ou substituição das rotas atuais, bem como na definição das políticas públicas para o incentivo ao desenvolvimento do setor, tão importante e estratégico para o Brasil.

II. JUSTIFICATIVA

A infraestrutura de *backbone* de fibra óptica do Brasil já foi objeto de estudos anteriores. Motta [8] procura aplicar a teoria dos grafos para compreender os nós de rede e suas interações, com a aplicação do algoritmo de Fruchterman-Reingold para identificar as centralizações da rede (Figura 3). A partir disso, resume a importância da expansão do acesso e da robustez da rede, de forma a garantir sua estabilidade.

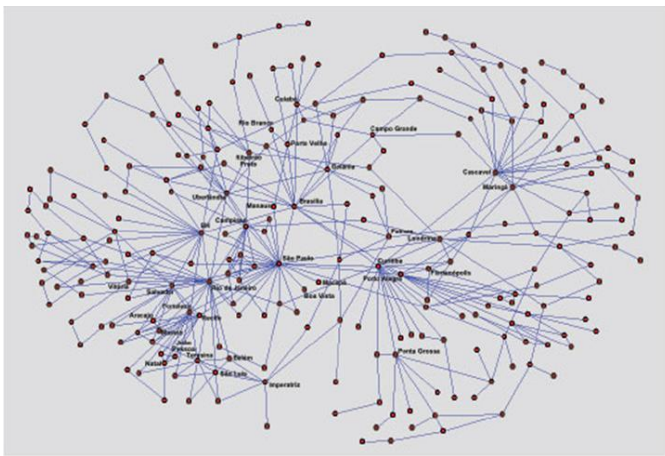


Fig. 3. Topologia da rede de *Backbones*
Fonte: Motta [8]

Outro estudo de relevância foi proposto pelo IPEA (Instituto de Pesquisa Econômica Aplicada) em 2017, em que os autores têm como propósito apresentar um modelo de avaliação econômica da expansão da rede de banda larga para os municípios brasileiros não conectados [3].

Esse estudo é particularmente interessante por também propor rotas de custo mínimo para a expansão do *Backhaul* de fibra óptica, como pretende este trabalho. Isso é feito através da teoria dos grafos [4, 6, 9], com a aplicação do problema da árvore geradora mínima, que consiste em encontrar, dado um grafo conexo e ponderado, uma árvore geradora de menor peso possível, em que o peso é dado pela soma dos pesos das arestas escolhidas (Figura 4).

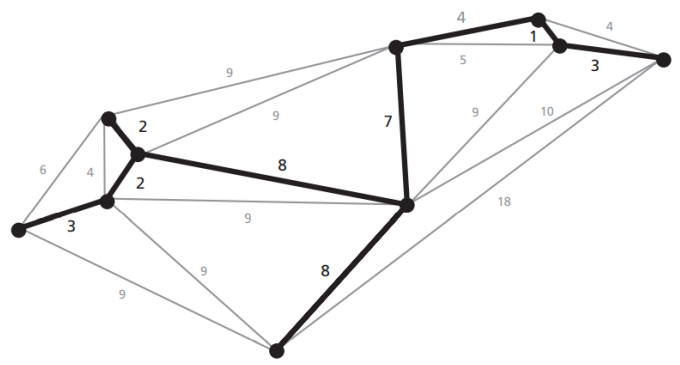


Fig. 4. Árvore Geradora Mínima
Fonte: Carvalho, Mendonça e Silva [3]

A partir dos dados obtidos, os autores buscaram aplicar a solução do problema da árvore geradora mínima para obter a rede de custo mínimo, que conecta todos os municípios sem fibra. Com isso, propuseram um grafo que represente uma rede factível de municípios conectados.

Assim, pode-se constatar que a bibliografia aponta métodos para a solução do problema de conexão dos municípios brasileiros que não possuem fibra óptica, que este trabalho procura atualizar e estender, de forma a subsidiar as políticas públicas no setor de telecomunicações do Brasil.

Este estudo propõe uma análise detalhada da viabilidade de expansão da rede de fibra óptica entre municípios brasileiros, utilizando o algoritmo de Prim [2], um método eficaz e comumente utilizado para resolver problemas de Árvore Geradora Mínima (AGM). Optamos por este algoritmo em vez do algoritmo de Kruskal [2], que também poderia ser utilizado para solucionar problemas de AGM, devido à sua eficiência na análise de grandes conjuntos de dados, o que se adequa à vasta quantidade de municípios brasileiros.

III. METODOLOGIA

Os dados utilizados nesse estudo foram obtidos nos sites da Anatel, IPEA e IBGE. Os dados são públicos e estão disponíveis nos painéis de dados das respectivas instituições.

O problema estudado (conectar os municípios brasileiros com fibra óptica) tem um enquadramento típico de Árvore Geradora Mínima, pois o que se busca é prover essas conexões com o menor custo possível, de forma que todos os municípios tenham pelo menos uma conexão à rede de *Backbone/Backhaul* de fibra óptica.

Para aplicar os métodos, foi necessário criar uma matriz com as distâncias entre os municípios brasileiros, obtidas na base georreferenciada do IBGE (SIGAS 2000). Para se gerar a matriz, foi necessário calcular as coordenadas das centróides de cada município da base, e a partir daí se calculou as distâncias geodésicas por algoritmos de precisão [5] entre cada coordenada. A matriz resultante tem 5570 linhas x 5570 colunas (Figura 5).

Os nomes das colunas e linhas são os códigos IBGE dos municípios, fundamentais para uma boa indexação dos resultados sem que possa haver algum problema com caracteres especiais, acentuação ou mudanças nos nomes dos municípios.

	1100015	1100023	1100031	1100049	1100056	1100064	1100072	1100080	1100098
1100015	0.00000	288.35937	209.32462	165.61638	136.72189	202.13264	138.05983	197.32952	204.10120
1100023	288.35937	0.00000	464.41482	232.78196	404.47796	441.34434	386.93016	271.02835	283.64911
1100031	209.32462	464.41482	0.00000	251.72436	73.68294	36.25255	77.95979	399.17483	235.44633
1100049	165.61638	232.78196	251.72436	0.00000	210.55230	222.12927	181.58190	312.38576	59.19779
1100056	136.72189	404.47796	73.68294	210.55230	0.00000	77.20077	35.71843	325.59709	211.34590
1100064	202.13264	441.34434	36.25255	222.12927	77.20077	0.00000	64.11094	397.23450	201.60992
1100072	138.05983	386.93016	77.95979	181.58190	35.71843	64.11094	0.00000	333.79721	177.43463
1100080	197.32952	271.02835	399.17483	312.38576	325.59709	397.23450	333.79721	0.00000	367.56219
1100098	204.10120	283.64911	235.44633	59.19779	211.34590	201.60992	177.43463	367.56219	0.00000
1100106	278.15788	228.73993	487.04480	350.75666	413.70796	479.84340	415.75059	106.73125	409.69384
1100114	210.44529	81.75553	382.77903	158.49389	322.84766	360.20662	305.21739	235.49749	213.89860
1100122	229.17349	143.08746	354.78957	104.16468	308.03409	326.04241	282.10447	312.84388	144.82226
1100130	362.42287	132.83669	495.53739	243.82721	449.91174	465.64270	423.82545	395.21360	272.82042
1100148	91.48950	201.19029	269.89065	119.90076	205.13696	252.01817	192.03920	201.62633	166.31095
1100155	206.00589	112.68126	358.11491	120.10417	303.54419	332.74295	281.99053	266.97671	172.38727
1100189	169.58768	316.57841	176.05891	84.92423	152.19206	143.18280	117.66408	352.65962	60.28523
1100205	428.89677	172.47571	623.43323	409.93896	557.95797	603.83027	545.49814	332.24578	456.01501
1100254	150.57324	178.87537	288.71872	62.23099	235.16170	263.59311	212.72358	261.37967	121.10194
1100262	317.51285	38.56152	484.20412	243.49150	427.51734	459.15708	407.59176	309.56604	289.47357
1100288	98.26991	235.80383	228.63059	68.20576	171.93756	206.24377	151.38572	253.32584	115.60061
1100296	79.35901	267.45402	197.32882	90.15542	138.66965	176.98713	119.55983	257.01880	124.75892
1100304	224.28887	379.26900	158.39220	146.59388	164.68886	122.13969	129.81317	414.80662	100.98478
1100320	122.07372	183.98048	327.40169	184.26592	256.81421	314.92377	251.79917	131.14004	241.62535
1100338	345.47379	189.28571	553.02218	375.34313	481.79216	540.33560	477.46838	204.93731	433.76628
1100346	131.51825	157.00050	316.53715	133.22463	251.65068	298.13394	238.64790	189.15300	192.19614
1100379	44.93670	314.97818	166.19969	155.27359	95.27334	157.38080	93.27160	241.95114	182.03438
1100403	335.14796	58.47331	518.95499	291.06640	456.90178	497.03898	441.14837	285.69491	342.10337
1100452	323.49329	108.44239	522.30003	317.84820	454.89678	504.82450	444.67745	232.11036	373.99374
1100502	89.74009	213.82365	252.93750	92.80574	190.67777	233.48748	174.98183	221.30326	146.26431
1100601	249.01392	42.12808	431.60973	211.08543	369.22565	410.34708	353.73264	232.43574	265.92926

Fig. 5. Trecho da matriz de distâncias dos municípios – elaboração dos autores

O resultado da aplicação dos métodos tem o formato (origem, destino, distância), onde origem e destino são municípios interligados com fibra óptica, e a distância é a distância que corresponde à quilometragem mínima a ser percorrida com a fibra.

Um dos objetivos do trabalho é propor uma solução exata e uma solução heurística para o problema. É importante destacar que a quantidade de dados é considerável (a matriz de distâncias possui 5570 linhas por 5570 colunas e todos os municípios foram considerados passíveis de conexão, o que resultou em 31.019.330 distâncias entre municípios), o que influenciou na escolha e avaliação dos algoritmos utilizados.

A solução exata será encontrada através do *solver* Gurobi a partir da formulação típica para o problema da AGM. A solução heurística será encontrada através do algoritmo de Prim, indicado para redes largas e de alta complexidade em termos de número de nós e conexões [7], programada no *Rstudio*.

Também propõe-se comparar os resultados encontrados com amostras de diferentes dimensões do problema. Para isso, serão calculados também a AGM para cada estado da federação, de forma independente, pelos dois métodos.

Assim, a solução do problema ocorrerá através dos seguintes passos:

1. Gerar a matriz de distâncias entre todos os municípios;
2. Carregar os dados para a aplicação do método exato no AMPL (uma linguagem de programação matemática para resolução de problemas de otimização);
3. Formular o problema no AMPL;
4. Solucionar o problema exato no AMPL, através do *solver* Gurobi;

5. Carregar os dados para a aplicação do método heurístico no *Rstudio*;

6. Programar o algoritmo de Prim no *Rstudio*;

7. Solucionar o problema no *Rstudio*;

8. Compilar a lista de municípios a ser atendidos no padrão (i, j, d), onde i e j são os municípios da ligação, e d é a distância entre as centróides dos municípios.

O computador usado para a execução dos códigos contém um processador AMD *Ryzen 5 5600X*, com 32 Gigabytes de memória RAM, com sistema operacional *Windows* instalado.

A partir dessa proposta, espera-se encontrar a lista de conexões de fibra ótica entre os municípios que têm essa carência e o custo mínimo (representado pela soma das distâncias resultantes a serem percorridas por cabos de fibra ótica) necessário para prover essas conexões. Importa destacar que esse é um problema com várias simplificações (por exemplo, o relevo não é considerado e as distâncias são tomadas a partir das centróides dos municípios), mas que trará boas diretrizes para estudos posteriores e para embasar a decisão das políticas públicas no setor.

A seguir são apresentados os resultados da implementação proposta.

IV. RESULTADOS

Ao se executar o algoritmo exato no AMPL, já era esperado que o tempo de processamento fosse significativamente maior que o algoritmo heurístico, que tem justamente como uma de suas grandes vantagens o tempo de execução. No entanto, a tarefa se mostrou inviável em grandes conjuntos de dados. O enorme volume de dados decorrente dessa aplicação levou à exaustão dos recursos computacionais à disposição, como a quantidade de memória disponível (figura 6).

```

Console
AMPL
ampl: solve;
### solve ...
#compile          0          0          767560
Error executing "solve" command:
error processing set 5:
    range too large
ampl:
    
```

Fig. 6. Mensagem de erro no AMPL ao tentar-se calcular a solução ótima

Dessa forma, uma grande vantagem do algoritmo heurístico já se confirmou, pois sua demanda por recursos computacionais é muito menor. Para permitir uma comparação entre os desempenhos e precisão dos algoritmos, foram então separados diversos conjuntos de dados, representados pelo Brasil e seus estados, com seus resultados na Figura 7.

Ao se verificar os resultados encontrados, percebe-se que o tempo de execução com o algoritmo exato aumenta muito mesmo nos pequenos estados onde foi possível sua execução. No estado do Acre, com somente 22 municípios, já foram necessários 606 segundos para encontrar a solução ótima de 1.379,32 quilômetros. A partir do estado de Rondônia, com somente 52 municípios, já se tornou inviável a execução do algoritmo exato, em que todas as tentativas de execução resultaram em mensagens de erro.

Já o algoritmo heurístico teve uma ótima performance em todos os conjuntos de dados. Ao se aplicar ao conjunto de todos os municípios do Brasil, o tempo de execução foi de 24,69 segundos (o tempo do algoritmo exato para somente o estado do Acre foi mais de 24 vezes maior). Na comparação direta nos três estados onde os dois algoritmos foram executados com sucesso, percebe-se que o algoritmo heurístico teve a mesma precisão do exato, encontrando também a solução ótima.

	Municípios	Solução Ótima Exato (km)	Solução Ótima Heurístico (Km)	Tempo Exato (segundos)	Tempo Heurístico (segundos)
Brasil	5570	-	129.460,80	-	24,69
Minas Gerais	853	-	16.461,46	-	0,31
São Paulo	645	-	9.770,34	-	0,16
Rio Grande do Sul	497	-	8.087,10	-	0,17
Bahia	417	-	10.942,50	-	0,07
Paraná	399	-	6.978,87	-	0,06
Santa Catarina	295	-	4.108,83	-	0,01
Goiás	246	-	6.700,65	-	0,01
Piauí	224	-	5.661,40	-	0,02
Paraíba	223	-	2.882,05	-	0,01
Maranhão	217	-	6.180,76	-	0,01
Pernambuco	185	-	3.760,21	-	0,02
Ceará	184	-	3.998,33	-	0,02
Rio Grande do Norte	167	-	2.359,98	-	0
Pará	144	-	7.956,76	-	0,02
Mato Grosso	141	-	8.371,30	-	0
Tocantins	139	-	4.554,50	-	0
Alagoas	102	-	1.358,63	-	0
Rio de Janeiro	92	-	1.586,38	-	0
Mato Grosso do Sul	79	-	3.920,60	-	0
Espírito Santo	78	-	1.497,87	-	0
Sergipe	75	-	993,88	-	0
Amazonas	62	-	6.977,56	-	0
Rondônia	52	-	2.422,06	-	0
Acre	22	1.379,32	1.379,32	606,13	0
Amapá	16	1.023,44	1.023,44	2,36	0
Roraima	15	1.358,08	1.358,08	0,81	0

Fig. 7. Resultados dos algoritmos exato e heurísticos – elaboração dos autores.

Dessa forma, a solução encontrada pelo algoritmo heurístico para todo o Brasil é que são necessários 129.460,80 km de fibras ópticas para interligar todos os municípios brasileiros (Figura 8).

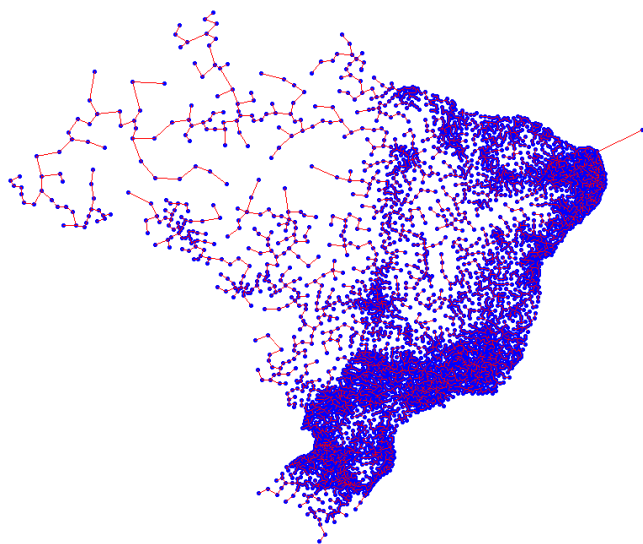


Fig. 8. Menor rota de fibra para conectar todos os municípios do Brasil – elaboração dos autores.

No entanto, é importante considerar os municípios que já possuem conexão com fibra, pois não é economicamente viável reconectar todos os municípios do Brasil com fibra. Assim, a partir dos dados da Anatel sobre os municípios atualmente

conectados, foi elaborado o grafo com a indicação dos novos enlaces (Figura 9). Foram calculados 53.809,57 km necessários com fibra para essas novas conexões.

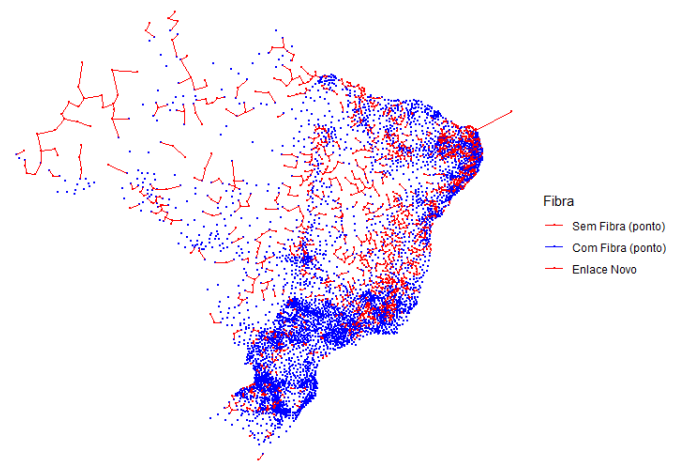


Fig. 9. Menor rota para conectar os municípios sem fibra do Brasil – elaboração dos autores.

Os resultados demonstram um potencial significativo para a expansão da infraestrutura de fibra óptica, o que traria vários benefícios práticos. Por exemplo, uma maior conectividade de banda larga poderia facilitar a digitalização de vários setores da economia, aumentando a produtividade e melhorando o acesso a serviços digitais para os cidadãos. Além disso, uma expansão bem planejada da rede de fibra óptica poderia ajudar a reduzir a disparidade no acesso à internet de alta velocidade entre diferentes regiões do Brasil.

Este estudo contribui para a literatura existente, oferecendo uma metodologia robusta e um estudo de caso detalhado sobre a expansão da infraestrutura de fibra óptica. Ele apoia estudos anteriores que destacam a importância do planejamento estratégico na expansão de redes de telecomunicações, mas também oferecem novas perspectivas ao aplicar o algoritmo de Prim neste contexto.

CONCLUSÕES

O estudo realizado confirmou que é possível se encontrar a solução ótima, através da aplicação de algoritmos, para se determinar o custo mínimo necessário para interligar todos os municípios do Brasil.

Os algoritmos escolhidos para a comparação foram um exato, programado na AMPL, e um heurístico (Algoritmo de Prim), programado no *R Studio*.

Os algoritmos tiveram a mesma precisão nos modelos onde ambos puderam ser executados, favorecendo sobremaneira o algoritmo de Prim no tempo de execução. Além disso, grupos maiores de municípios excederam a capacidade do software e do hardware utilizados para a execução do algoritmo exato, somente viáveis através do algoritmo heurístico.

A solução ótima encontrada foi de 129.460,80 km para interligar todos os municípios brasileiros com fibra óptica, bem como 53.809,57 km necessários com fibra para as novas conexões. Também foram apresentados resultados em grupos menores de municípios, separados pelos estados aos quais pertencem.

Dessa forma, conclui-se que os modelos utilizados são muito úteis para o cálculo de soluções de otimização para o setor de telecomunicações, e podem ser ferramentas muito úteis para auxiliar na tomada de decisões de investimentos e na proposição de políticas públicas para o setor.

Em conclusão, este estudo destaca a importância de uma análise estratégica cuidadosa na expansão da infraestrutura de fibra óptica, oferecendo uma nova perspectiva sobre esta questão através da aplicação do algoritmo de Prim. Com as implicações práticas do estudo claramente delineadas, esperamos que nossa pesquisa possa contribuir para os esforços de expansão da infraestrutura digital no Brasil, resultando em benefícios significativos para a economia e a sociedade brasileira.

Como forma de estender esse trabalho, propõe-se que em estudos posteriores sejam excluídas algumas simplificações adotadas, como o relevo, rotas facilitadas para as fibras, como estradas e leitos de rios, a posição geográfica das cidades e dos terminais de possíveis conexões das fibras ópticas, o teste de outros algoritmos (como o de Kruskal) e a inclusão de novos dados, como as operadoras responsáveis por cada enlace.

REFERÊNCIAS

- [1] BRASIL **Plano Estrutural de Redes de Telecomunicações (Edição 2023)**. Brasília, DF, 2023. <https://www.gov.br/anatel/pt-br/dados/infraestrutura/pert-1>.
- [2] CORMEN, Thomas H. et al. **Introduction to algorithms**. MIT press, 2022.
- [3] DE CARVALHO, Alexandre Ywata; DE MENDONÇA, Mário Jorge Cardoso; DA SILVA, José Jaime. **Avaliação de priorizações para expansão da rede de acesso à banda larga nos municípios brasileiros. Texto para Discussão**, 2017.
- [4] HARARY, F. **Graph theory**. Boston: Addison-Wesley, 1969.
- [5] KARNEY, Charles FF. **Algorithms for geodesics**. Journal of Geodesy, v. 87, n. 1, p. 43-55, 2013.
- [6] LUCCHESI, C. L. **Introdução à teoria dos grafos**. In: Colóquio Brasileiro De Matemática Do Instituto Nacional De Matemática Pura E Aplicada (IMPA), 12., 1979, Rio de Janeiro. Anais... Rio de Janeiro: Impa, 1979.
- [7] MARIANO, Artur et al. **Hardware and software implementations of Prim's algorithm for efficient minimum spanning tree computation**. In: International Embedded Systems Symposium. Springer, Berlin, Heidelberg, 2013. p. 151-158.
- [8] MOTTA, Marcelo Paiva da. **Topologia dos backbones de internet no Brasil**. Sociedade & Natureza, v. 24, p. 21-35, 2012.
- [9] SZWARCFITER, J. L. **Grafos e algoritmos computacionais**. 2. ed. Rio de Janeiro: Editora Campus Ltda., 1986.