

Planejamento de Cobertura de Redes 5G utilizando Inteligência Artificial através de programa de Geolocalização, RNA e Bayesianas

Ytalo de J.A.C. Sousa, Alex S. Macedo, Lizandro D. A. Zenteno, Caio M.M. Cardoso, Brenda S. Barbosa, Fabrício J. B. Barros, Jasmine P. L. Araújo.

Resumo— Este trabalho contribui para a formulação de uma metodologia destinada a propor um modelo para a predição de perda de percurso em um ambiente misto, composto por áreas urbanas arborizadas e/ou com floresta. Com base na rede *FeedForwardNet* usando *MATLAB*, empregando diversas funções de treino, arquiteturas e números de neurônios, visando otimizar o desempenho do modelo na predição da perda de percurso em ambientes mistos. O modelo é dividido em camadas e recorre a informações de geolocalização, perdas devido à ausência de visada direta, densidade urbana e densidade florestal para a predição da perda de percurso. Dentro desse modelo, também é aplicada a otimização Bayesiana. O modelo apresentou resultados promissores, com um erro médio quadrático de 0,037 dB e coeficiente de correlação médio acima de 0,8.

Palavras-Chave— 4G-LTE, 5G, predição da perda de percurso, redes Bayesianas, redes neurais, região amazônica.

Abstract— This work contributes to the formulation of a methodology designed to propose a Path Loss prediction model in a mixed environment, composed of urban areas with trees and/or forest. Based on the *FeedForwardNet* network using *MATLAB*, employing various training functions, architectures and numbers of neurons aiming to optimize the model's performance in predicting path loss in mixed environments. The model is divided into layers and uses geolocation information, losses due to lack of direct sight, urban density and forest density to predict Path Loss. Within this model, Bayesian optimization is also applied. The model showed promising results, with a mean squared error of 0.037 dB and an average correlation coefficient above 0.8.

Keywords— 4G-LTE, 5G, predicting Path Loss, Bayesian networks, neural networks, Amazon region.

I. INTRODUÇÃO

A transição para o sistema 5G implica uma significativa evolução na Qualidade de Serviço (QS) oferecida aos usuários e na arquitetura geral do sistema. Com essas mudanças, as operadoras de redes móveis precisam planejar para atender às novas demandas dos usuários 5G, enquanto garantem que os requisitos dos usuários LTE não sejam comprometidos devido ao compartilhamento da infraestrutura, característica da transição [1]. Para esse fim, são aplicados diversos modelos de Perda de Percurso (em inglês, Path Loss - PL), embora estes não consigam abranger completamente a complexidade dos ambientes mistos como será mostrado na comparação de modelos clássicos e o proposto.

Neste estudo, foram realizadas campanhas de *Drive Test*, amplamente utilizadas para coleta de dados, como evidenciado

Os autores pertencem ao Laboratório de Computação e Telecomunicações, Universidade Federal do Pará, Belém-Pará, Email: ytalocontato@gmail.com

em vários artigos [3]–[6]. Essas campanhas utilizaram o *G-NetTrack* para analisar a QS e coletar parâmetros de avaliação como RSRP, RSRQ e SINR. O RSRP, em particular, é utilizado para calcular a PL.

Recentemente, diversos estudos têm aplicado IA para predição em ambientes mistos, empregando múltiplas arquiteturas e variáveis de desempenho [7], geolocalização como entrada [8], e linha de visada como entrada para redes neurais [2]. Esses estudos utilizam RSRP e diferentes arquiteturas para prever a PL em redes 4G e 5G. No entanto, há uma carência de estudos que considerem o ambiente amazônico com suas características específicas.

Neste contexto de estudar um ambiente misto e aplicar modelos de Perda de Percurso usando Inteligência Artificial (IA), este trabalho realizou a coleta de dados por meio de campanhas de medição, pré-processou e analisou os níveis de sinal recebidos em redes 4G e 5G. Foram coletadas métricas de RSRP, RSRQ e SINR, utilizando uma Estação Rádio Base (ERB) situada em um bairro central de Belém.

Por fim, a rede neural foi aplicada e otimizada com base nos parâmetros do cenário e no modelo proposto com sua função de custo. A Perda de Percurso calculou-se a partir dos dados medidos [2]. Isso permite predizer áreas de sombra e avaliar a adequação do sinal em regiões urbanas e vegetadas, possibilitando a implementação de Small Cells para melhorar a cobertura.

Resumindo as principais contribuições desse trabalho são:

- Planejamento de Cobertura utilizando Inteligência Artificial (IA);
- Propor uma metodologia para calcular e predizer a perda de propagação em ambientes mistos;

II. METODOLOGIA DA CAMPANHA DE MEDIÇÕES

A metodologia utilizada neste trabalho será descrita nesta seção, conforme apresentado no fluxograma da Figura 1.

Para a campanha de medições, selecionou-se uma área da cidade de Belém do Pará, contemplando cinco bairros: Nazaré, Umarizal, São Brás, Guamá e Cremação. Foram coletados dados de nível de sinal 4G e 5G em um raio de 1 km, abrangendo várias bandas e larguras de banda. A rota definida é ilustrada na Figura 3, gerada pelo *Google Earth*, onde a ERB indicada por PABL66 opera em 3,311 GHz.

Durante a coleta, utilizou-se o método *Drive Test* a uma velocidade média de 5,56 m/s, empregando o programa *G-NetTrack Pro* [9]. Utilizaram-se dois *smartphones* com chips

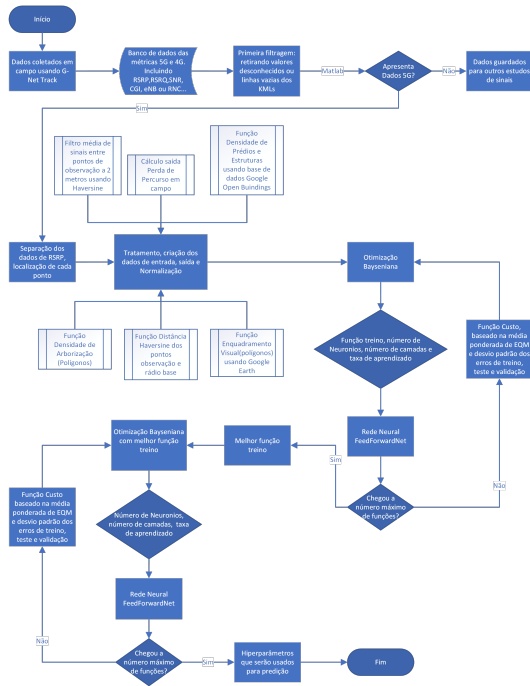


Fig. 1. Fluxograma total das fases do experimento

diferentes, ambos com suporte ao 5G, conforme ilustrado na Figura 2. No total, obtiveram-se 17.499 amostras válidas, após 10 voltas na mesma rota. Os dados registrados pelo *G-NetTrack* foram armazenados na memória dos dispositivos e, posteriormente, extraídos e analisados [5], [6].



Fig. 2. Software *G-NetTrack* para a coleta de dados. Fonte: Autor

III. MODELOS DE PROPAGAÇÃO

Os modelos de PL são fundamentais no planejamento de redes sem fio. Eles ajudam a determinar a quantidade de potência de transmissão necessária para alcançar a cobertura desejada e a configurar a infraestrutura da rede [10], além de serem essenciais para a otimização [11].

Portanto, para fins de comparação dos resultados, serão plotados os modelos clássicos junto ao modelo gerado utilizando IA. Os modelos utilizados incluem: Modelo *Free-space* [12], Modelo ABG (Alfa, Beta e Gama) [13] e Modelo Ericsson [14].

IV. SOFTWARE DE PLANEJAMENTO

A metodologia de planejamento adotada é a divisão por camadas, conforme apresentada na Figura 4. Esta abordagem é inspirada nos métodos utilizados por *softwares* de GIS (Sistema de Informação Geográfica) [15].



Fig. 3. Rota percorrida. Fonte: Autor

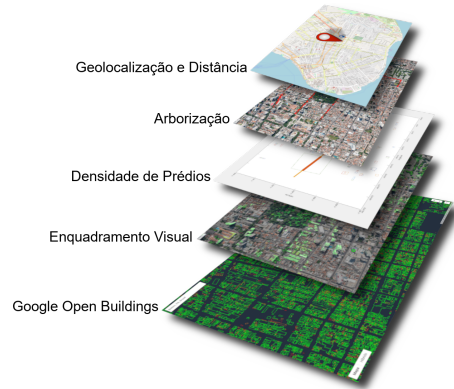


Fig. 4. Divisão em Camadas. Fonte: Autor

1) *Camada de Geolocalização*: Esta camada é utilizada em metodologias que predizem a perda de percurso pela distância [8]. Nela, serão realizados cálculos das distâncias entre o ponto de observação e a ERB.

2) *Camada de Arborização*: A divisão utiliza o *software Google Earth* [17] e sua posterior adição aos dados no *MATLAB*® [18]. Esta camada também é considerada no cálculo de PL em alguns modelos de propagação [19]. Ela é dividida em duas partes: uma referente à rua, onde há elevada arborização (*densidadeArvores*), e outra referente à área entre o observador e a ERB (*densidadeParksFlorestais*). Ambas as camadas podem ser vistas na Figura 5.

3) *Camada Densidade de Prédios*: Utilizando os dados coletados do *Google Open Buildings* [21], esta camada é usada para calcular a quantidade de prédios na linha de visada entre a antena e o receptor, conforme mostrado na Figura 5. A perda por propagação ocorre devido a vários fenômenos na propagação do sinal [20], justificando a inclusão dessa variável na rede (*densidadePredios*).

4) *Camada de Enquadramento Visual*: Utilizou-se o *Google Earth* e sua ferramenta de Enquadramento Visual, que, a partir de um determinado ponto e altura, gera o *viewshard*. Esta métrica também é usada para a modelagem do PL em outros trabalhos [22]. A Figura 6 mostra sua simplificação com polígonos do enquadramento visual gerados pelo *Google Earth*.

5) *Camada de dados Google Open Buildings*: A camada *Open Buildings* [21] contém os vértices e formas dos prédios (planta baixa), incluindo casas, obtidos a partir de imagens de satélite de alta resolução. Destas estruturas, foram utilizados apenas os prédios com mais de 4 andares, que têm maior im-

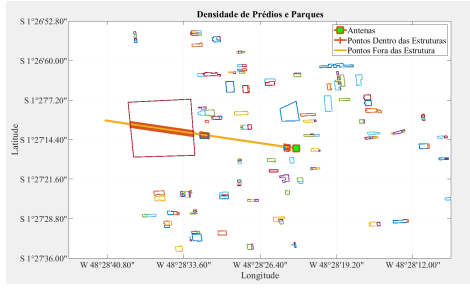


Fig. 5. Exemplo de contagem de passo dentro das estruturas, prédios e parques florestais. Fonte: Autor

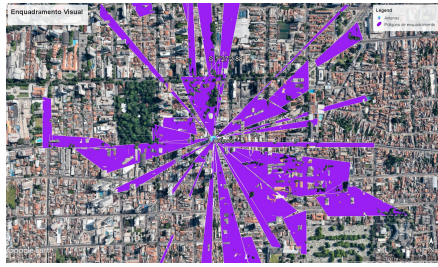


Fig. 6. Enquadramento visual. Fonte: Autor

pacto no *viewshared* da rádio base, totalizando 169 polígonos.

V. INTELIGÊNCIA COMPUTACIONAL E TÉCNICAS

Neste trabalho, utilizam-se técnicas inspiradas em *software* proprietário que usa camadas para determinar o PL [15], além da aplicação de Redes Neurais Artificiais (RNA), especificamente a *FeedForwardNet* do *MATLAB* com suas diversas funções de treino [25]. Para detalhar as técnicas e metodologias empregadas, o estudo é dividido nas seguintes subseções:

A. Processamento de dados por camadas

Perante os dados de 5G já separados na Figura 7, inicia-se o cálculo da média do PL entre pontos de observação próximos, em um raio de 2 metros, utilizando o método de Haversine [23]. Isso resulta em uma redução de 75% das amostras, diminuindo o erro médio quadrático e atenuando valores extremos ou muito baixos (*outliers*). Ademais, o método é empregado para calcular a distância entre a ERB e o usuário.

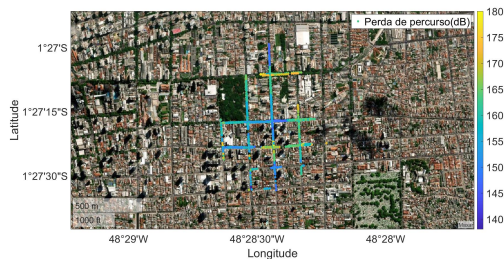


Fig. 7. Dados tratados usados para saída da rede de perda de percurso do Node escolhido. Fonte: Autor

Quanto à saída, o PL foi calculado para cada ponto de medição usando a equação simplificada proposta por [24], conforme

mostrado na Equação 1, seguindo trabalhos semelhantes [14]. A versão simplificada pode ser vista na Equação 2.

$$RSRP_{(dBm)} = P_{RS} + G_{BS} - \overset{0}{ATX} + \overset{0}{GUE} - PL, \quad (1)$$

Substituindo os valores com os dados da antena, como ganho G_{BS} de 24 dBi e potência de 200 W, obtém-se a Equação (2).

$$PL_{(dB)} = 10 \log_{10} \left(\frac{200W}{0,001} \right) + 24 - RSRP_{(dBm)}; \quad (2)$$

A entrada da rede segue conforme a sub-rotina apresentada na Equação (3).

$$X_{entradas} = [densidadeParksFlorestais, \quad densidadePredios, long, lat \quad (3) \\ , viewShared, densidadeArvores]$$

O PL é a saída da rede neural, normalizada com uma função própria, variando de -1 a 1.

B. Bayesiana como otimizador

Essa técnica de otimização [26] utilizou a função de custo, treinando uma rede *FeedForwardNet* a cada iteração (200). As entradas são os hiperparâmetros descritos na subseção correspondente deste artigo. A saída dessa função segue conforme apresentado na Equação 4.

$$\begin{aligned} \text{Custo} &= \text{média}(A) + \text{média}(B), \quad \text{onde} \\ A &= 1 \cdot \text{mse}(\text{erroTreino}) \\ &+ 2 \cdot \text{mse}(\text{erroTeste}) \\ &+ 1 \cdot \text{mse}(\text{erroValidação}), \quad (4) \\ B &= 1 \cdot \text{desvio padrão}(\text{erroTreino}) \\ &+ 2 \cdot \text{desvio padrão}(\text{erroTeste}) \\ &+ 1 \cdot \text{desvio padrão}(\text{erroValidação}) \end{aligned}$$

C. Sobre os hiperparâmetros

Os hiperparâmetros otimizados e testados incluem: número de neurônios (variando de 1 a 30 na busca pela melhor função e de 1 a 80 na busca pela função de treino vencedora), número de camadas ocultas (*hidden layers*, variando de 1 a 4 camadas), taxa de aprendizado da rede (variando de 0,1 a 0,0001) e as funções de treino utilizadas, conforme descrito em [25].

VI. RESULTADOS E DISCUSSÃO

Com as métricas coletadas, processou-se um número considerável de dados filtrados para o caso de estudo, totalizando 4.482 amostras. Estas, separadas por *ENodeID*, são: 1.779 amostras no *ENodeID* 910395, 1.582 amostras no *ENodeID* 910430 e 433 amostras no *ENodeID* 913202. Alguns dados foram coletados principalmente na área de interesse de sombra da rede. Selecionou-se o *ENodeID* 910430 com *CellID* igual a 31.

Ao final da predição, gera-se o gráfico para a rede com uma camada oculta (*hidden layer*), conforme mostrado na Figura

8. No gráfico, "PL Entrada" representa o PL utilizada para treinar a rede neural, enquanto "PL Predição" é a predição da rede neural para a PL.

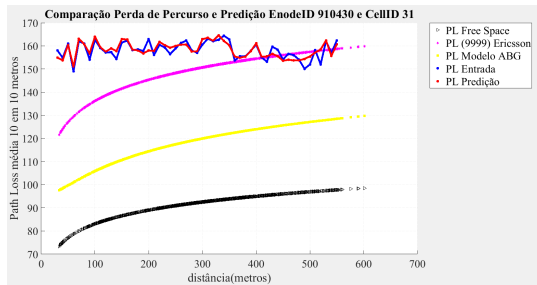


Fig. 8. Usando uma camada oculta e modelos de PL. Fonte: Autor

A. Perda de percurso no caso geral com uma rede de uma camada oculta

Os resultados são provenientes de uma otimização Bayesiana para a escolha dos hiperparâmetros, conforme mostrado na Figura 9.

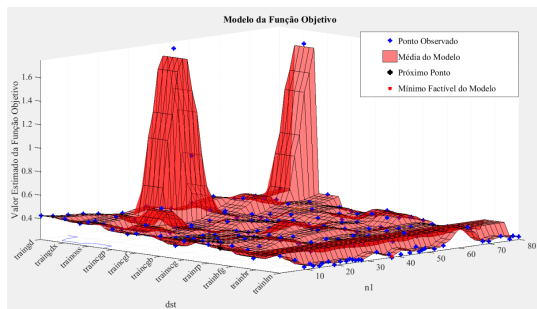


Fig. 9. Plano criado pela Bayesiana no espaço das variáveis usadas. Fonte: Autor

A função de custo combina o erro quadrático médio (EQM) e o desvio padrão ponderado para capturar a magnitude média dos erros e sua variabilidade. Apesar das unidades distintas (dB^2 para EQM e dB para desvio padrão), a métrica composta mostrou-se eficaz, na prática, guiando o *Bayesopt* para minimizar erros relativos e reduzir os erros globais de treinamento, teste e validação, comprovando sua adequação.

Na Figura 9, pode-se observar que cada função de treino (*dst*) da rede *FeedForwardNet* do *MATLAB* testou uma camada oculta contendo de 1 a 80 neurônios (*n1*, número escolhido devido às limitações de hardware para uma única camada oculta). Consequentemente, formou-se uma superfície que busca os melhores valores para minimizar a função de custo projetada.

Na mesma Figura 9, também é mostrada a superfície gerada pela otimização Bayesiana, que minimiza a função de custo. O tempo de conclusão ocorreu em 463,89 segundos. O resultado está na geração da arquitetura de rede apresentada na Figura 10, onde a melhor função de treino foi a *trainlm*. No geral, a rede apresentou um erro médio quadrático de 0,037343 nos dados de entrada.

Após a otimização, os valores encontrados são aplicados na função *FeedForwardNet*, e os resultados podem ser observados

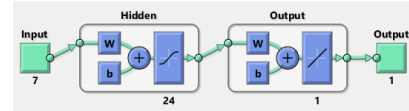


Fig. 10. Estrutura otimizada de apenas uma camada oculta. Fonte: Autor

na Figura 12. Na Figura 11, verifica-se que, ao longo das épocas, o treinamento é interrompido ao atingir um dos hiperparâmetros fixos: 6 erros de validação.

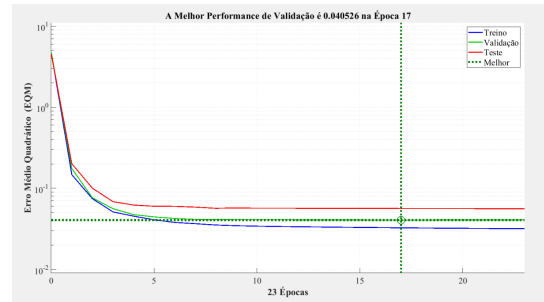


Fig. 11. Função performance da rede com uma camada. Fonte: Autor

Na regressão, conforme mostrado na Figura 12, observou-se um bom aprendizado da função de treino, evidenciado por um valor de 0,81971. Além disso, o teste apresentou um valor próximo, atingindo 0,74716. Vale ressaltar que esse desempenho foi alcançado com apenas uma camada e uma única otimização. Como resultado, obteve-se uma média final de 0,80188.

Aplicando essa rede treinada aos pontos de sombra escolhidos da rede 5G, chega-se à predição mostrada na Figura 13.

Os resultados mostram uma maior perda por propagação em predições mais distantes que envolvem prédios ou vegetação, conforme indicado nas camadas de entrada da rede para o treino. Observam-se altos valores de PL preditos, o que justificaria a mudança da tecnologia 5G para 4G no local (região de sombra).

Por outro lado, podemos comparar o sinal original e o PL da predição na região de sombra do 5G, como mostrado na Figura 14. Em ambos os gráficos, observa-se uma maior perda de propagação nessa região, onde não foram registrados valores do 5G com atenuações maiores.

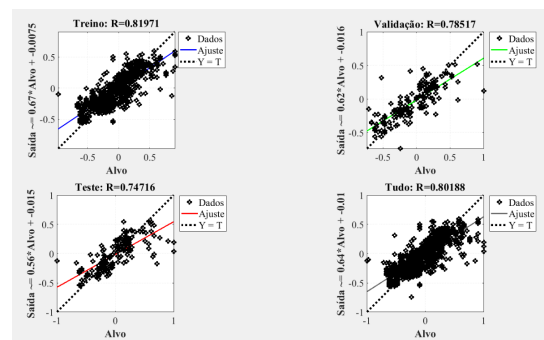


Fig. 12. Uma camada e sua regressão. Fonte: Autor

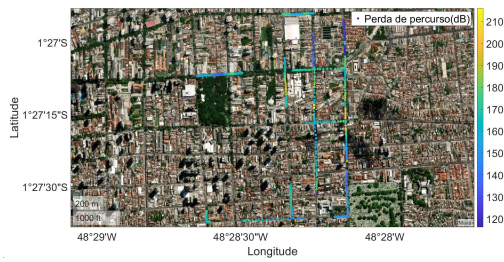


Fig. 13. Predição da perda de percurso com uma camada oculta. Fonte: Autor

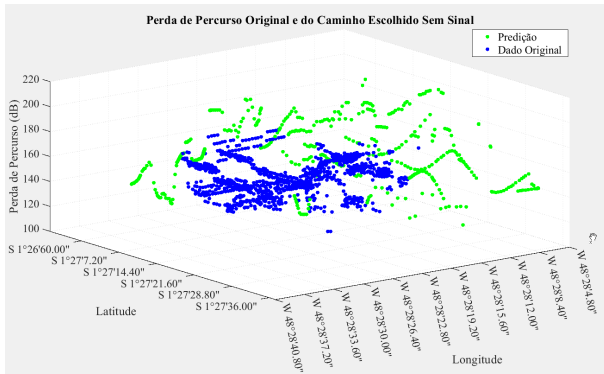


Fig. 14. Apenas uma camada. Entrada da rede e predição

VII. CONCLUSÕES

O desempenho da rede neural mostrou-se mais eficiente na predição do PL em comparação com outros modelos, conforme visto na Figura 8. Para além, a otimização Bayesiana demonstrou ser eficaz na escolha dos melhores hiperparâmetros para a rede, resultando em boas métricas de desempenho, com um erro médio quadrático de 0,037343 dB e valores significativos na regressão, apresentando um coeficiente de correlação médio acima de 0,8.

A metodologia proposta pode ser aplicada a outras localidades, permitindo a escalabilidade na demanda da rede celular e a expansão dos estudos para cenários mais complexos e robustos. No entanto, é necessária uma nova análise ao treinar a rede neural com uma nova Estação Rádio Base, sendo crucial criar uma nova *viewshared* e realizar o treinamento novamente. Para trabalhos futuros, pretende-se avaliar outros parâmetros, como *throughput*, *downlink*, *uplink* e camadas mais profundas da rede neural em redes 5G.

REFERÊNCIAS

[1] C. Henrique, A. Jr., and L. Pereira, “Coexistência entre 5g nr e lte-a em sistemas fiber/wireless,” in Anais de XXXVIII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, ser. SBrt2020, 2020, acesso em: 30 de novembro de 2023. [Online]. Available: <https://www.sbrt.org.br/sbrt2020/papers/1570649771.pdf>

[2] H. Cheng, S. Ma, H. Lee, and M. Cho, “Millimeter wave Path Loss modeling for 5g communications using deep learning with dilated convolution and attention,” IEEE Access, vol. 9, pp. 62 867–62 879, 2021.

[3] Widyasmoro, I. Surahmat, et. al., “Comparative performance analysis of 4g and 5g cellular network technology in indonesia: Case study in the city of jakarta,” in 2022 2nd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), pp. 158–163, 2022.

[4] M. A. Al Jahdhami, et. al., “Performance analysis of mobile broadband networks in ibra city, oman,” in 2021 International Conference on Artificial Intelligence and Big Data Analytics, pp. 1–6, 2021.

[5] V. Filho, Ytalo Sousa, et. al., “Análise Comparativa de Cobertura e Qualidade entre as Redes Celulares 4G e 5G em Belém-PA,” in Anais da XIII Conferência Nacional em Comunicações, Redes e Segurança da Informação, ENCOM, 2023.

[6] V. Filho, et.al, “Análise de Cobertura e Qualidade de Rede Celular 5G: Estudo de Caso de Belém-PA,” in Anais da XIV ENCOM, 2023.

[7] L. Wu, D. He, B. Ai, J. Wang, H. Qi, K. Guan, and Z. Zhong, “Artificial neural network based Path Loss prediction for wireless communication network,” IEEE Access, vol. 8, pp. 199 523–199 538, 2020.

[8] S. I. Popoola, et. al., “Determination of neural network parameters for Path Loss prediction in very high frequency wireless channel,” IEEE Access, vol. 7, pp. 150 462–150 483, 2019.

[9] G. Solutions. (2023) *G-nettrack Pro*. Acesso em: 30 de novembro de 2023. [Online]. Available: <https://gykovsolutions.com/>

[10] A. Ikpehai, et. al., “Low-power wide area network technologies for internet-of-things: A comparative review,” IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2225–2240, Apr. 2019.

[11] P. Tang, et. al., “Channel measurement and Path Loss modeling from 220 ghz to 330 ghz for 6g wireless communications,” China Communications, vol. 18, no. 5, pp. 19–32, May 2021.

[12] M. Veleli c. and M. Šunjevari c, “On the cramer-rao lower bound for rss-based positioning in wireless cellular networks,” AEU - International Journal of Electronics and Communications, vol. 68, no. 8, pp. 730–736, Aug. 2014.

[13] S. Sun, et. al., “Propagation Path Loss models for 5g urban micro-and macro-cellular scenarios,” in 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), May 2016.

[14] Z. E. Khaled, W. Ajib, and H. Mcheick, “An accurate empirical Path Loss model for heterogeneous fixed wireless networks below 5.8 ghz frequencies,” IEEE Access, vol. 8, pp. 182 755–182 775, 2020.

[15] Pathloss. (2023) Pathloss 6. Acesso em: 30 de novembro de 2023. [Online]. Available: <https://www.pathloss.com/pathloss6.html>

[16] Y. Wei, et. al., “Automatic distributing schemes of physical cell identity for self-organizing networks,” International Journal of Distributed Sensor Networks, vol. 8, no. 10, p. 973713, Oct. 2012.

[17] Google. (2023) Google earth. Acesso em: 30 de novembro de 2023. [Online]. Available: <https://www.google.com/earth/>

[18] “Matlab,” <https://www.mathworks.com/products/matlab.html>, acesso em: 30 de novembro de 2023.

[19] J. C. Silva, G. L. Siqueira, and P. V. G. Castellanos, “Propagation model for Path Loss through vegetated environments at 700 – 800 mhz band,” Journal of Microwaves, Optoelectronics and Electromagnetic Applications, vol. 17, no. 1, pp. 179–187, Mar. 2018.

[20] E. Promise and O. P. O., “Comparison of empirical Path Loss propagation models with building penetration Path Loss model,” International Journal on Communications Antenna and Propagation (I.Re.C.A.P), vol. 6, no. 2, 2016.

[21] Google. (2023) Google open buildings. Acesso em: 30 de novembro de 2023. [Online]. Available: <https://sites.research.google/open-buildings/>

[22] M. K. Samimi, T. S. Rappaport, and G. R. MacCartney, “Probabilistic omnidirectional Path Loss models for millimeter-wave outdoor communications,” IEEE Wireless Communications Letters, vol. 4, no. 4, pp. 357–360, Aug. 2015.

[23] V. Patil and P. K. Atrey, “Geosecure-r: Secure computation of geographical distance using region-anonymized gps data,” in 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM), Sept. 2020.

[24] M. Sousa, A. Alves, P. Vieira, M. P. Queluz, and A. Rodrigues, “Analysis and optimization of 5g coverage predictions using a beamforming antenna model and real drive test measurements,” IEEE Access, vol. 9, pp. 101 787–101 808, 2021.

[25] The MathWorks, Inc. (1994–2023) feedforwardnet - generate feedforward neural network. © 1994–2023 The MathWorks, Inc. [Online]. Available: <https://www.mathworks.com/help/deeplearning/ref/feedforwardnet.html>

[26] MathWorks. (2023) Bayesian optimization in matlab. Acesso em: 30 de novembro de 2023. [Online]. Available: <https://www.mathworks.com/help/stats/bayesopt.html>