

Modelo Adaptativo para Previsão de Recursos de Rede em Provedores de Internet Modernos

Dyego H. L. Oliveira e Rafael L. Gomes (Orientador)

¹Universidade Estadual do Ceará (UECE)
Fortaleza – CE – Brasil

{dyego, rafael.l.gom}@larces.uece.br

Abstract. *Modern Internet Service Providers (MISPs) need to address situations such as the elastic demand for network resources that can affect the Quality of Service (QoS). A promising approach to deal with elastic demand is the usage of a network traffic prediction technique. Within this context, this summary presents an adaptive network prediction model for MISPs that adjusts seasonality and trend and removes time series error cycles according to the behavior observed in network traffic. The results, using a real bandwidth data set, suggest that the proposed model improves the existing prediction models.*

Resumo. *Provedores de Internet Modernos (MISPs) precisam lidar com circunstâncias como a demanda elástica de recursos de rede que podem causar queda na Qualidade de Serviço (QoS). Uma abordagem promissora para lidar com demanda elástica é o uso de técnicas de previsão de tráfego. Neste contexto, este resumo apresenta o modelo adaptável de previsão de rede para MISPs, que ajusta sazonalidade e tendência e remove os ciclos de erro na série temporal de acordo com o comportamento observado no tráfego de rede. Os resultados, usando um conjunto de dados real, sugerem que o modelo proposto aprimora as técnicas de previsão tradicionais.*

1. Introdução

Recentemente, a Internet surgiu como o principal caminho para os serviços computacionais modernos (por exemplo, compartilhamento de conteúdo, sistemas inteligentes, automação de tarefas e outros), tornando-se uma entidade crucial para a sociedade moderna. A maioria desses serviços são baseados em Serviço de Acesso a Internet (*Internet Access Service - IAS*) por meio de Provedores de Internet (*Internet Service Provider - ISP*). Assim, os ISPs expandiram sua prestação de serviços, oferecendo diferentes alternativas, cabeada e sem fio, para acesso a Internet de seus clientes (como residências, empresas, universidades, etc.) [Foukas et al. 2017].

Estes serviços precisam abordar aspectos chave: baixo atraso, flexibilidade, resiliência, despesas de capital (*Capital Expenditure - CAPEX*) e despesas operacionais (*Operational Expenditure - OPEX*) compatíveis. Recursos que influenciam a qualidade de serviço (QoS) e a qualidade de experiência (QoE) dos usuários finais [Hou et al. 2018]. Nesse novo cenário, surge a ideia de Provedores de Internet Modernos (*Modern Internet Service Providers - MISPs*), que precisam aprimorar o gerenciamento de recursos, a flexibilidade da rede e a personalização do comportamento das infraestruturas de rede. Uma abordagem promissora para MISPs é a implantação de fatias de rede, ou seja, a divisão

dos recursos em redes lógicas isoladas. Cada fatia de rede pode ter a configuração mais adequada para melhor atender aos requisitos do cliente [Foukas et al. 2017].

Da mesma forma, um ponto importante para os MISPs manterem a qualidade do IAS é a demanda elástica por recursos de rede ao longo do dia, o que ocorre devido à mobilidade humana nas cidades. Nesse cenário, os MISPs precisam expandir ou reduzir dinamicamente a alocação de Largura de Banda (*Bandwidth* - Bw), caracterizando um comportamento elástico e permitindo um IAS sob demanda. Esse comportamento, quando não tratado adequadamente, pode causar problemas de lentidão, interrupção do serviço e desconexões constantes.

Uma abordagem promissora para lidar com serviços elásticos é o uso de um modelo de previsão de tráfego de rede [Bayati et al. 2018]. Contudo, os modelos tradicionais de previsão carecem de adaptabilidade para acompanhar as mudanças no comportamento das observações, uma vez que originalmente não realizam ajustes na sazonalidade e não removem possíveis ciclos de erro. Esse fato dificulta o uso dos modelos existentes nos MISPs para lidar com a situação de demanda elástica.

Este artigo resume a dissertação sobre o Modelo de Previsão de Rede Adaptável, chamado *PRA*, para permitir a alocação adequada de recursos de rede em cenários de demanda elástica. O modelo decompõe a série, testa a estacionariedade e corrige os erros de ciclo. Após a aplicação do *PRA* o conjunto de dados é submetido a técnicas de previsão como: Média Móvel Integrada Auto-Regressiva (ARIMA) e Rede Neural Auto-Regressiva (NNAR). Os experimentos foram realizados usando um conjunto de dados de demanda de largura de banda real da Universidade Estadual do Ceará (UECE). Os resultados indicam que o *PRA* minimiza a taxa de erro dos valores previstos, atingindo 30% de melhoria quando comparado aos modelos de previsão de rede existentes.

2. Inovação e originalidade em relação ao estado da arte

A partir do levantamento bibliográfico realizado, nota-se que os trabalhos existentes na literatura focam em predição de largura para atender diversos aspectos relacionados a melhora da Qualidade de Serviço, Planejamento e Segurança (como por exemplo as referências [Ruan et al. 2018, Hou et al. 2018, Katris and Daskalaki 2019, Wang et al. 2018]), mas nenhum desses trabalhos se concentrou no desenvolvimento de uma solução de previsão de largura de banda adaptável para demanda elástica em MISPs, que é o foco da dissertação.

O modelo desenvolvido realiza várias tarefas para aprimorar o processo de previsão: processamento de dados, ajuste sazonal, remoção de ciclos de erros, análise de tendências, verificação residual e amortização de ciclos irregulares, criando um modelo padronizado de amostras que facilita o processo de previsão das técnicas, reduzindo o tempo computacional de criação dos modelos. Mais detalhes sobre os trabalhos relacionados podem ser encontrados no Capítulo 3 da dissertação.

3. Modelo de Previsão de Rede Adaptável (*PRA*)

O Modelo de Previsão de Rede Adaptável (*PRA*) desenvolvido tem por objetivo ajustar o conjunto de dados referente à utilização de largura de banda para situações de demanda elástica, possibilitando assim uma maior eficiência das técnicas de predição existentes.

No geral, o modelo realiza as seguintes tarefas: (1) Decomposição de dados, (2) Teste de Estacionariedade e (3) Remoção de Ciclos. Posteriormente, após a aplicação do modelo PRA, os dados são submetidos a técnicas de predição. A descrição completa das etapas realizadas é feita no Capítulo 5 da dissertação.

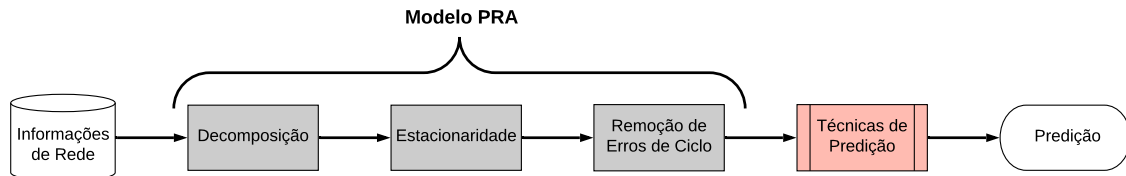


Figura 1. Funcionamento do modelo PRA proposto.

Inicialmente, o modelo recebe o conjunto de dados original referente a utilização de largura de banda, constituindo uma série temporal a ser analisada e ajustada de forma adaptável pelo modelo. Durante esta seção serão apresentados os resultados das etapas realizadas a fim de ilustrar de forma prática o funcionamento do modelo proposto. Serão detalhadas a seguir cada uma das etapas, onde a Decomposição dos dados, Testes de Estacionariedade e Remoção de Erros de Ciclo são descritas nas Seções 3.1, 3.2 e 3.3, respectivamente. Por fim, a Seção 3.4 apresenta o papel das técnicas de predição, bem como a interação das mesmas com o modelo PRA proposto.

3.1. Decomposição de dados

O objetivo da etapa de decomposição dos dados é compreender a série temporal presente no conjunto de dados de entrada. Essa compreensão engloba identificar a sazonalidade da série, identificação de tendência e verificação de ciclos. A *sazonalidade* em uma série corresponde às oscilações que sempre ocorrem em um determinado período do ano, do mês, da semana ou do dia. A *tendência* indica o seu comportamento a longo prazo e qual a frequência de mudanças. Os *ciclos* são caracterizados pelas oscilações de subida e de queda nas séries, de forma suave e repetida, sem regularidade, ao longo da componente de tendência. A Figura 2 ilustra um exemplo de *sazonalidade* e *tendência* da série temporal do conjunto de dados originais.

De posse dessas três informações sobre a série, pode-se encontrar *Erros*, que são informações/valores não compreendidos na sazonalidade e na tendência. Um componente de tendência é determinado usando uma média móvel, o valor sazonal é calculado pela média, em cada unidade de tempo e em todos os períodos, centralizando o componente sazonal.

3.2. Teste de Estacionariedade

O Teste de Estacionariedade verifica o comportamento a longo prazo da série temporal, sendo este um comportamento exponencial ou regular/constante. Esta etapa do modelo PRA é baseada no teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS), o qual verifica a hipótese de estacionariedade dos dados. O teste aplicado assume que a série temporal do comportamento da largura de banda (X_i , para $i = 1, 2, \dots, n$) pode ser representada pela soma de uma tendência determinística (k), uma caminhada aleatória (W_k) e um erro estacionário (E_k).

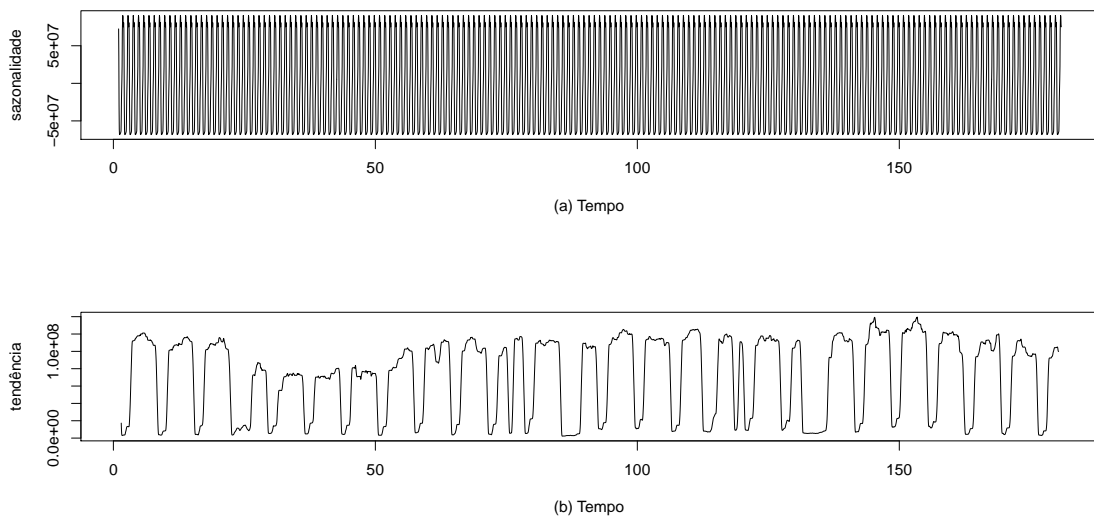


Figura 2. Exemplo de Sazonalidade e Tendência da série temporal original.

Adicionalmente, considera-se uma caminhada aleatória, onde U_k são uma coleção de variáveis aleatórias Independentes e Distribuídas Identicamente (*Independent and Identically Distributed - IID*) entre zero e σ_U^2 , ou seja, $U_k = IID(0, \sigma^2)$. A partir disso, a hipótese de estacionariedade é $\sigma^2 = 0$, visto que assume-se que E_k é estacionário sobre a hipótese nula que X_i tem tendência estacionária. Portanto, o resultado do teste KPSS revela o fator de estacionariedade da série temporal em todo o conjunto de ciclos, não rejeitando a hipótese nula.

3.3. Remoção de Erros de Ciclos

O uso médio de largura de banda em dias regulares atingem valores muito maiores que o uso médio em dias irregulares (finais de semana, feriados, eventos, etc.), caracterizando a baixa amplitude dos sinais em dias de ciclos irregulares.

No modelo PRA é aplicada uma técnica de média móvel/limiar de ordem 24 para remover as amostras correspondentes aos ciclos de baixa amplitude da série. O modelo executa adicionalmente uma robusta decomposição sazonal e de tendência usando Loess para estimar as relações de variáveis não lineares (valores omissos), removendo os *outliers* existentes.

A Figura 3(a) ilustra a série original sem ajustes. A partir da Figura 3(b) é possível observar o efeito da série temporal submetida ao algoritmo de correção de baixas amplitudes (ciclos irregulares). Posteriormente, na Figura 3(c), a série aparece praticamente livre de erros de ciclos, quando submetida ao processo de remoção de erros de ciclo. Por fim, é possível observar na Figura 3(d) que a etapa de remoção de ciclos completa as lacunas baseando-se em valores da série original.

Consideramos o modelo aditivo para relacionar as componentes não observáveis (fatos atípicos ou fortuitos) com as amostras que serão criadas pelo modelo de previsão, por identificar que as flutuações sazonais e tendência não variam com o nível da série, são constantes e regulares. Caso contrário, o modelo multiplicativo poderia ser utilizado.

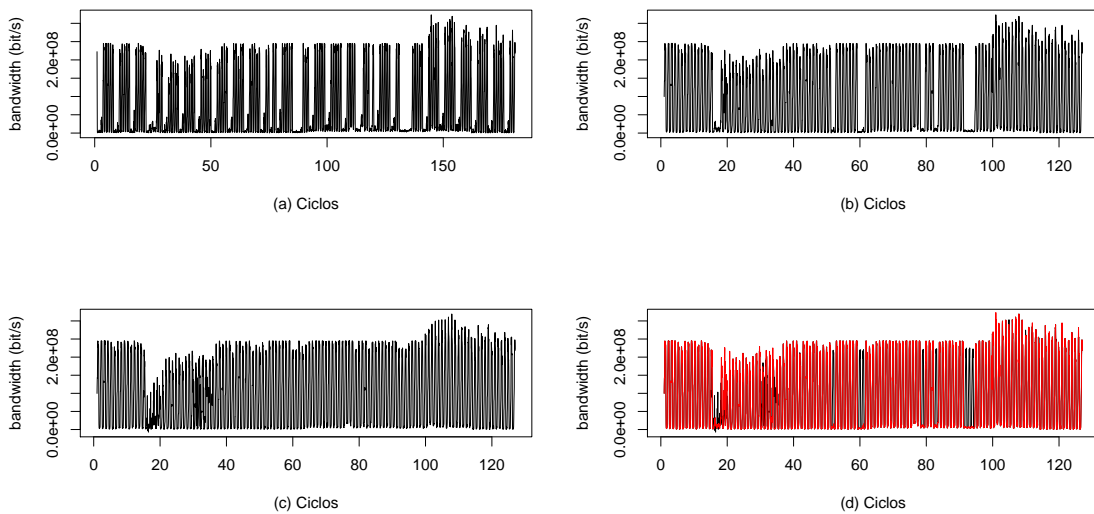


Figura 3. Evolução da correção dos ciclos irregulares nas amostras da série.

3.4. Técnicas de Predição

As técnicas de predição calculam passos adiante através de amostras passadas. Como exemplo de técnicas de predição tem-se: ARIMA, um modelo clássico da estatística que lida com sazonalidade $(p,d,q)(P,D,Q)$, onde as letras minúsculas se referem aos operadores não-sazonal, as maiúsculas à partes sazonais, onde, p é a ordem da parte autorregressiva, d o grau de diferenciações e q , ordem de média móvel; e NNAR (p,P,K) , um método de previsão que se baseia em modelos matemáticos simples do cérebro, permitindo relacionamentos não lineares entre a variável de resposta e seus preditores.

4. Resultados

O conjunto de dados usado descreve o uso real da largura de banda no *backbone* central da UECE [Oliveira and Gomes 2020]. Os dados foram monitorados durante um período de seis meses, sendo calculados a cada 60 minutos. Um período de 24 horas conclui um ciclo (assim, o conjunto de dados tem 4320 ciclos). Devido à dependência de tempo das séries, uma variação da técnica de Validação Cruzada de Série Temporal (TSCV) foi usada na previsão [Maleki et al. 2018].

Ambos os modelos (ARIMA e NNAR) foram utilizados para modelar as amostras afim de serem submetidas a previsão (ou seja, as próximas 24 horas de uso da largura de banda). Através do método *HoldOut* os subconjuntos Treino (utilizados na criação do modelo) e Teste (utilizados na validação) foram selecionados. Primeiro usando 72 amostras anteriores como conjunto de treinamento, depois 144 amostras e assim por diante (72, 144, 216, ...).

O *PRA* é usado como etapa prévia às técnica de predição, avaliando o incremento na capacidade de modelagem das mesmas. O Erro Quadrático Médio Raiz (RMSE) por um período T foi usado como métrica de avaliação. Figura 4 mostra os resultados de RMSE para previsão de 1 ciclo usando o modelo proposto e para o ARIMA e o NNAR tradicionais (sem a utilização do *PRA*). Similarmente, a Figura 5 ilustra o comportamento das predições realizadas pelo modelo proposto para determinados ciclos.

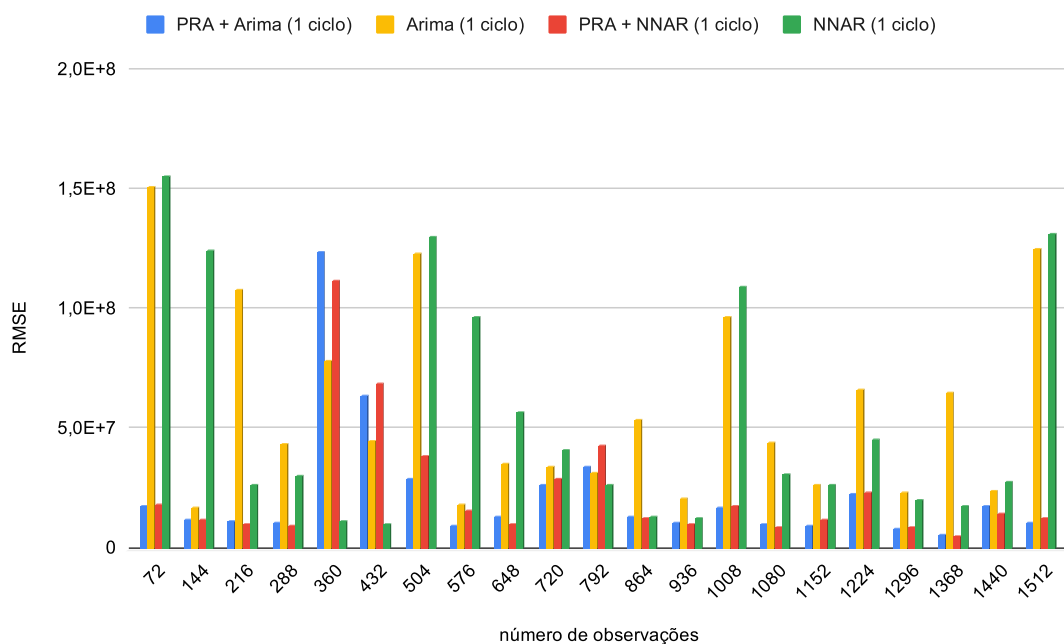


Figura 4. Comparação de Erro de Previsão.

Usando ARIMA e NNAR como técnicas de modelagem para previsão, obtêm-se os valores aceitáveis de generalização de 87% e 85%, respectivamente. Esse desempenho, que foi comparado ao conjunto de *Teste*, demonstra similaridade entre as duas técnicas, devido ao fato de a série temporal ser sido submetida ao *PRA*. É possível observar que os maiores erros obtidos, com 360 e 432 amostras de treinamento são efeitos do padrão que não fora corrigido o bastante, devido a grandes flutuações nos tempos 16 a 19 (similarmente ao ilustrado na Figura 3). Adicionalmente, os resultados de RMSE demonstraram uma proximidade entre as técnicas ARIMA e NNAR quando o modelo *PRA* é aplicado em conjunto (“*PRA+ARIMA*” e “*PRA+NNAR*”, respectivamente). Por outro lado, este fato não acontece quando as técnicas são usada sozinhas. Como pode ser visualizado na Figura 5, o modelo *PRA* consegue acompanhar as amostras da série temporal do conjunto de dados real, seguindo o comportamento sazonal e flutuações de amplitude, dando importância a variação das amostras, com aproximadamente 2, 4, 6 e 8 semanas.

Além disso, o modelo *PRA* proposto conseguiu reduzir o processamento computacional na fase de treinamento (estimação dos parâmetros) em cerca de 85% utilizando o ARIMA e cerca de 12% utilizando NNAR. Esse fato indica que o modelo *PRA* é capaz de reduzir o tempo de processamento para outros métodos de previsão, habilitando assim sua aplicação em um cenário de previsão em tempo real. A descrição completa dos experimentos pode ser encontrada no Capítulo 6 da dissertação, onde os gráficos de RSME para cada caso são analisados de forma mais abrangente.

5. Conclusão

O modelo de Previsão de Rede Adaptável (*PRA*) desenvolvido aprimora as técnicas de predição existentes e assim torna-se uma abordagem eficaz para lidar com as situações de demanda elástica. Os experimentos realizados, com uma base de dados real, mostraram

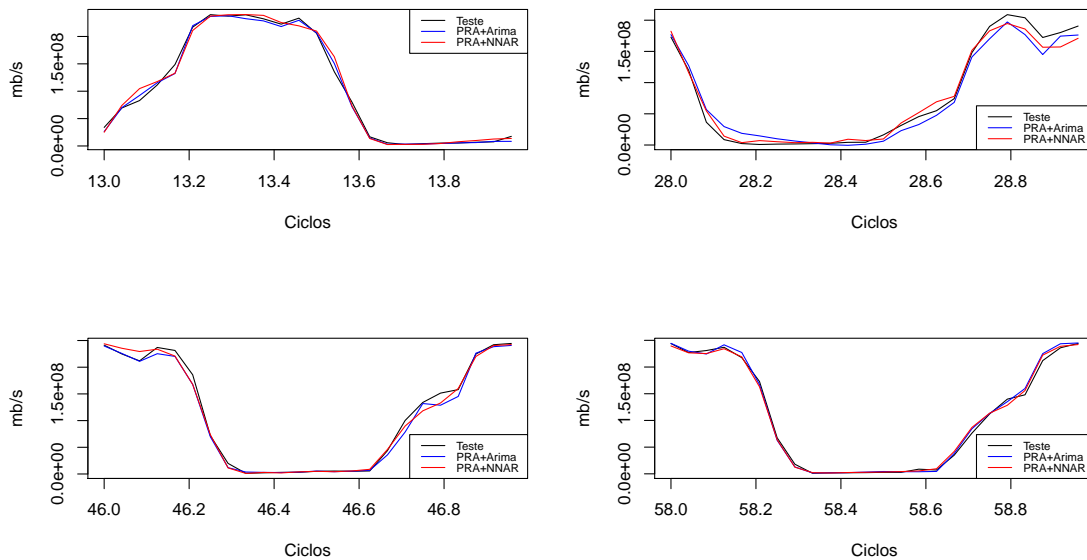


Figura 5. Relação entre os valores reais (Teste) e os valores previstos.

que a aplicação do modelo em conjunto com técnicas de predição evolui a capacidade de predição dessas técnicas quando comparado a utilização delas isoladamente. A utilização do modelo PRA proposto diminuiu em cerca de 30% a taxa de erros RMSE. Além disso, os resultados de RMSE demonstraram uma proximidade entre as técnicas utilizadas na previsão, comprovando flexibilidade do modelo proposto e a independência do mesmo em relação às técnicas de predição usadas.

6. Contribuições e Produção Científica

A dissertação apresentada neste artigo apresenta as seguintes contribuições: (I) Uma análise sobre os trabalhos existentes relacionados a predição de demanda de recursos; (II) Um modelo de previsão de demanda adaptável para cenários com demanda elástica; (III) Um conjunto de dados real sobre a demanda de largura de banda de 6 meses da UECE [Oliveira and Gomes 2020]; e, (IV) A aplicação do modelo desenvolvido em uma abordagem de fatiamento de rede (realizada após a data de defesa, mais detalhes na referência [Oliveira et al. 2021]). As contribuições elencadas possibilitaram a publicação de **1 Artigo Qualis B4** [Oliveira et al. 2020], **1 Artigo Qualis A2** [Oliveira et al. 2020], **1 Artigo Qualis A1** [Oliveira et al. 2021] e **1 Registro de Software no INPI** (Nº 512020001816-5 em 07/07/2020).

Referências

- Bayati, A., Khoa Nguyen, K., and Cheriet, M. (2018). Multiple-step-ahead traffic prediction in high-speed networks. *IEEE Communications Letters*, 22(12):2447–2450.
- Foukas, X., Patounas, G., Elmokashfi, A., and Marina, M. K. (2017). Network slicing in 5g: Survey and challenges. *IEEE Communications Magazine*, 55(5):94–100.
- Hou, Z., She, C., Li, Y., Quek, T. Q. S., and Vucetic, B. (2018). Burstiness aware bandwidth reservation for uplink transmission in tactile internet. In *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6.

- Katris, C. and Daskalaki, S. (2019). Dynamic bandwidth allocation for video traffic using farima-based forecasting models. *Journal of Network and Systems Management*, 27(1):39–65.
- Maleki, A., Nasserli, S., Aminabad, M. S., and Hadi, M. (2018). Comparison of ARIMA and NNAR Models for Forecasting Water Treatment Plant's Influent Characteristics. *KSCCE Journal of Civil Engineering*, 22(9):3233–3245.
- Oliveira, D., Filho, F., de Araújo, T., Júnior, J. C., and Gomes, R. (2020). Modelo adaptativo para previsão de recursos de rede em provedores de internet modernos. In *Anais do XXV Workshop de Gerência e Operação de Redes e Serviços*, pages 209–222, Porto Alegre, RS, Brasil. SBC.
- Oliveira, D. H. L., de Araujo, T. P., and Gomes, R. L. (2021). An adaptive forecasting model for slice allocation in softwarized networks. *IEEE Transactions on Network and Service Management*, 18(1):94–103.
- Oliveira, D. H. L., Filho, F. M. V., de Araújo, T. P., Celestino, J., and Gomes, R. L. (2020). Adaptive model for network resources prediction in modern internet service providers. In *2020 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–6.
- Oliveira, D. H. L. and Gomes, R. L. (2020). Bandwidth usage of university campus. DOI: 10.21227/jw40-y336. IEEE Dataport.
- Ruan, L., Mondal, S., and Wong, E. (2018). Machine learning based bandwidth prediction in tactile heterogeneous access networks. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–2.
- Wang, T., Guo, Z., Chen, H., and Liu, W. (2018). Bwmanager: Mitigating denial of service attacks in software-defined networks through bandwidth prediction. *IEEE Transactions on Network and Service Management*, 15(4):1235–1248.