

PREVISÃO DA RELAÇÃO SINAL/RUÍDO EM TELEFONIA MÓVEL VIA MODELAGEM COMPUTACIONAL

Palavras-Chave: CORRELAÇÃO ESTATÍSTICA, RELAÇÃO SINAL/RUÍDO,
SISTEMAS DE COMUNICAÇÃO

Autores:

LARA TRABACHINI, FT - UNICAMP

Prof. Dr. DIEGO SAMUEL RODRIGUES, FT - UNICAMP

INTRODUÇÃO:

Na pesquisa científica, é comum o uso de modelos matemáticos ou computacionais baseados em equações paramétricas hipotéticas para o estudo de processos e sistemas. Todavia, o uso desses modelos pode ser inviável devido a dinâmicas desconhecidas ou demasiadamente complexas. Dessa maneira, a utilização de modelos empíricos que inferem padrões e associações a partir de dados existentes representam uma abordagem alternativa altamente flexível [1]. A partir dessas abordagens, é possível inferir previsões, projeções e também buscar relações causais ou influência dinâmica entre variáveis. Para isso, antes da previsão, esse trabalho desenvolve uma análise de correlações de variáveis em dados de telefonia móvel disponibilizados de forma pública denominados de “4G LTE User Equipment Measurements along Kingston Transit 502 Bus Route” [2].

A relação entre sinal/ruído (SNR) é a razão, em decibéis, entre a potência do sinal útil e a potência do ruído que o acompanha. Em redes LTE (*Long Term Evolution*), o SNR funciona como métrica de primeira ordem para Qualidade de Serviço (QoS), pois se alinha estreitamente à potência de referência do sinal (RSRP) e à intensidade total recebida (LTERSSI) [3]. Assim, monitorar e prever SNR é crucial para decisões de *hand-over*, escolha de modulação, alocação de recursos de espectro e para manter conexões estáveis e de alto desempenho em sistemas de telecomunicações modernos.

A base de dados trabalhada é composta por medições de várias variáveis que formam a rede móvel “4G LTE” (SNR, RSRP, RSRQ, LTERSSI, DL, UL, velocidade do ônibus e altitude geográfica), cujos dados foram coletados com dois aparelhos celulares, Samsung Galaxy S9 (categoria LTE 18) e Samsung Galaxy S10e (categoria LTE 20) durante 30 viagens do ônibus público de Kingston Ontário, Canadá, em três horários distintos: 09h00, 12h00 e 18h00, utilizando mais de 190 mil pontos de dados e mais de 700 km percorridos.

METODOLOGIA:

Com a finalidade de investigar e compreender como as variáveis da base de dados se comportam e se relacionam entre si, foram utilizadas algumas métricas de estatística básica: Coeficiente de Correlação de Pearson, Coeficiente de Correlação de Kendall, Coeficiente de Correlação de Spearman e Teste de Hipóteses, a partir da linguagem de programação R. Desse modo, foram investigadas análises correlações lineares (Pearson) e outras que não exigem linearidade (Kendall e Spearman) com posterior confirmação de significância estatística por testes de hipóteses. Todos esses coeficientes de correlação possuem valores entre -1 e 1, em que a correlação é mais forte (negativamente ou positivamente) quando sua magnitude está próxima dos extremos -1 e 1 e, quanto mais próxima de zero, mais fraca a correlação.

O Coeficiente de correlação de Pearson (r) é descrito pela Equação 1, onde $cov(X,Y)$ é a covariância entre as variáveis X e Y e σ_X e σ_Y são os desvios padrões. Retorna uma medida estatística linear e adimensional para quantificar a força e a relação entre duas variáveis numéricas [4].

$$r = \frac{cov(X,Y)}{\sigma_X \sigma_Y}. \quad (\text{Equação 1})$$

O Coeficiente de Correlação de Kendall (τ) é descrito pela Equação 2, onde C é o número de pares concordantes e D é o número de pares discordantes. É uma medida estatística não paramétrica que avalia a associação entre duas variáveis ordinais, não assume uma relação linear entre as variáveis, e analisa a concordância ou discordância nas classificações dos dados [5].

$$\tau = \frac{C-D}{C+D}. \quad (\text{Equação 2})$$

O Coeficiente de Correlação de Spearman (ρ) é descrito pela Equação 3, onde d_i é a diferença entre os postos das duas variáveis para o i -ésimo indivíduo e n é a quantidade de pares de observações. É utilizado para descrever a relação entre duas variáveis monotonicamente relacionadas de forma não paramétrica, assim, baseia-se na distribuição dos postos (ou *ranks*) dos dados [6].

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)}. \quad (\text{Equação 3})$$

Um teste de hipóteses é um procedimento estatístico que permite tomar uma decisão utilizando os dados observados de uma amostra, e assim verifica uma suposição sobre o parâmetro populacional em análise [7]. Para a utilização desse procedimento, fixa-se o valor-p, que é a probabilidade de que a estatística do teste tenha valor discrepante do valor observado quando a hipótese nula é verdadeira. Definiu-se a hipótese nula como a correlação igual a zero e a hipótese alternativa como a correlação diferente de zero e definiu-se a correlação significativa com um valor-p menor do que ou igual a 0,05.

RESULTADOS E DISCUSSÃO:

A análise estatística dos dados, que agregou informações por horário, modelo de smartphone e em um conjunto de dados global, revelou correlações robustas e consistentes entre as variáveis da

rede móvel. Independentemente do método de correlação utilizado (Pearson, Kendall ou Spearman), a Relação Sinal-Ruído (SNR) demonstrou uma forte e significativa correlação positiva com a Potência Recebida do Sinal de Referência (RSRP) (valores entre 0,61 e 0,68) e com o Indicador de Intensidade de Sinal Recebido em LTE (LTERSSI) (valores entre 0,52 e 0,56). Todas as correlações relevantes foram validadas por meio de testes de hipóteses, confirmando sua significância estatística.

Os resultados apontam que RSRP e LTERSSI são as variáveis centrais que governam a qualidade do sinal, uma vez que o SNR acompanha expressivamente o comportamento de ambas. Além disso, identificou-se influências secundárias, mas estatisticamente significativas: a velocidade do ônibus apresentou uma correlação negativa de baixa magnitude com o SNR, sugerindo uma sutil deterioração da qualidade do sinal em movimento, enquanto a altitude demonstrou um discreto efeito positivo (valores entre 0,10 e 0,17). As taxas de transferência de dados (DL e UL) também se mostraram positivamente correlacionadas com as métricas de qualidade de sinal (valores entre 0,36 e 0,44), validando que melhores condições de recepção resultam em maior desempenho da rede.

As diferenças entre os coeficientes de correlação residem principalmente em suas escalas e pressupostos (linear e monotônico), mas os três métodos levaram a conclusões consistentes, reforçando a robustez das associações encontradas no sistema. Portanto, essa base de dados, com suas interações validadas, estabeleceu um fundamento sólido para a aplicação de uma modelagem computacional totalmente guiada por dados para a realização de previsão do SNR.

CONCLUSÕES:

A análise estatística estabeleceu, com sucesso, as relações fundamentais entre as variáveis da rede móvel, revelando a forte e significativa correlação positiva da Relação Sinal-Ruído (SNR) com a Potência Recebida do Sinal de Referência (RSRP) e a Intensidade Total do Sinal Recebido (LTERSSI). A consistência desses resultados, validada por múltiplos métodos de correlação e testes de hipóteses, forneceu uma base empírica sólida para a etapa de modelagem, destacando os principais indicadores da qualidade do sinal.

BIBLIOGRAFIA

- [1] SUGIHARA, George; PARK, Joseph; DEYLE, Ethan; SMITH, Cameron; YE, Hao. **Empirical Dynamic Modeling**. [S.l.]: [s.n.], 2022.
- [2] ELSHERBINY, Habiba; NAGIB, Ahmad M.; HASSANEIN, Hossam S. **4G LTE User Equipment Measurements along Kingston Transit 502 Bus Route**. [S.l.]: Borealis, 2020.
- [3] AFROZ, F.; SUBRAMANIAN, Ramprasad; HEIDARY, R.; SANDRASEGARAN, K.; AHMED, Safwat. **Sinr, Rsrp, Rssi and Rsrq Measurements in Long Term Evolution**. [S.l.], 2015.
- [4] KOIKE, Takaaki; LIN, Liyuan; WANG, Ruodu. Invariant correlation under marginal transforms. **Journal of Multivariate Analysis**, [S.l.], v. 204, 2024.

- [5] LIEBSCHER, E. Kendall regression coefficient. **Comput. Stat. Data Anal.**, [S.l.], v. 157, p. 107140, 2021.
- [6] DE WINTER, J. D.; GOSLING, S.; POTTER, J. Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. **Psychological methods**, [S.l.], v. 21, p. 273-290, 2016.
- [7] KING, M.; ZHANG, Xibin; AKRAM, M. Hypothesis testing based on a vector of statistics. **Journal of Econometrics**, [S.l.], v. 219, p. 425-455, 2020.
- [8] SUGIHARA, George; MAY, Robert M. Nonlinear Forecasting as a Way of Distinguishing Chaos from Measurement Error in Time Series. **Nature**, [S.l.], v. 344, p. 734-741, 1990.
- [9] SUGIHARA, George. Nonlinear Forecasting for the Classification of Natural Time Series. **Philosophical Transactions of The Royal Society of London. Series A: Physical and Engineering Sciences**, London, v. 348, p. 477-495, 1994.
- [10] CHANG, Chun-Wei; USHIO, Masayuki; HSIEH, Chihhao. Empirical Dynamic Modeling for Beginners. **Ecological Research**, [S.l.], v. 32, n. 6, p. 785-796, 2017.
- [11] WARNEKE, Konstantin; WAGNER, Carl-Maximilian; KEINER, Michael; HILLEBRECHT, Martin; SCHIEMANN, Stephan; BEHM, David George; WALLOT, Sebastian; WIRTH, Klaus. Maximal strength measurement: A critical evaluation of common methods-a narrative review. **Frontiers in Sports and Active Living**, [S.l.], v. 5, p. 1-11, 2023.