

Predição de Desempenho de Rede Resiliente a Falhas de Medição

Maria C. M. M. Ferreira¹, Silvio E. S. B. Ribeiro¹, Francisco V. J. Nobre¹,
Maria L. Linhares¹, Thelmo P. Araújo¹, Rafael L. Gomes¹

¹Universidade Estadual do Ceará (UECE), Fortaleza, Ceará, Brasil.

{clara.mesquita, silvio.eduardo, valderlan.nobre, malu.linhares}@aluno.uece.br

{thelmo.araujo, rafa.lopes}@uece.br

Resumo. *Serviços de monitoramento de rede são executados por diversas empresas e Provedores de Internet (ISP), que fornecem resultados de testes regulares de desempenho, tais como vazão, perda, atraso, dentre outros. Estas medições auxiliam a conhecer o comportamento da rede, bem como obter informações para um planejamento estratégico. Contudo, durante a realização das medições previstas durante o monitoramento da rede podem ocorrer falhas, as quais dificultam a execução de atividades mais complexas, tal como predição do desempenho de rede. Dentro deste contexto, este artigo apresenta um modelo resiliente e adaptativo para a predição de desempenho da rede, no qual inclui a identificação das falhas de medição, aplicando técnicas de imputação de dados a fim de adequar os dados para o processo de predição (baseado em Redes Neurais e Análise de Séries Temporais). Os experimentos realizados, usando dados reais da Rede Nacional de Ensino e Pesquisa (RNP), mostram que a solução proposta consegue atingir altos níveis de acurácia na predição com dados imputados, bem como supera outras abordagens de predição existentes.*

Abstract. *Network monitoring services are performed by several companies and Internet Providers (ISP), which provide results of regular performance tests, such as throughput, loss, and delay, among others. These measurements help to understand the behavior of the network, as well as obtain information for strategic planning. However, when carrying out the measurements planned during network monitoring, failures may occur, which makes it difficult to carry out more complex activities, such as forecasting network performance. Within this context, this article presents a resilient and adaptive model for forecasting network performance, which includes the identification of measurement failures and applying data imputation techniques to adapt the data for the forecasting process (based on Neural Networks and Time Series Analysis). The experiments carried out, using real data from the National Education and Research Network (RNP), show that the proposal can achieve high accuracy in prediction with imputed data, as well as outperforming other existing prediction approaches.*

1. Introdução

Provedores de Serviço de Internet (ISP) têm à sua disposição ferramentas de monitoramento de rede capazes de fornecer informações relevantes para a sua administração. Tais ferramentas executam testes regulares das mais variadas métricas de desempenho,

e a análise bem-sucedida de tais métricas se configura em conhecimento útil para um melhor planejamento estratégico. Corporações e instituições que operam sobre infraestruturas de rede devem ter um desempenho sólido na análise desses dados, uma vez que são fundamentais para a manutenção e evolução da rede. Tais informações abrangem aspectos primordiais da rede como o controle da Vazão de Bits, Perda de Pacotes, Atraso, dentre outras variáveis importantes [Vasileva et al. 2021, Yang et al. 2021]. A análise do tráfego da rede através de métricas de desempenho influencia, sobretudo, decisões administrativas, das quais se destacam: expansão da capacidade de enlace, expectativas de Qualidade de Serviço (QoS) e de Experiência (QoE) [Gomes et al. 2017], agendamento de manutenção, avaliações de Acordos de Nível de Serviço (*Service Level Agreement - SLA*) e análise de requisitos dos recursos da rede [Mok et al. 2021].

Em particular, o bom desempenho da Vazão é crítico para a QoS e disponibilidade de serviço de Internet, uma vez que uma baixa taxa de transferência pode resultar em atrasos, gargalos e congestionamentos na rede, levando a uma experiência de usuário insatisfatória e diminuição da eficiência da rede [Moreira et al. 2021]. Por outro lado, uma alta vazão de bits garante melhor transmissão de dados, melhorando a experiência do usuário e aumentando a produtividade do sistema [Portela et al. 2024, Gomes et al. 2016]. Além disso, a demanda elástica por recursos de rede também influencia no desempenho da rede, uma vez que a sua variação ao longo do dia ocorre devido à mobilidade humana nas cidades e diferentes padrões de comportamento na mesma rede local, dentre outros aspectos [Gomes et al. 2020]. Portanto, garantir o bom desempenho da rede é crucial para fornecer serviços de alta qualidade aos usuários, exigindo monitoramento constante e otimização das principais métricas de rede, cumprindo o SLA e evitando problemas de conexão e atrasos elevados [Portela et al. 2023].

Apesar da importância evidente de um alto desempenho da rede, é crucial reconhecer que as ferramentas de monitoramento empregadas podem trazer problemas na realização dos testes - por diversos aspectos e limitações - ocasionando em valores de medição perdidos, além de uma consequente baixa qualidade dos dados para análise [Costa et al. 2021]. Este fato pode comprometer técnicas de mineração de dados que analisam o comportamento temporal das informações, bem como prejudicar a boa compreensão e a atuação de ações de gerenciamento da rede, além de estratégias de planejamento e prospecção [Silveira et al. 2023]. Dessa forma, técnicas de pré-processamento e imputação de dados sintéticos ajudam na melhor compreensão do desempenho da rede, contornando eventuais limitações que venham a ocorrer durante as medições, e permitindo que problemas sejam antecipados proativamente [Naf et al. 2022]. Além disso, previnem falhas e alterações iminentes, uma vez que envolve a utilização de dados reais e técnicas de modelagem de imputação para estimar valores e fornecer uma melhor análise de desempenho da rede sob condições variadas [da Silva et al. 2020].

Dentro deste contexto, este artigo apresenta um modelo adaptativo de predição de desempenho de rede resiliente. A proposta é considerada resiliente, uma vez que aplica técnicas selecionadas para imputação de dados a fim de adequar as informações para o processo de predição quando ocorrem falhas de medição (e, conseqüentemente, lacunas da série temporal a ser analisada). Adicionalmente, o modelo de predição proposto é considerado adaptativo, visto que ajusta a série temporal das medições realizadas através de análises estatísticas (tais como a decomposição, definição de tendência e remoção

de erros de ciclo), criando uma série temporal padronizada que possibilita um melhor treinamento de modelos de predição. A série temporal tratada é usada como insumo para os modelos de Inteligência Artificial (IA) que irão realizar o processo de predição de uma nova série com o desempenho de rede esperado, mesmo quando ocorrem casos de falhas de medição. Em nosso trabalho anterior [Ribeiro et al. 2023], o processo de predição de desempenho era realizado sem lidar com as lacunas de dados geradas pelas falhas de medição, prejudicando o desempenho dos modelos de IA. Assim, o presente trabalho evolui a proposta anterior, trazendo novos aspectos de resiliência e mitigação de falhas.

A fim de validar a solução e analisar seu desempenho no que se refere a capacidade de predição, foram realizados experimentos usando dados reais do Serviço de Monitoramento da Rede Ipê (Monipê)¹ da Rede Nacional de Ensino e Pesquisa (RNP). Os resultados obtidos mostram que o modelo proposto consegue atingir altos níveis de acurácia na predição, bem como supera o uso de modelos de predição existentes no que se refere a métricas de avaliação, dentro do contexto de gerenciamento das métricas de rede. Adicionalmente, as técnicas de imputação de dados sintéticos melhoram o processo de predição frente aos dados originais, de acordo com a avaliação do RMSE.

O restante deste artigo está organizado da seguinte forma: A Seção 2 apresenta os trabalhos relacionados existentes. A Seção 3 apresenta o modelo de predição de desempenho proposto, enquanto a Seção 4 descreve os resultados dos experimentos realizados. Por fim, a Seção 6 conclui o artigo e apresenta trabalhos futuros.

2. Trabalhos Relacionados

Esta seção descreve os principais trabalhos relacionados, e recentemente publicados pela comunidade científica, sobre predição e sobre imputação no contexto de redes de computadores, incluindo aspectos de desempenho e qualidade de serviço.

Liang et al. [Liang et al. 2023] apresentam um modelo baseado em deep learning que consegue estimar os valores de média e de variância da distribuição de probabilidades das previsões a partir da extração de informação do interior de uma estrutura de grafo. Além disso, propôs um framework de probabilidade de predição para estimar a distribuição do valor de QoS. Essa estimativa pode ser posteriormente adaptada para muitas outras tarefas subsequentes, como processos de tomadas de decisão e detecção de anomalias. De forma similar, Damaskinos et al. [Damaskinos et al. 2022] propôs um intermediário entre o sistema operacional Android e uma aplicação de aprendizagem de máquina que prediz e controla o impacto da aprendizagem de tarefas em dispositivos móveis, i.e., prediz e controla o tempo computacional e o consumo de energia de cada tarefa de aprendizado. No entanto, nenhuma das abordagens busca entender o comportamento da rede, além de não fazerem uma predição de demanda por recursos.

Yang et al. [Yang et al. 2021] apresentam um método de predição de tráfego de rede baseado no ARIMA (*Autoregressive Integrated Moving Average model*)-BPNN (*Back Propagation Neural Network*) a fim de aumentar a precisão para predição de tráfego futuro. O ARIMA-BPNN realiza uma avaliação do potencial de mineração de comportamentos lineares e não lineares dos dados históricos de tráfego de rede, melhorando assim a precisão da predição. No entanto, o método proposto não considera a predição do desempenho da rede, somente do tráfego em si.

¹monipe-central.rnp.br

França et al. [França et al. 2021] apresentam dois modelos de regressão baseados em redes neurais para imputação de dados faltantes em gateways IoT, considerando o tempo de execução e a quantidade de memória utilizada. Da mesma forma, Park et al. [Park et al. 2023] apresentam uma abordagem que usa um modelo de aprendizagem profunda para estimar os valores ausentes de uma variável em dados multivariados de séries temporais, com foco no preenchimento de uma lacuna longa e contínua (por exemplo, vários meses de observações diárias ausentes) em vez de observações individuais aleatórias faltantes. Similarmente, Ding et al. [Ding et al. 2020] avaliam uma série de métodos de imputação para preenchimento de falhas de séries temporais recolhidas de dispositivos IoT. Neste estudo, as técnicas de interpolação de dados *Radial Basis Functions*, *Moving Least Squares (MLS)* e *Adaptive Inverse Distance Weighted* que foram usadas para o preenchimento das lacunas foram comparadas entre si e com o KNN, e constatou-se que, nesse caso, a técnica de interpolação de MLS de Lancaster foi a que obteve o melhor resultado. Nesse contexto, o estudo não avalia a interferência desses métodos de imputação na predição das séries temporais.

A partir do levantamento bibliográfico realizado, nota-se que nenhum artigo da literatura se concentrou no desenvolvimento de um modelo de predição de desempenho de rede que seja adaptável ao contexto da medição a ser realizada, que é o foco deste artigo, além de considerar variadas técnicas de imputação de dados no processo. Esta proposta realiza ações estatísticas para aprimorar o processo de análise e predição, criando um modelo padronizado de amostras que melhora a eficiência e execução de modelos de Aprendizado de Máquina e proporcionam melhores resultados e contribui para um melhor planejamento estratégico da infraestrutura de rede.

3. Proposta

Esta seção apresenta a descrição detalhada do modelo adaptativo de predição resiliente proposto neste artigo. Esta proposta visa ajustar os dados de medição da vazão fim-a-fim entre dois pontos na rede, através de técnicas de análise em séries temporais, para situações de demanda elástica, bem como a identificação e correção de falhas de medição, possibilitando assim uma maior eficiência no processo de predição de desempenho. A solução é, de forma geral, executada em quatro etapas: (i) Coleta de dados necessários para a predição, (ii) Correção de falhas de medição através de técnicas de imputação de dados, (iii) Tratamento da série temporal (aplicando técnicas de decomposição e interpolação) e (iv) Aplicação das técnicas de predição. Uma visão geral da estruturação da solução proposta neste projeto é apresentada na Figura 1.

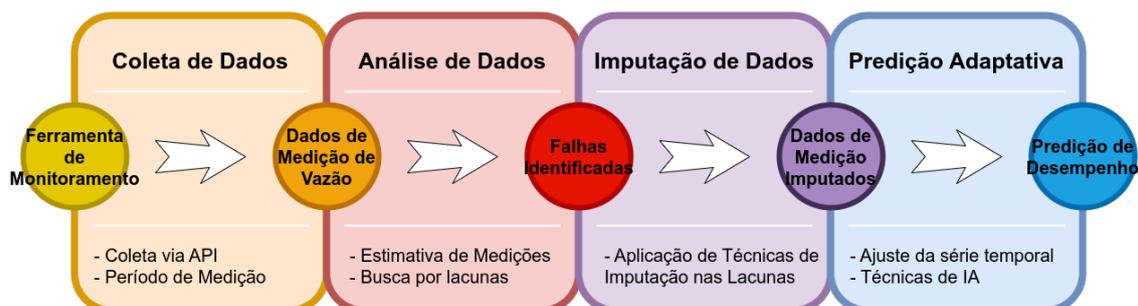


Figura 1. Visão Geral da Solução.

Inicialmente, a solução resgata os dados de medição oriundos dos serviços existentes (elemento em amarelo da Figura 1), onde qualquer ferramenta pode ser considerada, tais como Perfsonar, SolarWinds, Auvik, etc. Posteriormente, em posse dos dados, a solução verifica a existência de lacunas nas medições previstas para determinado período analisado, aplicando técnicas de imputação a fim de corrigir tais falhas. Em seguida, aplica-se um modelo proposto (com base nas medições de rede) na etapa de tratamento da série - elementos em azul da Figura 1 - com o objetivo de garantir um ajuste nos dados, ou seja, adequar a análise temporal ao contexto de demanda elástica e particularidades de cada contexto de comunicação. Por fim, o resultado do tratamento e análise da série temporal é usado para treinar um modelo de predição, isto é, aplica-se alguma técnica de predição para gerar uma nova série temporal de predição em relação ao desempenho analisado. A seguir, serão detalhadas as ações realizadas em cada uma das etapas apresentadas.

3.1. Técnicas de Imputação de Dados

Em geral, existem duas abordagens para técnicas de correção de dados ausentes [Thakur et al. 2021, Wu et al. 2015]: Interpolação e IA. As técnicas de interpolação são baseadas em um método matemático que ajusta uma função aos dados e usa essa função para gerar dados ausentes na série analisada. A interpolação linear é a mais simples, que calcula a média dos valores anteriores e subsequentes dos dados ausentes. Além disso, existem vários outros métodos de interpolação que podem ser aplicados, como Ponderada pelo Tempo, Média Móvel e Mediana Móvel, entre outros presentes na literatura.

Em relação ao uso de IA, essas podem ser aplicadas para analisar o conjunto de dados como um todo, definir as lacunas como alvos e propor um modelo para tratar esses espaços. Cada técnica de IA possui uma abordagem distinta, sendo aplicável no contexto deste projeto *K-Vizinhos Mais Próximos* (KNN), Modelos de Aprendizado Profundo, Redes Neurais, etc. Em seguida, descrevemos as técnicas aplicadas neste trabalho:

- **Média Móvel:** A imputação por média móvel é uma técnica estatística usada para lidar com dados ausentes em séries temporais ou sequências. Este método envolve a substituição de valores ausentes pela média de observações próximas dentro de uma janela de tempo específica. Ao calcular a média móvel, os pontos de dados ausentes são estimados, fornecendo uma representação suavizada dos dados originais. Esta abordagem é particularmente útil ao ajudar a manter a tendência geral e os padrões no conjunto de dados. A escolha da janela de tempo ou da média móvel (simples, ponderada ou exponencial) pode ser adaptada às características dos dados e aos objetivos.
- **Mediana Móvel:** A imputação por mediana móvel é um método empregado no contexto de imputação de dados ausentes para séries temporais ou sequências. Em vez de usar a média, esta técnica substitui os valores ausentes pela mediana dos pontos de dados dentro de uma janela de tempo especificada. A mediana móvel é uma alternativa robusta à média móvel, pois é menos afetada por valores extremos ou *outliers* no conjunto de dados. A Mediana Móvel se torna uma boa escolha ao lidar com sequências que podem conter flutuações irregulares ou esporádicas, onde manter a tendência central dos dados sem ser influenciado por valores extremos é importante. A seleção da janela de tempo é crucial nesta abordagem para controlar o nível de suavização e adaptar o método às características específicas dos dados.

- **Interpolação Linear:** A interpolação linear é um método direto e amplamente utilizado para lidar com dados ausentes em vários tipos de conjuntos de dados, incluindo séries temporais e sequências unidimensionais. Este método envolve a estimativa de valores ausentes traçando retas entre pontos de dados adjacentes conhecidos. A interpolação linear assume que a relação entre os dados é linear, o que significa que os valores mudam uniformemente entre dois pontos observados. Os pontos de dados ausentes são então calculados com base nessa relação linear. A interpolação linear é eficaz para conjuntos de dados em que os valores exibem um padrão ou tendência linear consistente.
- **K-Vizinhos Mais Próximos (KNN):** A imputação de K-Vizinhos Mais Próximos (*K-Nearest Neighbors* - KNN) é um método mais sofisticado de imputação que aproveita a similaridade entre pontos de dados para estimar valores ausentes. Neste método, determinamos os valores para os pontos de dados ausentes calculando médias ponderadas com base nos valores dos k vizinhos mais próximos [Ding et al. 2020]. A escolha de K , o número de vizinhos, é um parâmetro crítico, pois determina a influência dos pontos de dados próximos no valor imputado. A imputação de KNN é especialmente útil quando o conjunto de dados apresenta padrões e relações complexas, pois considera a vizinhança local de cada ponto de dados ausente. Pode ser aplicada tanto a dados numéricos quanto categóricos, sendo um método de imputação versátil que pode capturar dependências intrincadas nos dados. No entanto, a escolha de K e a métrica de distância utilizada (por exemplo, distância euclidiana ou de Manhattan) devem ser cuidadosamente selecionadas com base nas características do conjunto de dados e nos objetivos da análise.

3.2. Análise e Tratamento da Série Temporal

Em geral, os modelos tradicionais de previsão não tratam adequadamente falhas em conjuntos de dados, o que pode prejudicar a eficácia dos resultados. Esta abordagem dificulta a organização de séries temporais regulares e, conseqüentemente, análises precisas, afetando diretamente a qualidade das previsões, como evidenciado neste estudo. Por outro lado, modelos mais robustos que incluem uma etapa de pré-processamento dos dados, aumentam a tolerância e a resiliência a lacunas nas medições. Neste trabalho, aliamos isso a modelos de Redes Neurais que são caracterizados por uma capacidade computacional avançada e tolerância elevada a erros de medição [Lara-Benítez et al. 2021]. Uma série temporal é uma sucessão de observações registradas em ordem cronológica, durante intervalos de tempo regulares. O desafio de previsão consiste em ajustar um modelo para antecipar os valores futuros da série, levando em conta as observações passadas, i.e., o seu histórico. [Lara-Benítez et al. 2021]. Para isto, existem técnicas estatísticas, modelos Autorregressivos ou de Suavização, bem como técnicas mais robustas que utilizam Redes Neurais e Inteligência Artificial. Sendo assim, grandes falhas em conjuntos de dados de séries temporais se tornam um problema na predição, que se baseia nos dados que já existem para produzir novas observações. Além disso, a forma como esses dados são previamente tratados interfere diretamente nos resultados de predição [Tawn et al. 2020, Saad et al. 2020].

Reconhecendo a importância de minimizar os impactos das lacunas durante a análise e previsão, foi realizado um pré-processamento abrangente dos dados de treinamento. Esta abordagem incluiu a aplicação de técnicas de imputação para corrigir e tratar as falhas observadas previamente. A etapa de imputação é fundamental para melhorar o desempenho do preditor no treinamento que virá a seguir. Em ambientes onde lacunas

substanciais podem comprometer a confiabilidade das análises, técnicas de imputação emergem como um componente crítico do processo.

3.3. Modelos de Predição: LSTM e GRU

O modelo LSTM é um sistema neural recorrente especialmente projetado para superar os problemas de explosão/desaparecimento do gradiente, que normalmente surgem ao aprender dependências de longo prazo, mesmo quando os intervalos de tempo mínimos são muito longos [Rafi et al. 2021]. No geral, isso pode ser evitado usando um *Constant Error Carousels* (CEC), que mantém o sinal de erro dentro da célula de cada unidade. Na verdade, tais células são elas próprias redes recorrentes com uma arquitetura melhorada para estender o CEC com recursos adicionais, ou seja, a porta de entrada e saída formam a célula de memória. As conexões auto-recorrentes indicam feedback com atraso de um intervalo de tempo. O LSTM é adequado para classificar, processar e prever séries temporais com intervalos de tempo de duração desconhecida. A insensibilidade relativa do comprimento da lacuna dá ao LSTM uma vantagem sobre os modelos de RNNs tradicionais, modelos ocultos de Markov (MOM) e outros métodos de aprendizado de sequência.

Similarmente, a GRU é a nova geração de redes neurais recorrentes e é bastante semelhante a uma LSTM. As GRUs se livraram do estado da célula e usaram o estado oculto para transferir informações. Essa arquitetura possui apenas duas portas, uma de redefinição (*reset gate*) e uma de atualização (*update gate*), que são utilizadas para resolver o problema da dissipação do gradiente de uma RNN padrão. Basicamente, elas são dois vetores que decidem quais informações devem ser passadas para a saída. O que há de especial é que elas podem ser treinadas para manter informações de muito tempo atrás, sem dissipá-las com o tempo ou remover informações irrelevantes para a predição. Enquanto as LSTMs possuem dois estados diferentes passados entre as células, o estado da célula e o estado oculto, que carregam a memória de longo e curto prazo, respectivamente, as GRUs têm apenas um estado oculto transferido entre as etapas do tempo. Esse estado oculto é capaz de manter as dependências de longo e curto prazo ao mesmo tempo, devido aos mecanismos de restrição e cálculos pelos quais o estado oculto e os dados de entrada passam.

4. Experimentos

Esta seção irá descrever a forma como os experimentos realizados foram configurados (Subseção 4.1) e o método de avaliação dos resultados obtidos (Subseção 4.2). É válido ressaltar que o código desenvolvido, bem como os dados utilizados nos experimentos, estão disponíveis no repositório do projeto² e com as instruções necessárias para reprodutibilidade.

4.1. Configuração de Experimentos

A fim de realizar experimentos usando dados do mundo real, foram utilizados dados da RNP, através do Serviço de Monitoramento da Rede Ipê (MonIPÊ). O MonIPÊ usa o padrão de monitoramento internacional perfSONAR, onde as medições de Vazão ocorrem a cada 4 horas, enquanto as medições de Perda e Atraso ocorrem a cada 5 minutos. A partir deste cenário, foi decidido focar nos dados de vazão de rede, devido a maior

²<https://github.com/LarcesUece/Resilient-Performance-Forecasting>

representatividade em relação a capacidade de transmissão das aplicações sobre a rede, e pela oportunidade de avaliar o comportamento de diferentes tipos de controle de congestionamento para TCP (*Cubic* e BBR) frente ao processo de predição.

Faz-se necessário compreender a diferença entre os TCPs BBR e *Cubic*, visto que ambos possuem comportamentos diferentes [Miyazawa et al. 2020, Kanaya et al. 2020]. O TCP-BBR ao invés de responder diretamente a perda de pacotes, ele também leva em consideração largura de banda, RTT e outras métricas, calculando o “*padding gain*” para definir os valores entre envio de pacotes e a janela de congestionamento. Esta abordagem do BBR traz maior vazão e mais equilíbrio ao algoritmo. Por outro lado, o TCP *Cubic* utiliza uma função cúbica no tempo para o aumento da janela de congestionamento em tempo real, frente a perda de pacotes, fazendo assim com que o *Cubic* responda com mais qualidade à perda de pacotes, sem limitar consideravelmente o crescimento da janela quando não acontece a perda.

Nos experimentos, foram examinados dois pares de comunicação: PA-BA e PR-AM, abrangendo Pontos de Presença (PoP) da rede em quatro estados brasileiros: *Pará* (PA), *Amazonas* (AM), *Bahia* (BA) e *Paraná* (PR). Tais PoPs da Rede Ipê foram selecionados devido à diversidade de localizações geográficas, impactando na utilização dos links e na carga da rede, variando as capacidades da infraestrutura (enlaces variam de 200 Gbps a 1 Gbps), o que leva a comportamentos de comunicação distintos.

É válido ressaltar que a comunicação fim-a-fim entre os diversos pontos da rede possuem comportamentos distintos ao longo do dia e durante a semana, visto que a utilização da infraestrutura segue um comportamento social dos seus usuários. Portanto, foram considerados os quatro pontos de comunicação supracitados, que abrangem uma parte relevante do tráfego de bits na rede, interconectando as regiões Norte, Nordeste e o Sul do país. Mais especificamente, estes pontos foram escolhidos devido a: (i) Heterogeneidade em relação a localização geográfica, a qual influencia no número de enlaces utilizados no caminho fim-a-fim e na carga total gerada sobre a infraestrutura de rede; (ii) Capacidade da infraestrutura de rede, uma vez que os enlaces possuem uma alta variação de carga, fazendo com que a comunicação originada de um determinado PoP tenha um comportamento muito diferente em relação a outros dois que possuam uma variação de capacidade fim-a-fim distintas.

Antes do estudo do impacto dos métodos de imputação na predição, foi realizada uma avaliação preliminar das técnicas de imputação utilizadas. Para isto, foram selecionados conjuntos de dados dos mesmos dois pares de comunicação considerados, uma vez que, além das razões já especificadas, estes oferecem um maior contraste entre os resultados gerados por diferentes técnicas de imputação. Além disso, é importante notar que as porcentagens de quantidades de dados faltantes nos conjuntos de dados foi considerada durante o processo de seleção.

Tais informações foram derivadas de uma análise inicial, que quantificou a extensão dos dados faltantes em cada conjunto de dados. Especificamente, no período observado de aproximadamente seis meses, o conjunto de dados de PR-AM obteve 29.37% de dados faltantes e PA-BA mostrou-se com 28.86% de lacunas. Sendo assim, o procedimento adotado foi buscar pela maior sequência observada sem falhas de medição, e fazer a extração de valores aleatórios, representando algo em torno de 30% dos dados obser-

vados como falhas de medição. Após isto, foram aplicadas as técnicas de imputação em cada série e, na sequência, foi calculado o RMSE dos valores gerados frente aos valores originais. Dessa forma buscou-se avaliar a qualidade na geração de novas observações pelas técnicas de preenchimento frente aos valores originais existentes.

Com relação ao treinamento do Modelo de Predição, foram utilizadas séries temporais representando a vazão fim-a-fim entre os mesmos dois pontos de comunicação. Após o processo de análise e aplicação das técnicas de imputação para cada série gerada, foram separados 80% da série para treinamento e 20% para testes e validação dos modelos, i.e., por volta de 560 e 140 ciclos, respectivamente. As séries foram submetidas a quatro rodadas de treinamento diferentes variando os parâmetros dos modelos LSTM e GRU com intuito avaliar o desempenho de cada preditor.

Inicialmente, foram aplicadas as quatro técnicas de imputação descritas na Seção 3.1, para cada uma das séries temporais: Vazão PA-BA Cubic, Vazão PA-BA BBR, Vazão PR-AM Cubic e Vazão PR-AM BBR. Totalizando 16 séries temporais resultantes com valores preenchidos. Após isso, todas as séries resultantes foram submetidas ao processo de predição adaptativo presente em [Ribeiro et al. 2023] com foco para os modelos LSTM e GRU que obtiveram melhor desempenho na predição. Finalmente, foi calculado o RMSE para cada um dos resultados de predição frente aos valores reais.

4.2. Métodos de Avaliação

Para avaliar com melhor precisão as técnicas de imputação e o desempenho dos modelos de predição frente a valores reais, foi utilizado o *Root Mean Square Error* (RMSE). Com isso, é possível validar os valores gerados pelas técnicas, além de avaliar a predição pelos modelos utilizando as mesmas técnicas de preenchimento. O RMSE para um determinado período de tempo T é definido pela Equação 1, onde \hat{y}_t representa o valor predito e y_t significa o valor real do desempenho no tempo t . No contexto da métrica RMSE, valores maiores indicam desempenho de imputação menos precisos da abordagem.

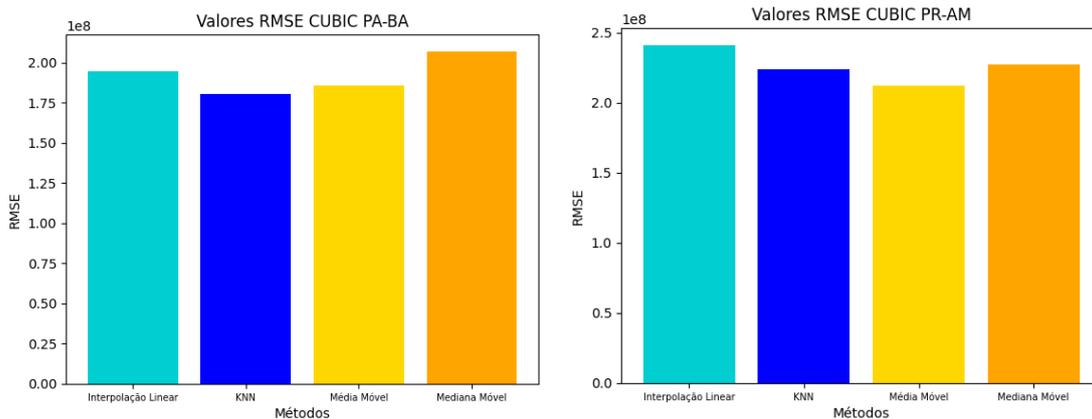
$$\text{RMSE}(T) = \frac{1}{\sqrt{T}} \left(\sum_{t=1}^T (\hat{y}_t - y_t)^2 \right)^{\frac{1}{2}} . \quad (1)$$

5. Resultados

Esta seção apresenta os resultados obtidos a partir da análise dos experimentos realizados com um conjunto de dados reais, onde as Subseções 5.1 e 5.2 discutem os principais pontos no que se refere as ações de imputação de dados e predição de desempenho, respectivamente.

5.1. Avaliação das Técnicas de Imputação

Com relação a imputação de dados, a Figura 2 ilustra a comparação de cada técnica com dados originais através do RMSE. Como exemplo, foi utilizado o tráfego Cubic, devido a sua alta variabilidade e maior dificuldade de predição de valores exatos. Foi retirado aleatoriamente 30% dos valores (representando, de forma geral, a ordem de dados faltantes nas medições coletadas) em um intervalo regular e estreito de 18 e 35 valores para PR-AM e PA-BA, respectivamente. É válido ressaltar que imperfeições nos dados são



(a) RMSE do PA-BA.

(b) RMSE do PR-AM.

Figura 2. Resultados do RMSE para as Imputações de Dados de Vazão

uma ocorrência comum que resultam em irregularidades e, conseqüentemente, afetam especialmente os aspectos temporais de análise.

Observa-se que a distribuição densa de dados nos casos de PA-BA e PR-AM (para o tráfego Cubic) reduz a necessidade de imputação dentro de intervalos estreitos e, como resultado, os cálculos de precisão com base em intervalos específicos tornam-se menos informativos. Conseqüentemente, a métrica RMSE tende a ser mais alta em tais situações devido ao espaço limitado para melhorias substanciais por meio de técnicas de imputação.

Sendo assim, do ponto de vista de eficiência de imputação de dados a Média Móvel e o KNN se destacam com menores valores de RMSE. Contudo, é válido ressaltar que os métodos resultam em erros de imputação mesmo que reduzidos. Assim, faz-se necessário ainda entender o impacto dessas imputações no processo de predição.

5.2. Avaliação do Processo de Predição com Dados Imputados

Esta seção apresenta os resultados obtidos a partir da análise dos experimentos realizados com os conjunto de dados reais provenientes das técnicas de imputação descritas anteriormente. Os resultados do RMSE são apresentados nas Figuras 3 e 4.

Analisando a Figura 3 é possível notar que, de forma geral, para os pontos de comunicação PA-BA, os preditores atingiram níveis satisfatórios de erro, uma vez que os valores de RMSE ficaram entre 50 e 250 Mbits/s. O TCP Cubic, devido ao comportamento "agressivo" da sua janela de congestionamento, fato que dificulta o ajuste dos modelos para a predição, o erro aumenta em todos os casos, uma vez que tal comportamento causa muita variação nas medições coletadas.

Importante ressaltar que para o TCP BBR, os modelos de predição conseguiram atingir resultados bastante satisfatórios de RMSE, superando em todos os casos o TCP Cubic, uma vez que ambos os preditores atingiram os menores valores de erro, com destaque para o modelo GRU que teve uma leve vantagem sobre o LSTM, em todos os casos.

É possível observar tanto na Figura 3 quanto na Figura 4 que o que mais interfere no resultado da predição é exatamente a técnica de imputação aplicada aos dados. A variação de algumas técnicas produz resultados semelhantes, principalmente a

Interpolação Linear e KNN, que para ambos os casos se mostraram eficientes.

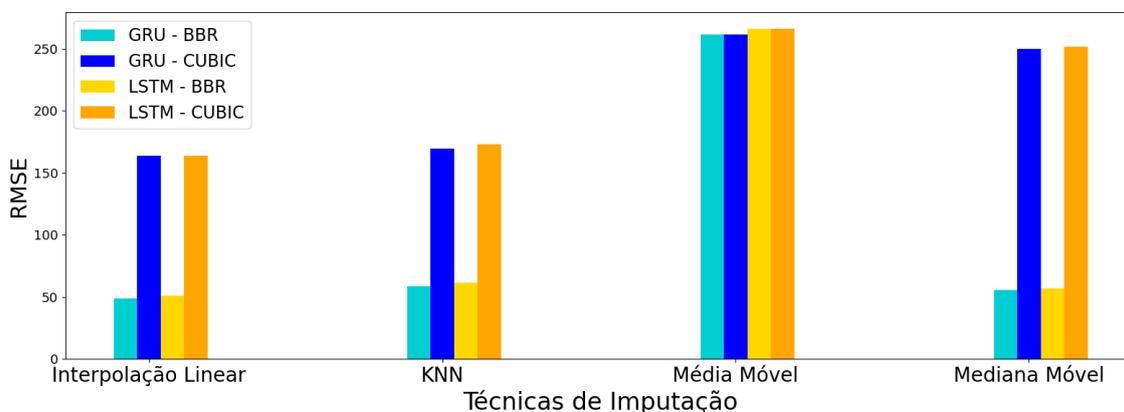


Figura 3. Resultados do RMSE para as Predições de Vazão entre os PoPs PA-BA.

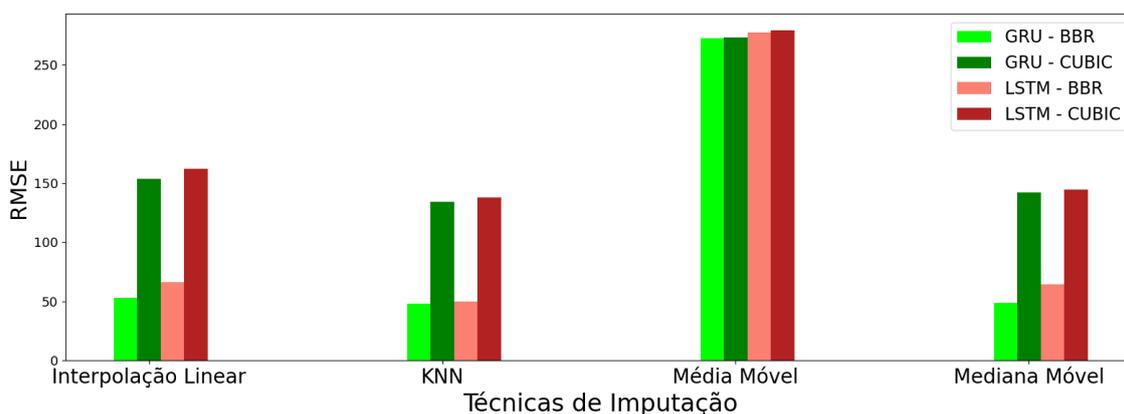


Figura 4. Resultados do RMSE para as Predições de Vazão entre os PoPs PR-AM.

É possível observar que há uma semelhança no desempenho do modelo utilizando Interpolação Linear, KNN e Mediana Móvel para os pontos de comunicação PR-AM. Embora KNN seja uma técnica de Inteligência Artificial reconhecida, possui desempenho semelhante quando comparado com a Interpolação Linear, por exemplo, esta que é uma técnica menos sofisticada.

Em ambos os casos, a Média Móvel mostra o pior desempenho quando considerado o processo de predição, em todos os cenários. É uma observação relevante, uma vez que foi observado na Seção 5.1 que, de uma forma geral, a mesma técnica foi a que mais se aproximou dos valores reais de medição. Isso reforça que um intervalo estreito de ocorrências não consegue representar bem a eficácia das técnicas.

A Figura 5 ilustra o ajuste do modelo aos dados originais para o tráfego PA-BA, tomando como exemplo um dos casos de predição. É possível notar que os preditores LSTM e GRU conseguem bons resultados frente aos valores selecionados para teste, reconhecendo com precisão flutuações e dependências de longo prazo. A robustez do uso das Redes Neurais para a predição mostra que, embora esta ocorra sobre dados imputados e séries temporais adaptadas, ainda assim os preditores reconhecem com alta precisão os

padrões e conseguem extrapolar valores futuros com qualidade, indicando uma adequada capacidade de generalização para este contexto.

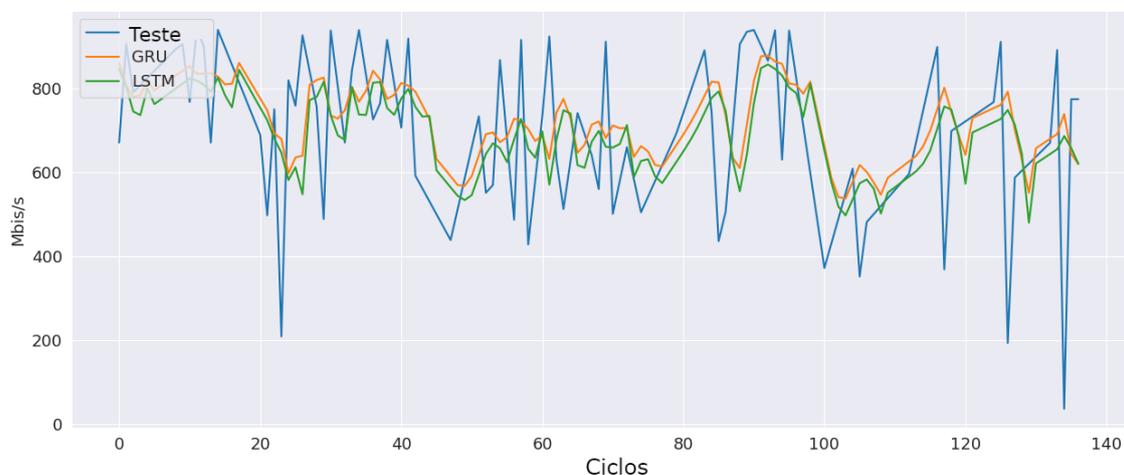


Figura 5. Predição realizada para os pontos de comunicação PA-BA.

6. Conclusão

A avaliação do desempenho de redes é uma ação voltada para adquirir dados relevantes para o planejamento estratégico de empresas e provedores de Internet. A capacidade de prever esse desempenho torna-se uma funcionalidade essencial para assegurar a eficácia dos serviços que operam na rede, contribuindo diretamente para a experiência dos usuários nesse contexto. Contudo, o processo de predição de desempenho é complexo, ainda mais diante da realidade atual de ocorrência de falhas de medição, ocasionando lacunas nos dados e, conseqüentemente, dificultando a realização de atividades mais complexas.

A fim de lidar com esta situação, este artigo apresentou um modelo adaptativo e resiliente de previsão de desempenho de rede. Esse modelo identifica falhas nas medições e utiliza técnicas de imputação de dados para preparar os dados para o processo de predição, com base em Redes Neurais e Análise de Séries Temporais. Os experimentos conduzidos, utilizando dados reais da Rede Nacional de Ensino e Pesquisa (RNP), demonstram que a solução proposta alcança elevados níveis de precisão na previsão com dados imputados.

Como trabalhos futuros, pretende-se evoluir a solução a fim de considerar séries temporais multi-variadas, incluindo outros tipos de medição de rede (tais como atraso e perda), bem como a verificação do comportamento da solução considerando modelos de aprendizado de máquina combinados.

Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) do Brasil (*N*º 303877/2021-9) e Rede Nacional de Pesquisa (RNP) através do Programa de Monitoramento de Redes (PMON 2023) e do Comitê Técnico de Ciência de Dados e Inteligência Artificial (CT-CDIA 2023) pelo apoio financeiro.

Referências

- Costa, W. L., Portela, A. L., and Gomes, R. L. (2021). Features-aware ddos detection in heterogeneous smart environments based on fog and cloud computing. *International Journal of Communication Networks and Information Security*, 13(3):491–498.
- da Silva, G., Oliveira, D., Gomes, R. L., Bittencourt, L. F., and Madeira, E. R. M. (2020). Reliable network slices based on elastic network resource demand. In *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*, pages 1–9.
- Damaskinos, G., Guerraoui, R., Kermarrec, A.-M., Nitu, V., Patra, R., and Taiani, F. (2022). Fleet: Online federated learning via staleness awareness and performance prediction. *ACM Trans. Intell. Syst. Technol.*, 13(5).
- Ding, Z., Mei, G., Cuomo, S., Li, Y., and Xu, N. (2020). Comparison of estimating missing values in iot time series data using different interpolation algorithms. *International Journal of Parallel Programming*, 48(3):534–548.
- França, C. M., Couto, R. S., and Velloso, P. B. (2021). Missing data imputation in internet of things gateways. *Information*, 12(10).
- Gomes, R., Bittencourt, L., Madeira, E., Cerqueira, E., and Gerla, M. (2017). Management of virtual network resources for multimedia applications. *Multimedia Systems*, 23(4):405–419.
- Gomes, R. L., Bittencourt, L. F., Madeira, E. R., Cerqueira, E., and Gerla, M. (2016). A combined energy-bandwidth approach to allocate resilient virtual software defined networks. *Journal of Network and Computer Applications*, 69:98–106.
- Gomes, R. L., Bittencourt, L. F., and Madeira, E. R. M. (2020). Reliability-aware network slicing in elastic demand scenarios. *IEEE Communications Magazine*, 58(10):29–34.
- Kanaya, T., Tabata, N., and Yamaguchi, S. (2020). A study on performance of cubic tcp and tcp bbr in 5g environment. In *2020 IEEE 3rd 5G World Forum (5GWF)*, pages 508–513.
- Lara-Benítez, P., Carranza-García, M., and Riquelme, J. C. (2021). An experimental review on deep learning architectures for time series forecasting. *International Journal of Neural Systems*, 31(03):2130001. PMID: 33588711.
- Liang, W., Li, Y., Xu, J., Qin, Z., Zhang, D., and Li, K.-C. (2023). Qos prediction and adversarial attack protection for distributed services under dlaas. *IEEE Transactions on Computers*, pages 1–1.
- Miyazawa, K., Yamaguchi, S., and Kobayashi, A. (2020). Mechanism of cyclic performance fluctuation of tcp bbr and cubic tcp communications. In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pages 1139–1144.
- Mok, R. K. P., Zou, H., Yang, R., Koch, T., Katz-Bassett, E., and Claffy, K. C. (2021). Measuring the network performance of google cloud platform. In *Proceedings of the 21st ACM Internet Measurement Conference, IMC '21*, page 54–61, New York, NY, USA. Association for Computing Machinery.
- Moreira, D. A., Marques, H. P., Costa, W. L., Celestino, J., Gomes, R. L., and Nogueira, M. (2021). Anomaly detection in smart environments using ai over fog and cloud

- computing. In *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*, pages 1–2. IEEE.
- Naf, J., Spohn, M.-L., Michel, L., and Meinshausen, N. (2022). Imputation scores.
- Park, J., Müller, J., Arora, B., Faybishenko, B., Pastorello, G., Varadharajan, C., Sahu, R., and Agarwal, D. (2023). Long-term missing value imputation for time series data using deep neural networks. *Neural Computing and Applications*, 35(12):9071–9091.
- Portela, A. L., Menezes, R. A., Costa, W. L., Silveira, M. M., Bittecourt, L. F., and Gomes, R. L. (2023). Detection of iot devices and network anomalies based on anonymized network traffic. In *NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium*, pages 1–6.
- Portela, A. L. C., Ribeiro, S. E. S. B., Menezes, R. A., de Araujo, T., and Gomes, R. L. (2024). T-for: An adaptable forecasting model for throughput performance. *IEEE Transactions on Network and Service Management*, pages 1–1.
- Rafi, S. H., Nahid-Al-Masood, Deeba, S. R., and Hossain, E. (2021). A short-term load forecasting method using integrated cnn and lstm network. *IEEE Access*, 9:32436–32448.
- Ribeiro, S., Menezes, R., Portela, A., Araújo, T., and Gomes, R. (2023). Aplicando redes neurais e análise temporal para predição adaptativa de desempenho de rede. In *Anais do XLI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 490–503, Porto Alegre, RS, Brasil. SBC.
- Saad, M., Chaudhary, M., Karray, F., and Gaudet, V. (2020). Machine learning based approaches for imputation in time series data and their impact on forecasting. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2621–2627.
- Silveira, M. M., Portela, A. L., Menezes, R. A., Souza, M. S., Silva, D. S., Mesquita, M. C., and Gomes, R. L. (2023). Data protection based on searchable encryption and anonymization techniques. In *NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium*, pages 1–5.
- Tawn, R., Browell, J., and Dinwoodie, I. (2020). Missing data in wind farm time series: Properties and effect on forecasts. *Electric Power Systems Research*, 189:106640.
- Thakur, S., Choudhary, J., and Singh, D. P. (2021). A survey on missing values handling methods for time series data. In Sheth, A., Sinhal, A., Shrivastava, A., and Pandey, A. K., editors, *Intelligent Systems*, pages 435–443, Singapore. Springer Singapore.
- Vasileva, P., McKee, S., Penev, A., and Vukotic, I. (2021). Ps-dash –analysis, monitoring and visualization of network measurements. In *2021 International Conference Automatics and Informatics (ICAI)*, pages 93–96.
- Wu, S.-F., Chang, C.-Y., and Lee, S.-J. (2015). Time series forecasting with missing values. In *2015 1st International Conference on Industrial Networks and Intelligent Systems (INISCom)*, pages 151–156.
- Yang, H., Li, X., Qiang, W., Zhao, Y., Zhang, W., and Tang, C. (2021). A network traffic forecasting method based on sa optimized arima–bp neural network. *Computer Networks*, 193:108102.