

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

Dyego H. L. Oliveira¹, Francisco M. V. Filho¹, Thelmo P. de Araujo¹,
Joaquim Celestino Jr.¹, Rafael L. Gomes³

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

{dyego,mardoniovf,thelmo,celestino,rafaellllgom}@larces.uece.br

Resumo. Atualmente, os Provedores de Internet (ISPs) tendem a evoluir para Provedores de Internet Modernos (MISPs), a fim de lidar com situações 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. Contudo, tais técnicas não atingem fatores de correção necessários quando o comportamento da rede não segue um padrão claro (como nas situações de demanda elástica). Neste contexto, este artigo apresenta um 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.

Abstract. Currently, Internet Service Providers (ISPs) tend to evolve to Modern Internet Service Providers (MISPs), addressing 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. However, such techniques do not reach the necessary correction factor when the behavior of the network traffic is not clear (as in elastic demand situations). Within this context, this paper 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.

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 um 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].

Independentemente do tipo de rede de acesso, todos precisam abordar aspectos chave: baixo atraso, flexibilidade, resiliência, despesas de capital (*Capital Expenditure* -

CAPEX) e despesas operacionais (*Operational Expenditure* - OPEX) compatíveis. Esses recursos 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, resultando em usuários frustrados e quebra dos requisitos do Acordo de Nível de Serviço (*Service Level Agreement* - SLA).

Uma abordagem promissora para lidar com serviços elásticos é o uso de um modelo de previsão de tráfego de rede [Bayati et al. 2018]. Um modelo de previsão de tráfego de rede permite a compreensão do comportamento da rede, por meio de observações anteriores do uso de recursos da rede (histórico), e a previsão de valores futuros da demanda da rede, permitindo a aplicação de tarefas proativas para evitar os problemas mencionados anteriormente e planejar a infraestrutura de rede.

As abordagens tradicionais não modelam as séries temporais para minimizar as baixas amplitudes dos ciclos regulares, usando apenas reduções automáticas baseadas na média móvel dos próprios modelos de previsão. Como exemplo, tem-se o Holt-Winters [Yar and Chatfield 1990, Chatfield and Yar 1988], técnica que escolhe estatisticamente o modelo aditivo ou multiplicativo e não elimina as componentes regulares (ciclos), trabalhando apenas com sazonalidade e tendência.

Dessa maneira, 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 [Adhikari and Agrawal 2013]. Esse fato dificulta o uso dos modelos existentes nos MISPs para lidar com a situação de demanda elástica devido à variabilidade do comportamento do cliente ao longo do tempo (seja durante o dia, semana ou mês). Portanto, os MISPs precisam de uma solução para realizar uma previsão adaptativa.

Dentro deste contexto, este artigo apresenta um modelo de Previsão de Rede Adaptável, chamado *PRA*, para permitir a alocação adequada de recursos de rede e um planejamento estratégico da infraestrutura de rede. O modelo *PRA* proposto decompõe os dados originais da série, corrige os erros de ciclo e testa a estacionariedade. Depois de aplicado o modelo *PRA* proposto, o conjunto de dados pode ser submetido a técnicas de previsão, como por exemplo: Média Móvel Integrada Auto-Regressiva (ARIMA) [Su et al. 2016] e Rede Neural Auto-Regressiva (NNAR) [Maleki et al. 2018], de acordo com o comportamento de tráfego de rede observado.

De maneira geral, aplicando o PRA, pode-se corrigir dados como tendência e componentes irregulares, habilitando a diferenciação a partir da tendência e remoção dos componentes irregulares, passando a série mais estruturada para a técnica preditiva. Sendo assim, os modelos sem o PRA, dão maior importância as amostras mais próximas da atual (que será predita), minimizando o peso de amostras antigas (priorizando as imediatamente anteriores), limitando a capacidade de predição das técnicas tradicionais.

Os experimentos foram realizados usando um conjunto de dados de demanda de largura de banda real da Universidade Estadual do Ceará (UECE - Brasil). Este conjunto de dados tem a utilização média da largura de banda para a Internet a cada hora dos primeiros seis meses de 2019. Os resultados indicam que o modelo de previsão adaptativa proposto minimiza a taxa de erro dos valores previstos, atingindo 30% de melhoria quando comparado aos modelos de previsão de rede existentes [Yoo and Sim 2015].

O restante deste artigo está organizado da seguinte forma: A Seção 2 apresenta os trabalhos relacionados existentes. A Seção 3 apresenta o modelo de previsão adaptativa proposto, enquanto a Seção 4 descreve os resultados dos experimentos realizados. Por fim, a Seção 5 conclui o artigo e apresenta trabalhos futuros.

2. Trabalhos Relacionados

Esta seção descreve os principais trabalhos relacionados sobre a demanda de recursos e estratégias de previsão para redes de computadores. A Tabela 1 resume os trabalhos existentes na literatura, destacando as diferenças em relação a proposta, onde a coluna *Contexto* apresenta o ambiente em que o trabalho relacionado atua, enquanto as colunas *Estratégia* e *Foco* informam a estratégia aplicada e o objetivo do artigo, respectivamente.

Bayati et al. [Bayati et al. 2018] propõem um algoritmo, para redes de alta velocidade, para modelar diferentes escalas de tempo usando a Regressão de Processo Gaussiana (*Gaussian Process Regression* - GPR), abordando a propagação de erros na previsão de tráfego com várias etapas à frente. A previsão em uma escala de tempo é feita usando os dados dessa escala e os resultados da previsão em escalas de tempo maiores. No entanto, esse algoritmo não considera uma abordagem adaptativa para lidar com o comportamento da demanda de recursos.

Hou et al. [Hou et al. 2018] aplicam o método Neyman-Pearson para classificar o processo de chegada de cada usuário, criando uma reserva de largura de banda com reconhecimento de rajadas para cenários da Internet Tátil. Os autores diferenciam os usuários nos estados de alto e baixo tráfego e otimizam a largura de banda reservada para atender aos requisitos de latência e confiabilidade. Contudo, a abordagem aplicada pelos usuários considera apenas a chegada de pacotes de acordo com o comportamento das rajadas de tráfego, limitando sua aplicabilidade em MISPs.

Ruan et al. [Ruan et al. 2018] apresentam uma abordagem de aprendizado de máquina para prever a demanda de largura de banda em redes ópticas sem fio heterogêneas na Internet Tátil, chamada MLP-DBA. O MLP-DBA concentra-se nos requisitos de minimização de atraso de upstream através da alocação de largura de banda preditiva de acordo com o status de cada unidade de rede óptica. Todavia, a aplicação do MLP-DBA nos MISPs é comprometida por sua restrição às informações das redes óptico-sem fio.

Wang et al. [Wang et al. 2018] propõem o BWManager, que mitiga os ataques de negação de serviço (DoS) no controlador de redes SDN. Basicamente, o BWManager prevê o consumo de largura de banda dos usuários para determinar os valores de confiança e as filas de prioridade para proteção. O mecanismo prevê a utilização da largura de banda para cada usuário com base em dados históricos do contador de bytes dos switches, analisando a diferença em dois momentos através das técnicas ARIMA e Heteroskedasticity Condicional AutoRegressiva (*Autoregressive Conditional Heteroskedasticity* - ARCH).

Yoo et al. [Yoo and Sim 2015] desenvolveram um modelo para prever a utilização esperada da largura de banda em redes de alta velocidade, com o objetivo de melhorar a eficiência da utilização de recursos e agendar movimentos de dados científicos em larga escala. O modelo de previsão é baseado na decomposição sazonal de séries temporais por Loess (STL) e ARIMA, nos quais são submetidos os dados de utilização do caminho SNMP (Simple Network Management Protocol). O modelo desenvolvido trabalha especificamente com os dados do protocolo SNMP, ou seja, é limitado à redes baseadas em SNMP. Esse fato restringe a aplicação do modelo desenvolvido nos MISPs.

Aldhyani et al. [Aldhyani and Joshi 2016] propõem um modelo de previsão que combina modelos de séries temporais com abordagens de agrupamento suave (como Fuzzy C-Means - FCM e Rough K-Means - RKM) para permitir a alocação de largura de banda e controle de congestionamento. Esse modelo integrado evita que os dados não-lineares e voláteis afetem as escolhas realizadas pelos modelos de previsão. No entanto, esse modelo integrado não executa uma abordagem adaptativa para atender aos requisitos de demanda de recursos elásticos.

Katris et al. [Katris and Daskalaki 2019] desenvolveram modelos de séries temporais no tráfego de vídeo como parte de três esquemas de alocação dinâmica de largura de banda. Os autores tentam melhorar a precisão das previsões de tráfego de vídeo usando a Média Móvel AutoRegressiva Integrada Fracionada (FARIMA) e a Heterocedasticidade Condicional AutoRegressiva Generalizada (GARCH) com redes neurais. As previsões de tráfego aplicam um método de seleção de não linearidade a ser combinado com o modelo. O uso do modelo de previsão implementado é específico para o tráfego de vídeo, uma vez que considera informações de buffer e taxa de utilização. Desta forma, a sua aplicação no contexto MISP é limitada.

Harstead et al. [Harstead and Sharpe 2015] descrevem um modelo que prevê demandas de largura de banda de assinantes agregados em redes residenciais de acesso fixo. O modelo permite que os operadores de rede dimensionem suas redes e façam os investimentos corretos para o futuro. Os autores usaram técnicas estatísticas para quantificar o número de fluxos de vídeo simultâneos, a mistura de resoluções padrão, alta definição e ultra-alta definição, ganho de multicast e a tendência de entrega multicast para unicast desses fluxos.

A partir do levantamento bibliográfico realizado, nota-se que nenhum artigo da literatura se concentrou no desenvolvimento de um modelo de previsão de largura de banda adaptável para demanda elástica em MISPs, que é o foco deste artigo. Nossa proposta 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 faci-

Tabela 1. Trabalhos Relacionados

Referência	Estratégia	Foco
Bayati et al. [Bayati et al. 2018]	Escalas de tempo usando GPR	Previsão de várias etapas à frente
Hou et al. [Hou et al. 2018]	Método de Neyman-Pearson	Reserva de Bw
Ruan et al. [Ruan et al. 2018]	Alocação de Bw preditiva	Minimização de atraso do <i>Upstream</i>
Wang et al. [Wang et al. 2018]	ARIMA e ARCH	Mitigação de ataques DoS
Yoo et al. [Yoo and Sim 2015]	ARIMA e STL	Utilização de recursos e escalonamento
Aldhyani et al. [Aldhyani and Joshi 2016]	Agrupamento suave e séries temporais	Previsão de tráfego de vídeo
Katris et al. [Katris and Daskalaki 2019]	FARIMA e GARCH com redes neurais	Predição de tráfego de vídeo
Harstead et al. [Harstead and Sharpe 2015]	Técnicas estatísticas para quantificação	Planejamento de infraestrutura de rede
Nossa Proposta	Modelo Adaptativo	Predição de Recursos em Demanda Elástica

lita o processo de previsão das técnicas, reduzindo o tempo computacional de criação dos modelos.

3. Proposta

Esta sessão apresenta a descrição detalhada do modelo de Previsão de Rede Adaptável (*PRA*) proposto. O modelo *PRA* proposto 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. Tais técnicas tentam lidar com os aspectos de sazonalidade e tendência, mas não atingem fatores de correção necessários e, conseqüentemente, não conseguem prever o comportamento dos dados de forma adequada quando estes não demonstram um padrão bem claro (como nas situações de demanda elástica).

O modelo *PRA* realiza as seguintes tarefas sequencialmente: (1) Decomposição de dados, (2) Remoção de Ciclos e (3) Teste de Estacionariedade. Posteriormente, após a aplicação do modelo *PRA*, os dados podem ser inseridos em alguma técnica de predição, para a predição ser realizada. A Figura 1 ilustra o processo descrito, onde as caixas em cinza representam as tarefas 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. Uma mostra com os dados originais da série, sem ajustes, pode ser visualizada na Figura 2. Durante esta seção serão apresentados os resultados das etapas realizadas utilizando essa mesma amostra de dados, a fim de ilustrar de forma prática o funcionamento do modelo proposto.

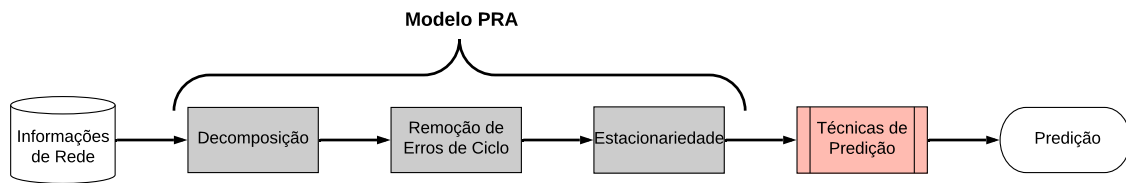


Figura 1. Funcionamento do modelo PRA proposto.

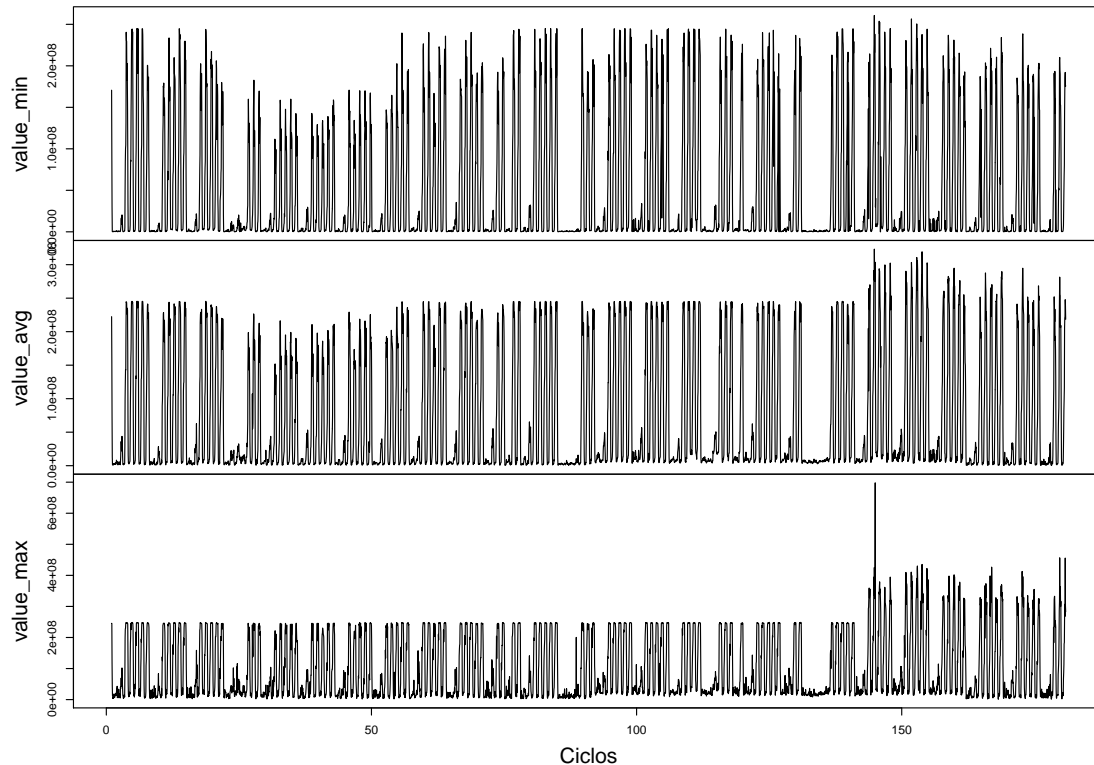


Figura 2. Dados de largura de banda da série temporal contendo: valor mínimo, médio e máximo.

A seguir serão detalhadas cada uma das etapas realizadas pelo modelo, onde a Decomposição dos dados, Testes de Estacionariedade e Remoção de Erros de Ciclo são descritas nas Seções 3.1, 3.3 e 3.2, respectivamente. Por fim, a Seção 3.4 apresenta o papel das técnicas de predição, bem 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. O método estatístico de decomposição clássica onde utiliza-se de médias móveis para extrair e examinar os componentes de sazonalidade (S_t), ciclos de tendência (R_t), e erro (N_t), onde (L_t) corresponde aos dados da série observados, descritos pela Função 1.

$$L_t = f(S_t, R_t, N_t) . \quad (1)$$

A *sazonalidade* (S_t) em uma série corresponde às oscilações que sempre ocorrem em um determinado período p do ano, do mês, da semana ou do dia, para cada:

$$S_{t+p} = S_t. \quad (2)$$

A *tendência* de uma série indica o seu comportamento a longo prazo (se ela cresce, decresce ou permanece estável) 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 3 ilustra um exemplo de *sazonalidade* e *tendência* da série temporal do conjunto de dados apresentado na Figura 2.

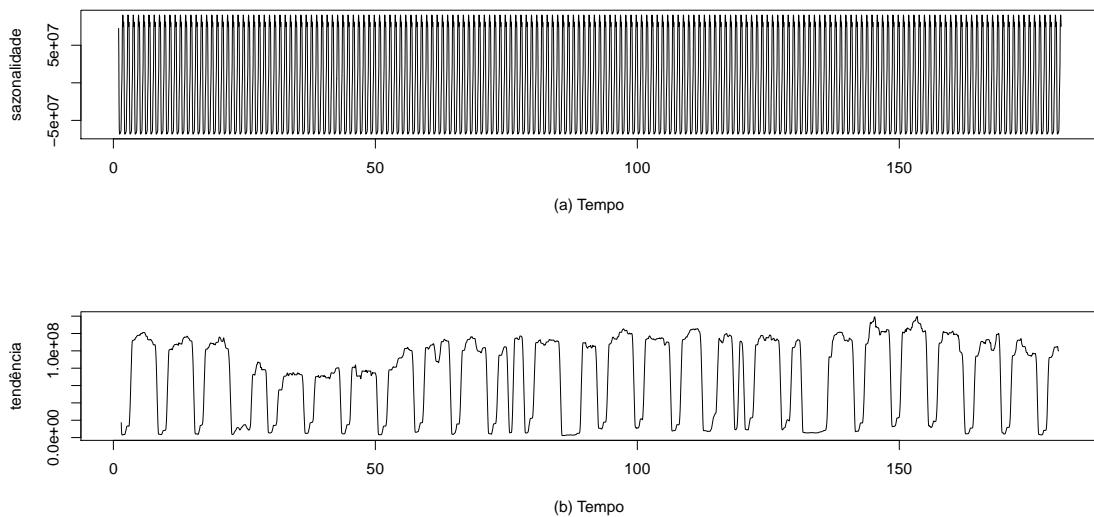


Figura 3. Exemplo de Sazonalidade e Tendência da série temporal original.

De posse dessas informações, pode-se encontrar *Erros*, que são informações/valores não compreendidos na sazonalidade e na tendência. Para isso, um componente de tendência é determinado usando uma média móvel, em seguida, o valor sazonal é calculado pela média, em cada unidade de tempo e em todos os períodos, centralizando o componente sazonal. Por fim, após remover tendência e sazonalidade das amostras originais, o erro é calculado.

A partir da obtenção desse conhecimento (sazonalidade, tendência, ciclos e erros) o modelo proposto passa para a etapa relacionada a Remoção de Erros de Ciclo, descrita a seguir.

3.2. Remoção de Erros de Ciclos

Devido às características do uso da largura de banda em situações de demanda elástica, o uso médio de largura de banda em dias regulares atinge valores muito maiores que o uso médio em dias irregulares (finais de semana, feriados, eventos, etc.). Neste tipo de situação, as técnicas de previsão não conseguem distinguir a baixa amplitude dos sinais em dias de ciclos irregulares.

Esses ciclos de baixa amplitude interferem na análise de tendência e sazonalidade e devem ser removidos. No modelo PRA proposto é aplicada uma técnica de média móvel centrada de ordem 24 para analisar a tendência a longo prazo e, a partir disso, remover as amostras correspondentes aos ciclos de baixa amplitude da série. Em séries temporais a fórmula geral é definida de acordo com a Equação 3, onde a média dada pela ordem S é calculada sobre o conjunto de amostras em cada período t .

$$MMC_t^S = \frac{y_{t+\frac{s}{2}} \dots + y_{t+1} + y_t + y_{t-1} + \dots + y_{t-\frac{s}{2}}}{S}. \quad (3)$$

Os períodos extremos são somados e multiplicados por 0,5, então, guarda-se o valor da média nesse índice. Se a ordem da média fosse ímpar, $s - 1$ seria utilizado no numerador. Entretanto, a aplicação da técnica não é suficiente para remover todas as imperfeições

A partir disso, o modelo executa adicionalmente uma robusta decomposição sazonal e de tendência usando Loess [Dokumentov and Hyndman 2015] para estimar as relações de variáveis não lineares (valores omissos), removendo os *outliers* existentes. A mudança na configuração da série de dados original e da série pós aplicação do modelo proposto (sem os ciclos de baixa amplitude) podem ser vistas na Figura 4.

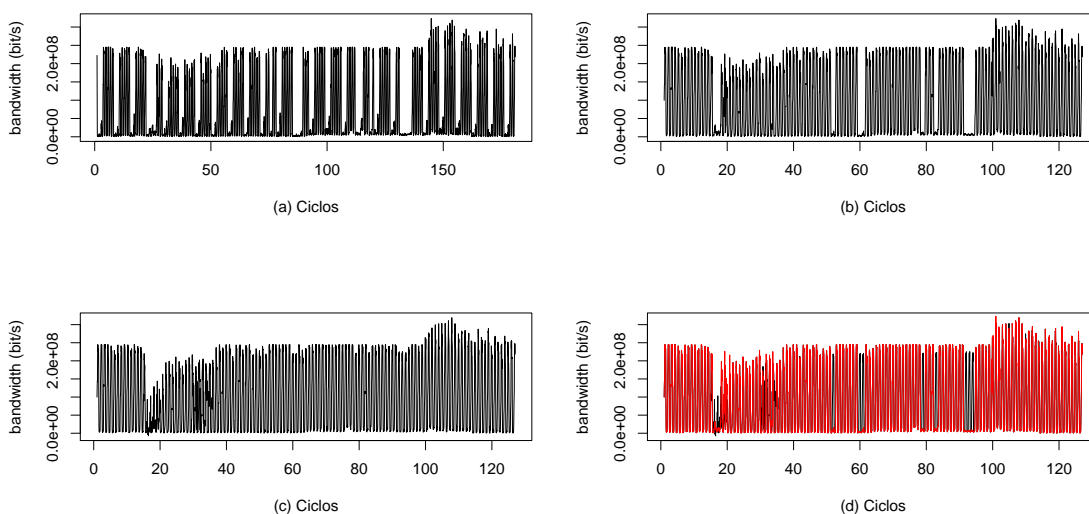


Figura 4. Evolução da correção dos ciclos irregulares nas amostras da série.

A Figura 4(a) ilustra a série original sem ajustes. A partir da Figura 4(b) é possível observar o efeito da série temporal submetida ao algoritmo de correção de baixas amplitudes (ciclos irregulares). Posteriormente, na Figura 4(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 4(d) que a etapa de remoção de ciclos completa as lacunas baseando-se em valores da série original. Uma vez removidos os ciclos de baixa amplitude, pode-se prosseguir com a análise da tendência e sazonalidade da série.

Apos a decomposição, 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. Este processo é representado na Equação 4, onde L_t são os dados no tempo t , com $t \in \{1, \dots, T\}$, S_t é o componente sazonal, R_t é a tendência, e N_t é uma média zero com ruído aleatório e variância constante.

$$L_t = S_t + R_t + N_t . \quad (4)$$

As Figuras 5(a) e 5(b) mostram a sazonalidade e a tendência, respectivamente, dos dados ajustados (sem os ciclos de baixa amplitude).

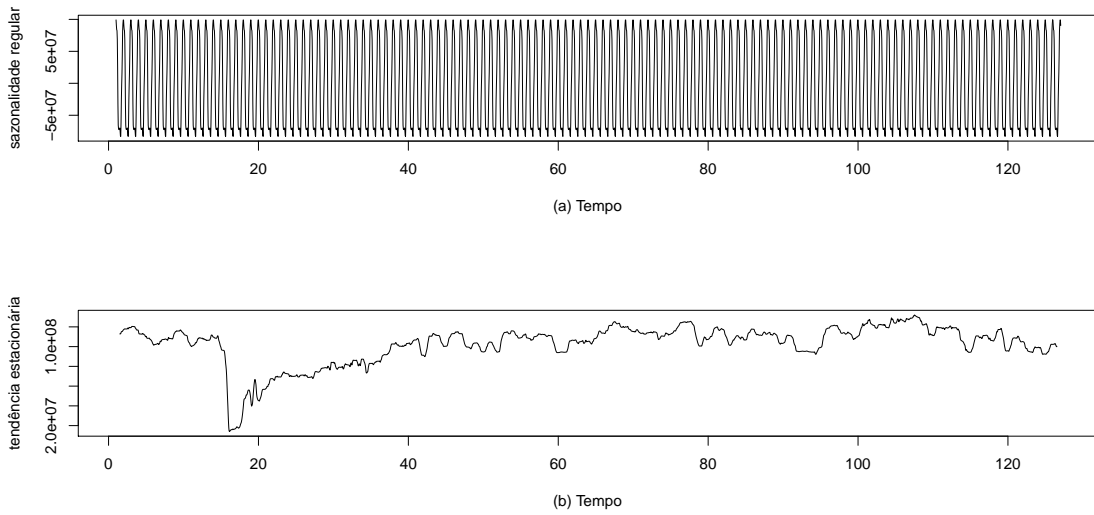


Figura 5. Sazonalidade constante e tendência estacionária da série temporal ajustada.

3.3. Teste de Estacionariedade

O Teste de Estacionariedade visa verificar o comportamento a longo prazo da serie temporal do conjunto de dados, sendo este um comportamento exponencial ou regular/constante. A série temporal pode não ser estacionária, implicando na mudança de nível ou inclinação em períodos curtos ou longos. Este fato afeta a capacidade de predição das técnicas existentes [De O. Schmidt et al. 2014, Schmidt et al. 2013].

Esta etapa do modelo PRA proposto é baseada no teste Kwiatkowski-Phillips-Schmidt-Shin (KPSS) [Carrion-I-Silvestre et al. 2001], 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), de acordo com a Equação 5.

$$X_i = k + W_k + E_k . \quad (5)$$

Adicionalmente, considera-se uma caminhada aleatória como representado na Equação 6, 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, $U_k = IID(0, \sigma^2)$.

$$W_k = W_{k-1} + U_k. \quad (6)$$

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 [Kwiatkowski et al. 1992]. Caso não seja atestada a estacionariedade, a série pode ser corrigida pelo processo de diferenciação, subtraindo o valor de um período t do seu antecessor, como mostra a Equação 7.

$$Y_{dt} = Y_t - Y_{t-1}. \quad (7)$$

A fim de aumentar a confiabilidade do modelo e padronização das amostras, o teste de estacionariedade proposto é aplicado em todos os conjuntos de treinamento, aceitando a estacionariedade e avaliando a presença de raiz unitária como nulas. Essa comprovação é obtida quando o teste de hipótese da série temporal atinge valores abaixo de 0,383, que é considerado crítico para um nível de confiança de confirmação conjunta [Carrion-I-Silvestre et al. 2001].

3.4. Técnicas de Predição

As técnicas de predição calculam passos adiante através de amostras passadas. Tais técnicas tentam acompanhar sazonalidade e tendência, mas possuem limitações de capacidade de correção em casos de séries temporais que não possuem um padrão bem definido.

Como exemplo de técnicas de predição tem-se Média Móvel Integrada Auto-Regressiva (ARIMA) [Su et al. 2016] e Rede Neural Auto-Regressiva (NNAR). ARIMA é 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 parte sazonais, onde, p é a ordem da parte autorregressiva, d o grau de diferenciações e q , ordem de média móvel.

Por outro lado, NNAR (p,P,K) [Maleki et al. 2018] é 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, análogo ao ARIMA citado anteriormente, mas sem funções lineares.

4. Resultados

Esta seção mostrará os resultados obtidos a partir da análise dos experimentos realizados a partir de um conjunto de dados real. O conjunto de dados descreve o uso da largura de banda (volume de tráfego) no *backbone* central da Universidade Estadual do Ceará (UECE). Os dados foram monitorados em tempo real durante um período de seis meses, com uso médio, mínimo e máximo sendo calculados a cada 60 minutos. Um

período de 24 horas conclui um ciclo, portanto, todo o conjunto de dados consiste em 4320 amostras.

Devido à dependência de tempo das séries, uma variação da técnica de Validação Cruzada de Série Temporal (TSCV) [Maleki et al. 2018] foi usada na previsão. 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 devido a grande quantidade de amostras. 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 a aplicação de uma técnica de predição, a fim de avaliar o incremento na capacidade de modelagem das mesmas, ARIMA e NNAR, com a aplicação do modelo proposto (assim como ilustrado na Figura 1).

A partir disso, o Erro Quadrático Médio Raiz (RMSE) por um período T foi usado como métrica de avaliação, seguindo a Equação 8, onde \hat{y}_t é o valor estimado e y_t é o valor real do uso da largura de banda no momento t .

$$\text{RMSE}(T) = \frac{1}{\sqrt{T}} \left(\sum_{t=1}^T (\hat{y}_t - y_t)^2 \right)^{\frac{1}{2}} . \quad (8)$$

A Figura 6 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 7 ilustra o comportamento das predições realizadas pelo modelo proposto para determinados ciclos.

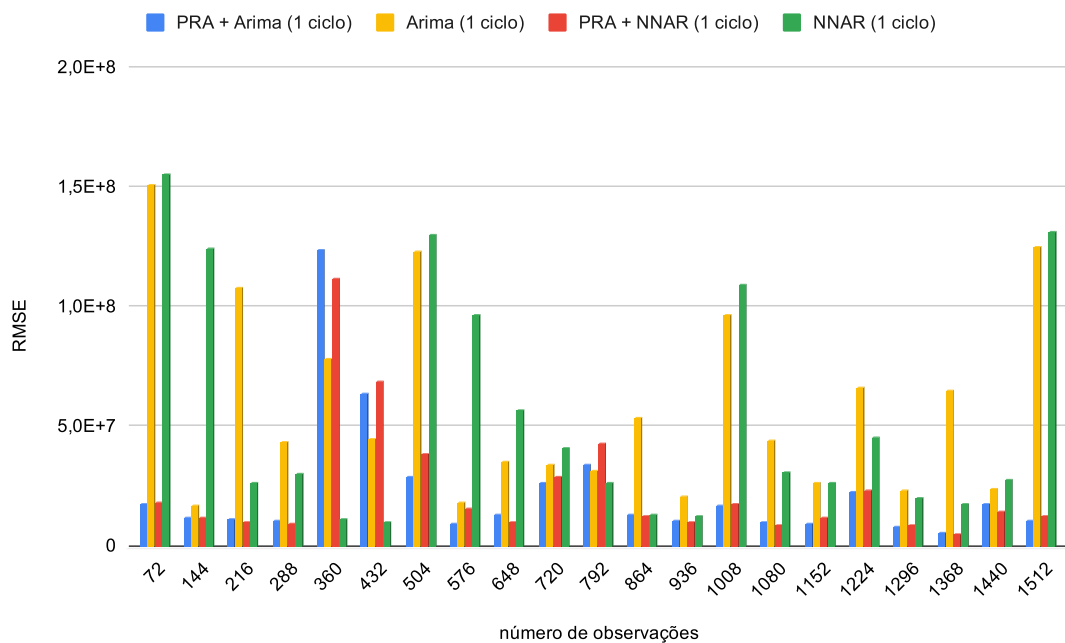


Figura 6. Comparação de Erro de Previsão.

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 4).

Adicionalmente, a partir dos dados apresentados na Figura 6, percebe-se que 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. Desta forma, pode-se concluir que o modelo *PRA* proposto possui flexibilidade e a independência em relação as técnicas de predição usadas.

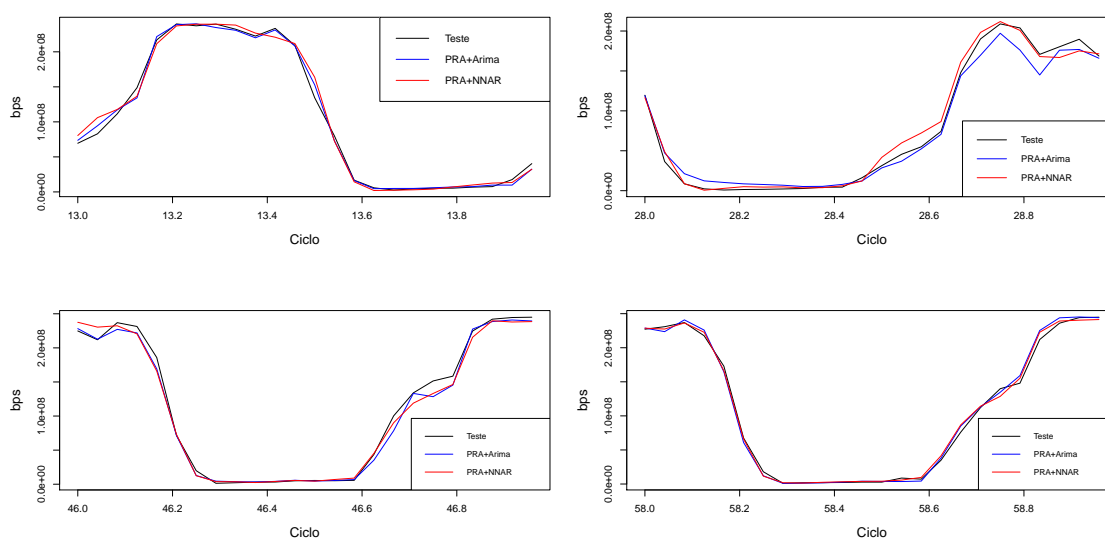


Figura 7. Relação entre os valores reais (Teste) e os valores previstos.

Como pode ser visualizado na Figura 7, 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. Esta situação ocorre pois o modelo *PRA* remove as variáveis de baixa importância durante a retirada dos erros de ciclos, ausentando variáveis utilizadas pelo método regressivo das técnicas de predição. 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.

5. Conclusão

A Internet se tornou uma ferramenta crucial para a prestação de serviços de rede. Os ISPs tendem a evoluir MISPs, a fim de lidar com situações que podem influenciar a

QoS da prestação de serviço. Uma delas é demanda elástica de recursos de rede, que pode causar problemas como lentidão, interrupção e desconexões constantes.

Neste trabalho apresentamos o modelo de Previsão de Rede Adaptável (PRA) para aprimorar as técnicas de predição existentes e assim tornar-se uma abordagem eficaz para lidar com as situações de demanda elástica. O modelo proposto adapta-se a utilização de diferentes médias de largura de banda por diferentes intervalos de ciclos. Adicionalmente, o modelo PRA realiza sequencias de ajustes que podem ser submetidas, atingindo valores ótimos de generalização.

Os experimentos realizados, com uma base de dados real, mostraram 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 utilizados na previsão, o que não ocorre quando as técnicas são usada sozinhas, comprovando flexibilidade do modelo proposto e a independência do mesmo em relação as técnicas de predição usadas.

Como trabalho futuro, pretende-se estender a forma de predição, incluindo aspectos como variáveis independentes e amostras da base de dados real. Espera-se que esta inclusão gere um novo modelo com maior capacidade de entender o comportamento elástico da largura de banda. Além disso, faz-se necessário a avaliação do modelo em outros conjuntos de dados para uma melhor análise da proposta.

Referências

- Adhikari, R. and Agrawal, R. K. (2013). An introductory study on time series modeling and forecasting. *CoRR*, abs/1302.6613.
- Aldhyani, T. H. H. and Joshi, M. R. (2016). Integration of time series models with soft clustering to enhance network traffic forecasting. In *2016 Second International Conference on Research in Computational Intelligence and Communication Networks (ICR-CICN)*, pages 212–214.
- 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.
- Carrion-I-Silvestre, J. L., Sansó-I-Rosselló, A., and Ortuño, M. A. (2001). Unit root and stationarity tests' wedding. *Economics Letters*, 70(1):1–8.
- Chatfield, C. and Yar, M. (1988). Holt-winters forecasting: Some practical issues. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 37(2):129–140.
- De O. Schmidt, R., Sadre, R., Melnikov, N., Schönwälder, J., and Pras, A. (2014). Linking network usage patterns to traffic Gaussianity fit. *2014 IFIP Networking Conference, IFIP Networking 2014*.
- Dokumentov, A. and Hyndman, R. J. (2015). STR : A Seasonal-Trend Decomposition Procedure Based on Regression. (13):23.
- 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.

- Harstead, E. and Sharpe, R. (2015). Forecasting of access network bandwidth demands for aggregated subscribers using monte carlo methods. *IEEE Communications Magazine*, 53(3):199–207.
- 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.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1):159 – 178.
- 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. *KSCE Journal of Civil Engineering*, 22(9):3233–3245.
- 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.
- Schmidt, R. D. O., Sadre, R., and Pras, A. (2013). Gaussian traffic revisited. *2013 IFIP Networking Conference, IFIP Networking 2013*, pages 1–9.
- Su, F., Dong, H., Jia, L., Qin, Y., and Tian, Z. (2016). Long-term forecasting oriented to urban expressway traffic situation. *Advances in Mechanical Engineering*, 8(1):1–16.
- 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.
- Yar, M. and Chatfield, C. (1990). Prediction intervals for the holt-winters forecasting procedure. *International Journal of Forecasting*, 6(1):127 – 137.
- Yoo, W. and Sim, A. (2015). Network bandwidth utilization forecast model on high bandwidth networks. In *2015 International Conference on Computing, Networking and Communications (ICNC)*, pages 494–498.