

APLICANDO A TÉCNICA DE SÉRIES TEMPORAIS EM GERENCIAMENTO PRÓ-ATIVO DE REDES DE COMPUTADORES

Francisco Watzko Neto

Edmundo Roberto Mauro Madeira

Instituto de Computação
IC - UNICAMP
CP 6176 - CEP 13083-970 - Campinas - SP
e-mail: {watzko, edmundo}@dcc.unicamp.br

Resumo

A crescente importância das redes de computadores impulsionou os estudos sobre gerenciamento pró-ativo e a tentativa de se gerenciar os sistemas de computação de uma forma integrada. Este artigo contribui no campo do gerenciamento pró-ativo apresentando os resultados obtidos com a aplicação da teoria de séries temporais no campo do gerenciamento de redes. Como principais resultados tem-se a estimativa do funcionamento futuro do segmento monitorado e a descrição dos passos a serem seguidos para a aplicação da técnica estudada no campo da monitoração de redes. O modelo de gerenciamento pró-ativo proposto e as funções de gerenciamento propostas servem de subsídios para a administração dos sistemas de computação de forma integrada.

Abstract

The importance of computer networks has been increased and has stimulated research on proactive network management and the effort to manage computational systems in an integrated form. This paper is on the field of proactive management and presents the results we have reached with the use of time series theory in network management. We have the possibility of estimate the future work of the segment, and the description of the steps to be taken to apply this theory in the field of network monitoring as main results. The proposed proactive management model and management functions contribute to manage computer systems in an integrated form.

1. INTRODUÇÃO

Este artigo apresenta os resultados dos estudos sobre o gerenciamento pró-ativo de redes de computadores. Suas principais contribuições são a proposição de um modelo de gerenciamento pró-ativo que se utiliza de informações sobre as aplicações que estão trafegando em segmentos monitorados, a aplicação da técnica de séries temporais na análise do perfil de funcionamento destes segmentos e a proposição de uma série de funções de gerenciamento que utilizam de tal técnica e do nível de informações disponibilizadas pelo modelo proposto.

A Seção 2 comenta os conceitos básicos sobre gerenciamento pró-ativo. Na Seção 3 é analisada a aplicação da técnica de séries temporais na interpretação do perfil de funcionamento do segmento monitorado (*baseline*), e na Seção 4 são apresentadas as funções pró-ativas de gerenciamento baseadas nas informações sobre as aplicações que estão sendo executadas e dos estudos de séries temporais. A Seção 5 discute brevemente alguns trabalhos relacionados. Finalizando, apresenta-se a conclusão deste trabalho.

2. GERENCIAMENTO PRÓ-ATIVO

No campo do gerenciamento de redes, dois foram os principais modelos básicos de gerenciamento de redes propostos: um baseado no protocolo CMIP (*Common Management Information Protocol*) proposto pela ISO (*International Organization for Standardization*) para o

ambiente OSI (*Open Systems Interconnection*) e outro baseado no protocolo SNMP (*Simple Network Management Protocol*), utilizado na Internet.

Inicialmente o gerenciamento de redes utilizava-se do paradigma reativo procurando resolver os problemas após estes terem ocorrido. A crescente importância das redes levou ao estudo de um novo paradigma, o gerenciamento pró-ativo, que tenta identificar situações que possam levar à degradação do sistema e resolvê-las antes que uma paralisação ocorra. A identificação dos estados críticos é obtida através da comparação das informações sobre a atual situação do funcionamento do sistema com o seu perfil de funcionamento, também conhecido como *baseline*. Após esta etapa, deve-se procurar identificar a razão de tal estado e as respectivas soluções para tal.

A *baseline*, ou perfil de funcionamento, é uma caracterização estatística do funcionamento normal do segmento monitorado, podendo conter informações sobre diferentes horários, dias, meses e até anos. Sua atualização periódica é recomendável devido aos constantes avanços existentes no ramo da computação.

Uma das formas de se criar este perfil de funcionamento é através da utilização da RMon MIB (*Management Information Base*) [WAL 91]. Esta foi criada com o objetivo de coletar informações da rede como mais um elemento do ambiente de gerenciamento, tal como as máquinas ou roteadores que a compõe. Destacam-se as seguintes facilidades oferecidas: ela permite a identificação de situações pré-determinadas e a possibilidade de reportá-las à estação de gerenciamento quando da sua ocorrência, além de poder operar sem o contínuo contato com esta estação, distribuindo o gerenciamento da rede antes concentrado na estação de gerenciamento. Dentre os vários grupos definidos nesta especificação foram utilizados:

1. *statistics*: que informa estatísticas sobre a taxa de utilização e de erros do segmento monitorado;
2. *history*: mantém amostras estatísticas periódicas do grupo *statistics*;
3. *host*: possui vários contadores para diversos tipos de tráfego de/para máquinas que fazem parte do segmento monitorado;
4. *filter*: permite ao monitor observar e capturar pacotes que passem por um filtro configurado;
5. *packet capture*: gerencia a forma de se enviar as informações coletadas nos filtros.

Além destes grupos, são definidos ainda: *hostTopN*, *matrix* e *tokenRing*.

A recente constatação de que as facilidades oferecidas pelas redes de computadores são disponibilizadas através de diversas aplicações, gerou a necessidade de administração dos sistemas de rede com todos os seus recursos. Buscando suprir a carência de informações sobre as aplicações, foi proposta a RMon 2 MIB [WAL 95a], que é na verdade uma complementação à RMon MIB anterior, permitindo efetuar análises acima da camada de rede, através da leitura do cabeçalho do protocolo do nível de rede, visualizando-se os protocolos de alto nível que são executados acima do protocolo desta camada, através da leitura do cabeçalho do protocolo TCP ou UDP e até mesmo o cabeçalho do protocolo do nível de aplicação [BIE 97]. Esta nova MIB permite a geração de estatísticas sobre estes níveis como, por exemplo, a quantidade de tráfego destinado ou originado de uma determinada máquina de uma aplicação. A implementação da coleta de dados sobre os protocolos TCP e UDP neste trabalho foi feita através da utilização dos grupos *filter* e *packet capture* da RMon 1 implementada pelo agente remoto Beholder, desenvolvido na Universidade de Delft, the Netherlands, disponível através de ftp anônimo para o endereço eletrônico <ftp.dnpap.et.tudelft.nl>. No grupo *filter*, foram definidos dois filtros para cada aplicação monitorada, um com endereço destino especificando a porta *well-known* da aplicação e outro definindo a mesma porta para o endereço origem. Estes dois filtros são associados a um canal responsável pelo gerenciamento do fluxo de pacotes que são aceitos pelos filtros. Após os pacotes estarem no canal, eles serão armazenados em um *buffer* associado a este definido no grupo *packet capture*. Neste ponto, um programa desenvolvido no projeto é utilizado para ler as informações destes pacotes e armazená-las criando uma *baseline* da aplicação. O segmento de

rede, ou parte da rede, monitorado neste trabalho apresenta uma grande utilização e faz parte da rede do Instituto de Computação que é composta por 14 segmentos.

2.1. MODELO DE GERENCIAMENTO PRÓ-ATIVO DE APLICAÇÕES

Em [DEF 96], é proposto um modelo para o gerenciamento pró-ativo independente de aplicações que define uma estrutura genérica onde os conceitos de *baseline*, verificação e ações corretivas, interagem com uma plataforma de gerência responsável pela tomada de ações corretivas. Visando complementar este modelo para lidar com informações sobre as aplicações, foram incluídos no modelo proposto neste trabalho as facilidades de disponibilização e tratamento das informações em nível de aplicação, a técnica de séries temporais na análise da *baseline*, e novas funções para o módulo **Sistema de Correção Automática** que passou a fazer parte do **Módulo de Gerenciamento Pró-ativo** com o objetivo de distribuir o gerenciamento da rede [WAT 97]. A Figura 1 mostra o modelo proposto.

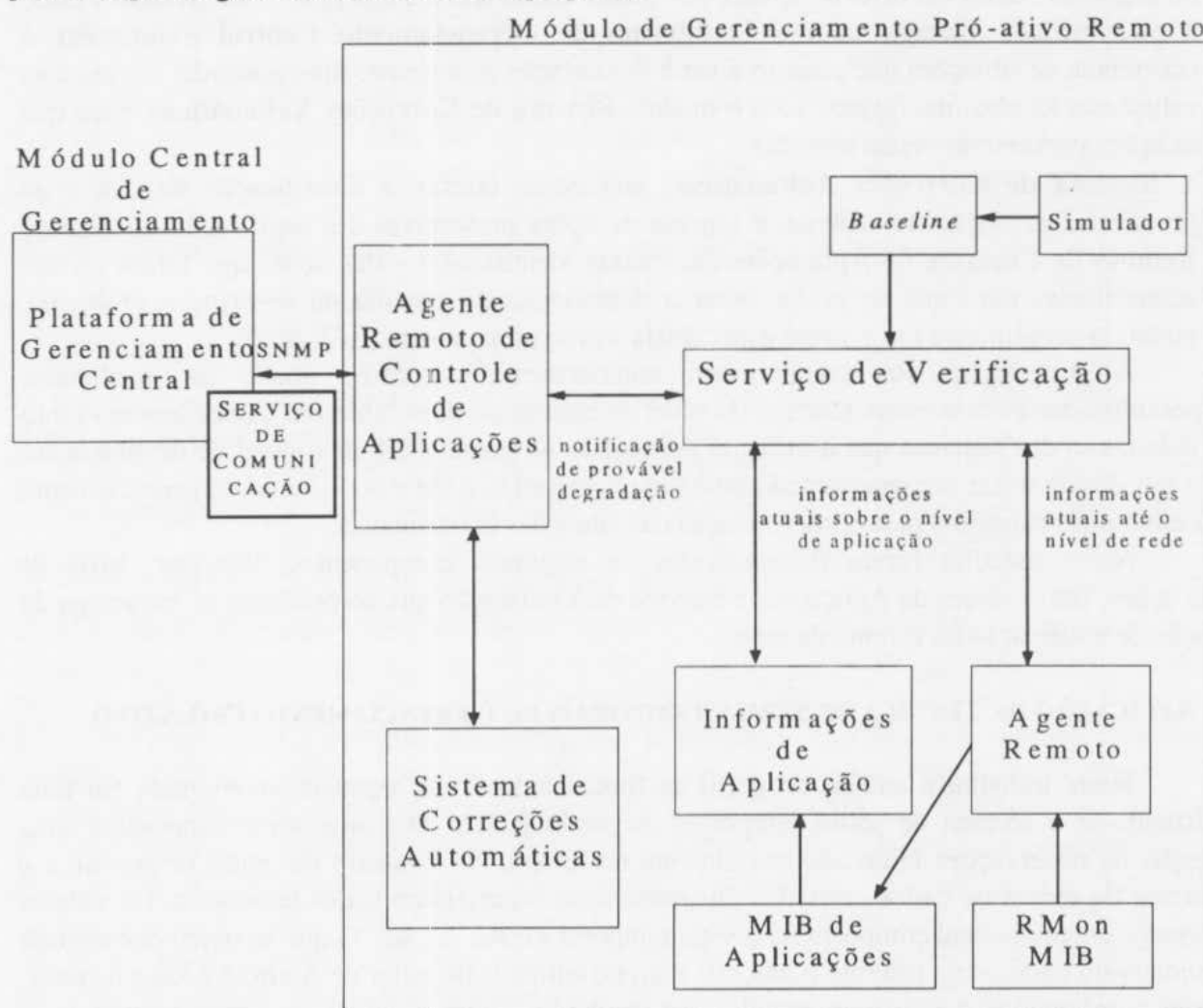


Figura 1 - Modelo de Gerenciamento Pró-ativo Proposto

- **Módulo Central de Gerenciamento:** este módulo é composto pela Plataforma Geral de Gerenciamento e tem como tarefa o gerenciamento centralizado da rede.
- **Módulo de Gerenciamento Pró-ativo Remoto:** este módulo tem como objetivo a implementação do gerenciamento pró-ativo propriamente dito. Seus componentes são:

* **Baseline:** este componente conterá informações sobre o funcionamento normal da rede. Atualmente ela é gerada através da monitoração da rede, mas também existe a possibilidade de definição de um modelo da rede, para que se possa estimar, através do emprego de simulações, uma *baseline* do comportamento mais próximo do ideal possível.

* **Agente Remoto:** tem como tarefa a coleta de informações sobre o segmento monitorado até o nível de rede.

* **Informações de Aplicação:** tem como função a disponibilização de informações sobre as aplicações que estão sendo monitoradas ao módulo **Serviço de Verificação**.

* **Serviço de Verificação:** tem como funções a identificação de situações que possam levar à degradação da rede e dos sistemas, baseando-se tanto em informações de mais baixo nível quanto sobre as aplicações que estão sendo monitoradas, e reportá-las ao módulo **Agente Remoto de Controle de Aplicações**.

* **Agente Remoto de Controle de Aplicações:** é responsável pelas tarefas de controle local do segmento, tanto em nível de aplicações quanto em níveis mais baixos de informação. Cabe a este módulo interagir com a **Plataforma de Gerenciamento Central** e informar a ocorrência de situações que possam levar à degradação do sistema, suas possíveis causas e as soluções adotadas, interagindo com o módulo **Sistema de Correções Automáticas**, para que as ações preventivas sejam tomadas.

* **Sistema de Correções Automáticas:** tem como tarefas a identificação da causa da provável degradação do sistema, a tomada de ações preventivas e a notificação do **Agente Remoto de Controle de Aplicações** das causas identificadas e das ações que foram ou são aconselhadas para que se possa evitar a degradação do sistema ou resolver o problema, podendo ser utilizado um sistema especialista, como proposto em [SAE 96a].

A proposição dos componentes anteriormente descritos, aliado às facilidades disponibilizadas para o gerenciamento do nível de aplicação, possibilita um gerenciamento tanto da rede como dos sistemas que a utilizam permitindo ao gerente definir o nível de detalhamento com que deseja tratar um determinado problema e provendo a ele e ao sistema de gerenciamento pró-ativo mais subsídios para a interpretação das situações identificadas.

Neste trabalho foram desenvolvidos os seguintes componentes: *Baseline*, MIB de aplicações, Informações de Aplicações e Serviço de Verificação que no protótipo se encarrega da função de notificação ao gerente da rede.

3. APLICAÇÃO DA TÉCNICA DE SÉRIES TEMPORAIS NO GERENCIAMENTO PRÓ-ATIVO

Neste trabalho a análise do perfil de funcionamento do segmento monitorado foi feita utilizando-se a técnica de séries temporais. Segundo [CHA 84], uma série temporal é uma coleção de observações feitas sequencialmente no tempo. No contexto de gerência pró-ativa a *baseline* da rede e os dados coletados dinamicamente representam séries temporais. Os valores coletados em tempo real compõem uma série temporal x_1, x_2, \dots, x_N . O que se objetiva é estimar um intervalo onde esteja inserido o valor de x_{N+Q} no tempo N do valor de X em Q pontos à frente. Serão apresentados a seguir os estudos desenvolvidos sobre a teoria de séries temporais, a modelagem realizada e a verificação das previsões obtidas, adaptados à monitoração de redes de computadores.

Segundo [BOW 87], a escolha de uma técnica de previsão deve levar em consideração os seguintes fatores:

1. **forma de previsão.** Basicamente esta pode ser feita de duas formas: previsão de um ponto ou de um intervalo;
2. **intervalo de tempo.** Identifica granularidade de tempo que se deseja trabalhar já que as previsões são feitas para pontos no tempo;

3. **padrão dos dados.** Identificação de padrões, tendências e sazonalidades existentes nos dados;
4. **custo de previsão.** Dentre os fatores que devem ser levados em consideração, destacam-se: custo de desenvolvimento do modelo, o custo do armazenamento das informações necessárias à previsão e o custo do cálculo da previsão;
5. **precisão desejada.** Probabilidade do valor lido estar dentro da faixa prevista;
6. **disponibilidade de dados;**
7. **facilidade de operação e compreensão.**

Foram estabelecidas para este trabalho:

1. Pretende-se prever um intervalo de confiança onde o próximo valor a ser observado estará contido;
2. Um intervalo de tempo de 30 minutos tanto para a coleta dos dados quanto para a previsão;
3. A análise dos dados e conhecimentos prévios levaram a identificação de uma sazonalidade na série temporal analisada entre as amostras;
4. Por ser uma nova forma de aplicação desta técnica no campo do gerenciamento pró-ativo, o custo de desenvolvimento do modelo foi um pouco alto devido principalmente aos estudos realizados no campo estatístico. O custo de armazenamento das informações necessárias à previsão e o custo do cálculo da previsão podem ser desprezados;
5. Trabalha-se com um intervalo de confiança de 95%, ou seja, 95% dos intervalos dos valores futuros previstos estarão corretos e que 10% destes intervalos previstos estarão errados;
6. A disponibilidade dos dados foi atendida com a utilização do agente Beholder e de programas desenvolvidos para a geração de informações sobre o nível de aplicação.

A metodologia de previsão, que está sendo aplicada foi desenvolvida por G. E. P. Box e G. M. Jenkins [BOX 70], consiste de quatro passos:

1. **Tentativa de Identificação.** Tentativa de identificação de um modelo através do emprego da função de autocorrelação da amostra (SACF - *Sample Autocorrelation Function*) e da Função de Autocorrelação Parcial da Amostra (SPACF - *Sample Partial Autocorrelation Function*);
2. **Estimativa.** Uma vez que o modelo tenha sido identificado, passa-se à estimativa dos parâmetros.
3. **Checagem de Diagnóstico.** É feita uma verificação das estimativas realizadas com o modelo especificado. Se o modelo se mostrar inadequado ele deve ser modificado e aprimorado;
4. **Previsão.** Quando o modelo já foi modificado e aprimorado, passa-se à aplicá-lo para a previsão de valores futuros.

Esta metodologia é largamente aplicada no campo da economia. Pretendeu-se neste trabalho analisar qual o seu comportamento quando aplicada no campo de Gerenciamento de Redes. Ela foi escolhida por poder ser utilizada tanto em dados discretos quanto contínuos desde que estes sejam medidos em intervalos iguais e discretos. Pode ser aplicada também tanto sobre dados sazonais quanto não sazonais.

Serão apresentados a seguir, os estudos feitos sobre séries temporais: Modelo ARIMA, função de autocorrelação, processos de médias móveis e mistos, processos integrados, processos sazonais e não sazonais combinados.

3.1. MODELO ARIMA (AUTOREGRESSIVE INTEGRATED MOVING AVERAGE)

Uma equação deste modelo define um valor em uma série temporal como sendo linearmente relacionado ao seu próprio valor passado, dando-se maior importância aos valores mais recentes, ou seja, tenta-se verificar a influência dos valores anteriormente observados no valor atual de uma determinada variável monitorada. Qualquer previsão de um modelo deste tipo é uma média ponderada dos valores anteriormente observados.

Um modelo ARIMA é baseado nos dados amostrados, no caso, as diversas variáveis que compõem a *baseline*. Se o modelo é uma boa aproximação do processo real, este tende a imitar o comportamento do processo real, desta forma as previsões deste fornecerão informações úteis sobre os valores da série, permitindo a identificação de situações que levem à degradação do sistema antes que estas venham a ocorrer.

3.1.1. MÉDIA DE VARIÂNCIA ESTACIONÁRIA

A análise padrão ARIMA baseia-se na simplificação de que o processo que gerou a série temporal a ser analisada é estacionário, ou seja, é um processo cuja média, variância e a função de autocorrelação são constantes no tempo. Caso este fato não ocorra, pode-se modificá-la para torná-la estacionária. Como as modificações são conhecidas, estas podem ser revertidas para que as previsões possam ser comparadas aos valores originais da série.

A verificação da estacionariedade da série costuma ser feita através da observação tanto do gráfico da série propriamente dita quanto dos gráficos da função SACF e SPACF da mesma.

A estabilização da variância é feita através da transformação dos dados, sendo duas as principais transformações a serem aplicadas:

1. Se o desvio padrão de uma série é proporcional ao seu nível, a utilização do logaritmo natural levará a uma série com a característica desejada;
2. Se a variância da série original é proporcional ao seu nível, a utilização da raiz quadrada levará a uma série com a característica desejada.

Para que a média de uma série seja considerada estacionária, ela deve convergir rapidamente para uma média constante. A estabilização da média pode ser feita das seguintes formas:

1. **diferenciação não sazonal:** cria-se uma nova série da seguinte forma: $w_t = z_t - z_{t-1}$, chamada de primeira diferença não sazonal. Se w_t ainda não apresentar uma média geral constante, calcula-se a primeira diferença desta série, chamada de segunda diferença não sazonal, $w_t^* = w_t - w_{t-1}$ ou $w_t = (z_t - z_{t-1}) - (z_{t-1} - z_{t-2})$. Se d é o grau da diferença não sazonal, então teremos $d = n$ para a n -ésima diferença não sazonal;
2. **diferenciação sazonal:** Através da observação do gráfico da série ou de conhecimentos prévios pode-se constatar a existência de variações periódicas ou sazonais: o nível de uma observação (w_t) é similar ou relacionado ao nível de outras observações que estão separadas por $s = 48$ períodos de tempo (w_{t-96}, w_{t+96}), $2s$ períodos de tempo (w_{t-96}, w_{t+96}), e assim por diante. Uma série pode ser diferenciada apenas não sazonalmente, apenas sazonalmente ou ambas. Seja D o grau de diferença sazonal, se $d=0$ e $D=1$, a série será da forma: $w_t = z_t - z_{t-s}$.

3.2. AUTOCORRELAÇÃO

A Autocorrelação procura verificar de que forma o valor corrente da série (z_t) está relacionado aos seus valores antecessores ($z_{t-1}, z_{t-2}, z_{t-3}, \dots$) ou com seus valores futuros ($z_{t+1}, z_{t+2}, z_{t+3}, \dots$).

3.2.1. FUNÇÃO DE AUTOCORRELAÇÃO DA AMOSTRA (SACF)

Dada a série temporal: z_1, z_2, \dots, z_K . A função de autocorrelação da amostra no intervalo k (r_k) é:

$$r_k = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2} \quad \text{onde} \quad \bar{z} = \frac{\sum_{t=1}^n z_t}{n}$$

Esta quantidade mede o relacionamento linear entre as séries temporais observadas separadas por um intervalo de k unidades de tempo. A SACF também é utilizada na verificação da estacionariedade de uma determinada série temporal, segundo [BOW 87]:

1. Se o SACF dos valores de uma série temporal desaparecem muito rapidamente ou decrescem muito rapidamente, então estes valores devem ser considerados estacionários;
2. Se o SACF dos valores de uma série temporal decrescem extremamente devagar, então estes valores devem ser considerados não estacionários.

Para a monitoração de redes, este parâmetro indica a correlação existente entre o valor da variável de utilização no tempo t com o observado no tempo $t-k$, ou seja, se o valor observado no tempo t é maior do que a média dos valores observados, e o valor observado no tempo $t+3$ também é maior que a média, e todas as possibilidades desta comparação.

3.2.2. FUNÇÃO DE AUTOCORRELAÇÃO PARCIAL DA AMOSTRA (SPACF)

A função de autocorrelação parcial da amostra de intervalo k é:

$$r_{kk} = \begin{cases} r_1, k = 1 & k = 1 \\ \frac{r_k \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} & \forall k \geq 2 \end{cases}$$

A SPACF pode ser interpretada como sendo a SACF das séries temporais observadas separadas por um intervalo de k unidades de tempo com os efeitos da intervenção das observações eliminados.

Na monitoração de redes, é o conjunto de valores composto pelos pesos atribuídos ao valor no tempo $t+K$, com o intervalo $k=1, 2, 3, \dots, K$.

3.2.3. PROCESSOS ARIMA ESTACIONÁRIOS

Um modelo ARIMA é baseado nos dados coletados e seu equivalente teórico é um processo ARIMA que possui o seu próprio ACF e PACF associados. Sua identificação segue os seguintes passos:

1. Construem-se o ACF e PACF da série temporal formada pelos dados disponíveis;
2. Compara-se o ACF e o PACF da série temporal da amostra com os ACFs e PACFs teóricos até que se encontre uma semelhança aceitável entre ambos;
3. Seleciona-se o processo que possui os ACFs e PACFs semelhantes como um modelo experimental ARIMA.

3.2.3.1. PROCESSO AUTOREGRESSIVO

Dois processos ARIMA comuns e bastante simples são apresentados na Fórmula 1 .

$$z_t = C + \phi_1 z_{t-1} + a_t \qquad z_t = C + \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t \qquad C = \mu_z \left(1 - \sum_{i=1}^p \phi_i \right)$$

Fórmula 1 - Modelos de AR(1) e AR(2) e C

Nestes processos, os valores de z passados com os seus coeficientes são chamados termos autoregressivos (AR), a ordem (p) é o comprimento do maior intervalo entre os termos e o termo C é uma constante que relaciona a média (μ_z) da série z_t e os coeficientes AR. Os ACFs e PACFs teóricos associados aos processos AR(1) e AR(2) são apresentados na Figura 2 e seu comportamento, resumido da seguinte forma:

1. A ACF teórica decai exponencialmente, com o padrão de uma onda senoidal ou ambos.
2. A PACF teórica possui saltos até o intervalo p e em seguida torna-se zero.

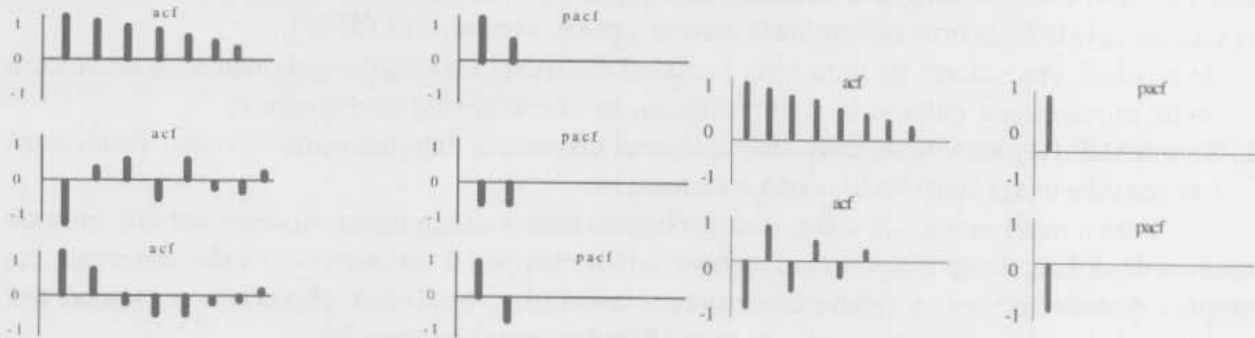


Figura 2 - ACFs e PACFs teóricos de AR(1) e AR(2)

Para a monitoração de redes este modelo representa que o valor observado atualmente é decorrente de um valor observado em um tempo $t-k$ com um determinado peso, acrescido de um valor C e de um erro a . A constante C está relacionada com a média (μ_z) dos valores observados e com os coeficientes do modelo.

3.2.4. MÉDIAS MÓVEIS E PROCESSOS MISTOS

A Fórmula 2 apresenta os processos MA(1), MA(2) e ARMA(1,1). Nestas, o distúrbio aleatório passado (a_t) juntamente com os seus coeficientes são chamados termos de médias móveis (MA - *Moving Average*) que, por convenção, são escritos com sinais negativos. A ordem (q) deste processo é o tamanho do maior tamanho de intervalo nos seus termos.

$$Z_t = C - \theta_1 a_{t-1} + a_t \quad (1) \quad z_t = C - \theta_1 a_{t-1} - \theta_2 a_{t-2} + a_t \quad (2) \quad z_t = C - \phi_1 z_{t-1} - \theta_1 a_{t-1} + a_t \quad (3)$$

Fórmula 2 - Exemplos de Processos (1) MA(1), (2) MA(2) e (3) ARMA(1,1)

Um processo que combine os processos AR e MA é conhecido como ARMA, chamado média móvel autoregressiva. As ACFs e PACFs teóricas associados aos processos MA(1), MA(2) e ARMA(1,1) são apresentados na Figura 3. Seu comportamento pode ser resumido da seguinte forma:

1. A ACF teórica de um processo MA possui saltos até o intervalo q e em seguida torna-se zero;
2. A PACF teórica de um processo MA decresce;
3. Ambos ACF e PACF teóricos de um processo misto decrescem.

Na monitoração de redes, os termos de um processo MA representam uma parte do valor da variável observada no tempo t , composta por um ruído aleatório que leva a uma distorção da função desta variável.



ACFs e PACFs teóricos do Processo MA(1) e MA(2)



Figura 3 - ACF e PACF teóricos do Processo ARIMA(1,1,1)

3.2.5. PROCESSOS INTEGRADOS

Todas as fórmulas apresentadas até aqui são processos aplicáveis a séries temporais que possuem média estacionária. Para se aplicar tais conceitos a processos que não possuem tal característica, a única diferença existente é que o modelo será escrito para a série diferenciada gerada a partir da série original. Por exemplo, para um processo cujo $d = 1$ e $D = 0$ e cuja série diferenciada w_t seja descrita por um processo ARMA(1,1), isto gerará um processo ARIMA(p,d,q)=ARIMA(1,1,1) para a série original z_t .

3.3. MODELOS SAZONAIS

Os modelos ARIMA podem também ser aplicados na representação de padrões periódicos. As ordens de um processo sazonal ARIMA puro são expressas da forma ARIMA(P, D, Q) $_s$, onde P_s é o tamanho do maior intervalo dos termos AR sazonais, Q_s é o tamanho do maior intervalo dos termos MA sazonais, D_s é o número de diferenciações sazonais realizadas na série original e s é a sazonalidade da série. Um dos modelos sazonais mais comuns existentes é o ARIMA(0,1,1) $_s$ que é apresentado por: $z_t = z_{t-s} + C - \theta_s a_{t-s} + a_t$

Apesar dos padrões sazonais e não sazonais costumarem ocorrer juntos na prática, é interessante procurar efetuar uma análise separada de ambos. As ACF e PACF de um processo teórico puramente sazonal de ordens P e Q são idênticos aos processos não sazonais de mesma ordem com uma exceção: para um processo puramente sazonal os padrões ocorrem nos intervalos $s, 2s, 3s, \dots$. Logo, as ACF e PACF de processo puramente sazonais podem ser caracterizadas da seguinte forma:

1. A ACF decresce vagarosamente nos múltiplos do intervalo s , caso a diferença sazonal seja necessária;
2. Em um processo AR puramente sazonal estacionário de ordem P , a ACF decresce nos múltiplos do intervalo s enquanto a PACF possui saltos nos múltiplos do intervalo s e em seguida torna-se zero;
3. Em um processo MA puramente sazonal estacionário de ordem Q , a ACF apresenta saltos nos múltiplos do intervalo s e então torna-se zero, sendo que o último salto não nulo ocorre no intervalo Qs . A sua PACF decresce nos múltiplos do intervalo s ;
4. Em um processo misto puramente sazonal de ordens P e Q , ambas as ACF e PACF decaem nos múltiplos do intervalo s .

Para a monitoração de redes, a representação da sazonalidade de uma série temporal implica na representação de que o valor da variável observado em um determinada hora t está relacionado com o valor observado nesta mesma hora no dia anterior, por exemplo. Pode-se pensar, por exemplo, que a observação feita a uma hora da tarde de uma segunda-feira está relacionada com a observação feita à uma hora da tarde da segunda-feira anterior e assim sucessivamente. Vários podem ser os níveis de sazonalidades a serem representadas no modelo: os diferentes dias de uma semana, os diferentes meses do ano e até anos diferentes.

3.4. COMO APLICAR A TÉCNICA DE SÉRIES TEMPORAIS NO GERENCIAMENTO DE REDES

Este trabalho propõe que os seguintes passos sejam seguidos para a aplicação de séries temporais na monitoração de redes:

1. Geração da *baseline*;

Neste passo é feita a coleta de dados, devem ser definidos com que intervalo de tempo serão feitas as coletas de dados e o programa necessário para esta atividade.

2. Análise da *baseline*,

Após a *baseline* ter sido coletada, deve-se escolher os parâmetros que pretende-se analisar segundo a técnica de séries temporais. Para esta escolha, deve ser levada em consideração a relevância da mesma para o gerenciamento de redes e a aleatoriedade da variável pois quanto maior for a aleatoriedade, menor será a precisão do modelo, podendo chegar ao ponto de ser impossível modelar a variável utilizando-se séries temporais. A geração de gráficos dos dados coletados é peça fundamental nesta fase. Os gráficos são importantes tanto para se verificar a aleatoriedade de uma variável (série temporal) quanto para a verificação de sazonalidades, tendências e períodos. A apresentação de todos os dados coletados em apenas um gráfico pode ajudar na identificação de tendências de crescimento ou de queda da variável ao longo do tempo. A geração de um gráfico para cada dia da semana ou para cada mês ajuda na identificação de sazonalidades e padrões de comportamento da variável como, por exemplo, os dias da semana ou semanas do mês que registram os valores mais altos. Já o gráfico que apresenta vários dias da semana juntos serve para a identificação de periodicidades com uma granularidade menor, como os horários onde aparecem os maiores valores ou o padrão de comportamento da variável durante o dia. Estes gráficos também são úteis na identificação da relação entre os valores observados, por exemplo, para verificar se a existência de um valor elevado em um tempo t influencia ou não nos valores futuros.

3. Efetua-se as diferenciações sazonal e não sazonal da série, quando necessário, e a geração dos gráficos da função ACF e PACF da série;

A geração dos gráficos da função ACF e PACF normalmente é feita com a utilização de um programa de estatística que possui também as facilidades para se fazer a estabilização da média e da variância da série.

4. Identificação de modelo-teste;

Neste passo é utilizado o conhecimento sobre o comportamento da série adquirido no passo 2. Os processos AR representam a relação entre o valor da variável em um determinado tempo t com os seu valores anteriores, ou seja, ele procura representar a tendência de um determinado valor permanecer acima ou abaixo da média em relação aos seus valores anteriores. Já os processos MA representam a correlação existente entre o distúrbio aleatório componente do valor observado no tempo t e seus antecessores. Por exemplo, os processos de primeira ordem (AR(1), MA(1), ARMA(1,1)) indicam que o valor da variável no tempo t sofre influência apenas do seu valor anterior, no gerenciamento de redes equivaleria a dizer que o nível de utilização em uma determinada hora está relacionado apenas com o nível de utilização da hora anterior, ou ainda que o nível de utilização observado agora não influenciará em nada o valor observado daqui a duas observações. As sazonalidades do processo são representadas também através dos processos AR, MA e da integração de ambos. Neste caso, eles estarão representando a correlação existente entre o valor da variável observado agora e o observado a s períodos de tempo. Neste caso, s pode representar a correlação entre os horários de um dia ou entre os dias de uma semana. Para o gerenciamento de redes, seria a representação no modelo da relação existente entre os horários de pico ou entre os dias de maior utilização do segmento, por exemplo. Um processo de ordem 1 representa a correlação existente entre um determinado horário de um dia e o respectivo horário no dia anterior mas caracterizando que o mesmo horário a dois dias atrás não teria

influencia sobre o atual. As funções ACF e PACF irão auxiliar na identificação do relacionamento entre os valores da série temporal que está sendo analisada. Procura-se comparar os gráficos de ACF e PACF da série observada com os gráficos dos modelos teóricos (AR, MA, ARIMA), buscando identificar os que mais se assemelham. Esta comparação é bastante empírica e muitas vezes pode levar a vários modelos que devem ser testados posteriormente;

5. Estimativa dos valores da função que representa o modelo-teste e a verificação da adequação do modelo;

Este passo normalmente é feito com o auxílio de um programa estatístico. Com este programa são calculados os valores da equação que representa o modelo especificado e são realizados os testes de adequação do modelo ao processo. Caso o modelo seja rejeitado, indica que existe alguma correlação entre os valores que ainda não foi representada, ou seja, a correlação entre os horários de um dia ou entre os dias da semana não está corretamente adequada. Neste caso o modelo deve ser melhorado através da tentativa de representação da correlação existente entre os diversos valores que compõem o resíduo do modelo de teste.

6. Estimativa de valores futuros.

Neste passo é feita a estimativa dos valores futuros da série. Sua precisão está relacionada com o nível de adequação do modelo ao processo, ou seja, com o nível de detalhamento que se procurou representar no modelo e com o nível de aleatoriedade da variável.

3.5. ANÁLISE DA UTILIZAÇÃO DO SEGMENTO MONITORADO ATRAVÉS DA TÉCNICA DE SÉRIES TEMPORAIS

A série temporal que será utilizada como base de dados para este estudo terá como dados o nível de utilização de um segmento da rede do Instituto de Computação da Unicamp. As coletas foram feitas em intervalos de 30 minutos durante 39 dias.

3.5.1. IDENTIFICAÇÃO

A Figura 4 apresenta a plotagem da série lida, o ACF e o PACF desta amostra com $d=D=0$. Observa-se que o seu ACF decai vagarosamente para zero, sugerindo que a média geral da série não é estacionária. Isto parece consistente com o comportamento do gráfico da plotagem da série. Faz-se então a primeira diferença não sazonal ($d=1$) desta série, apresentado na Figura 5 juntamente com seu ACF e PACF. Analisando-se o gráfico de plotagem, observa-se que este apresenta uma média global constante, fato este confirmado pelo gráfico ACF que decai para zero bastante acentuadamente, além de uma acentuada variação sazonal. Estes gráficos foram feitos utilizando-se dos recursos da linguagem S+, largamente utilizada no campo da estatística por oferecer uma variada gama de recursos necessários aos estudos estatísticos, maiores detalhes sobre esta linguagem podem ser encontrados em [EVE 94, STE 94].

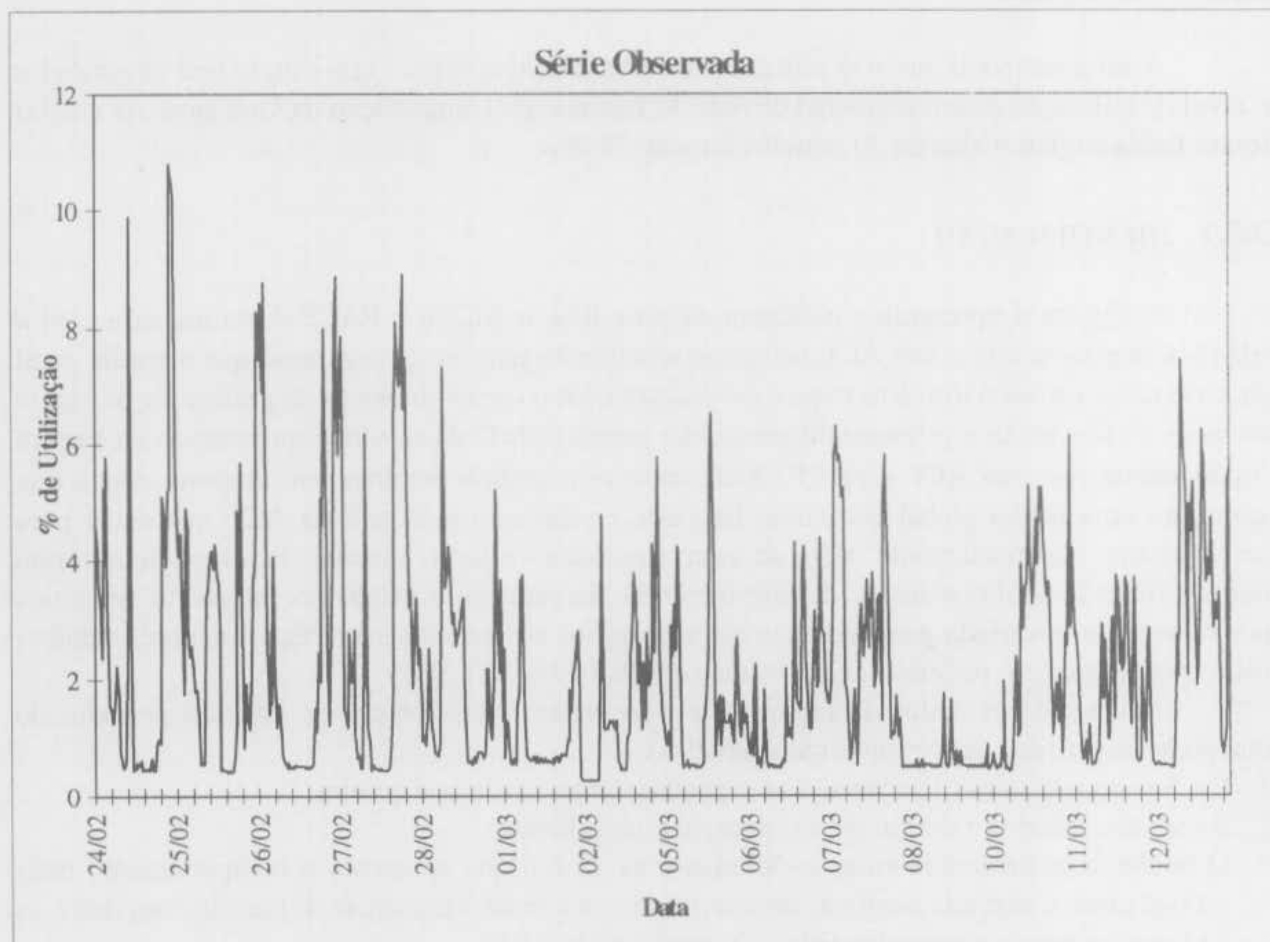
Analisando os dados da Figura 4 levando-se em consideração os diversos períodos do dia, pode-se verificar as seguintes características:

1. Os dias com maior utilização são as segundas, terças, quartas e quintas;
2. As sextas, sábados e domingos a utilização é mais baixa;
3. O horário compreendido entre as 8 horas e as 18 é o que apresenta o comportamento mais estável entre a segunda e quinta, caracterizando-se por uma tendência de crescimento das 8 às 13 horas seguindo de uma tendência de queda até às 23:30;
4. O intervalo de tempo compreendido entre a 0 hora e as 8 horas apresenta um nível de aleatoriedade bem maior que os outros intervalos.

Baseando-se nestes pontos, pode-se constatar que a série que está sendo analisada apresenta alguma sazonalidade levando-se em consideração os diferentes horários do dia e uma sazonalidade um pouco mais fraca em relação aos dias da semana. Estas informações nos fornecem idéia que o modelo poderá apresentar uma parte sazonal e uma não sazonal.

O gráfico da série diferenciada ($d=1$ e $D=0$) nos indica a existência de sazonalidades, vindo a confirmar a análise dos dados feita anteriormente. Aplica-se então a diferença sazonal ($D = 1$). Esta nova série é apresentada na Figura 6 juntamente com a função ACF e PACF calculadas para a nova série. Esta apresenta uma sazonalidade bem menor que a série anterior. Conclui-se que não é mais necessário efetuar qualquer outra diferenciação sazonal. Comparando-se os gráficos da função ACF e PACF da série em análise com os apresentados na seção 2, pode-se estimar que a parte não sazonal se assemelha aos de um modelo ARMA(1,1), se tomarmos como interpretação dos gráficos que ambos decaem.

Procura-se então, analisar os gráficos tentando dar mais atenção aos valores sazonais. Mais uma vez, pode-se estimar que eles podem ser representados por um processo AR(2), caracterizado pelo salto no intervalo 48 e 96 do gráfico da função PACF e pela forma de decrescer do gráfico ACF, que sazonalmente não aparenta um comportamento muito forte pois a observação no intervalo 48 de -0,03 pode ser considerada irrelevante se comparada com o valor no intervalo 0 que é igual a 1. A sazonalidade com valor 48 é explicada pela periodicidade com que é coletada a amostra: de 30 em 30 minutos ou 48 amostras por dia. O modelo que representa o comportamento da série observada é ARIMA(1,1,1)(2,1,0)₄₈.



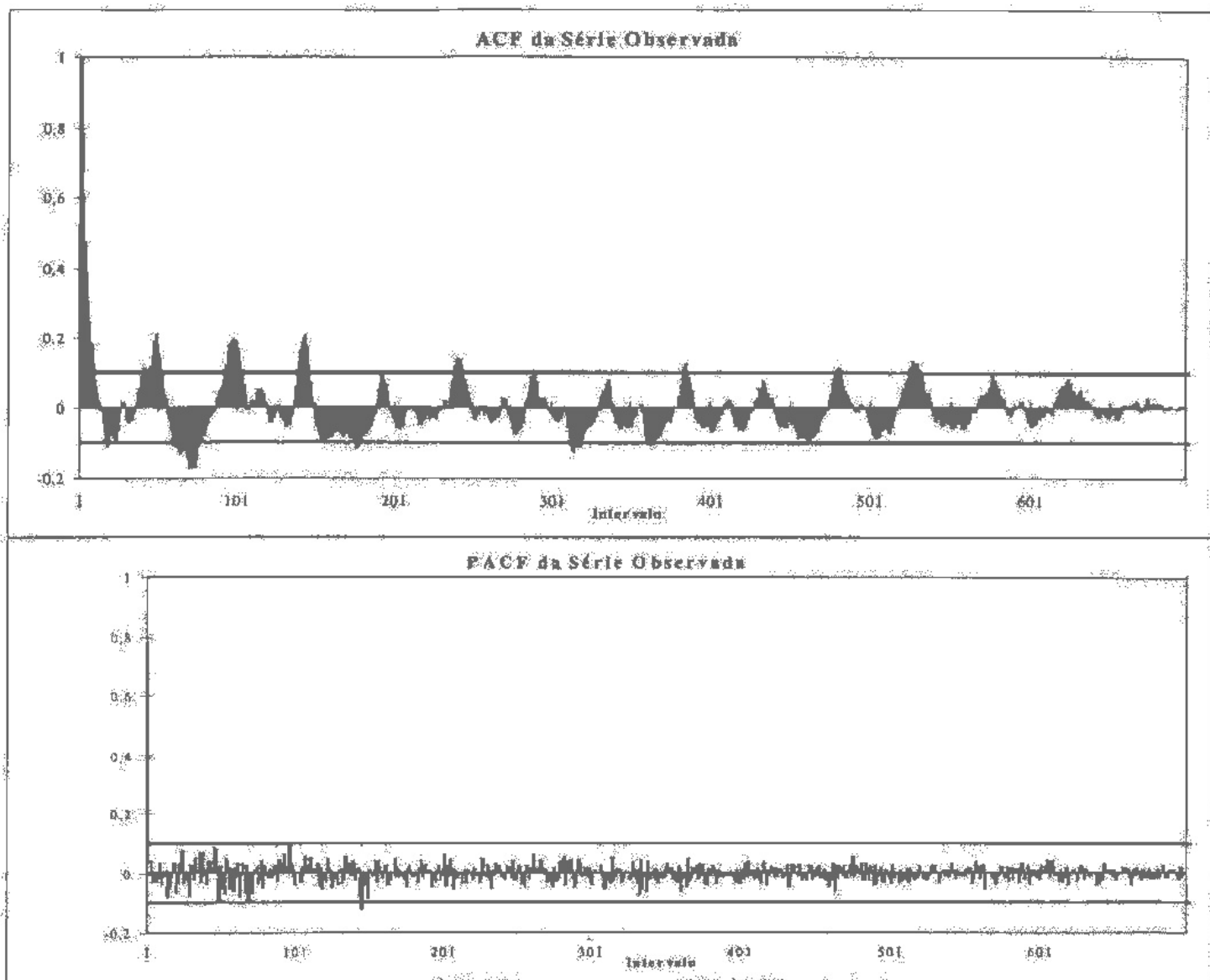
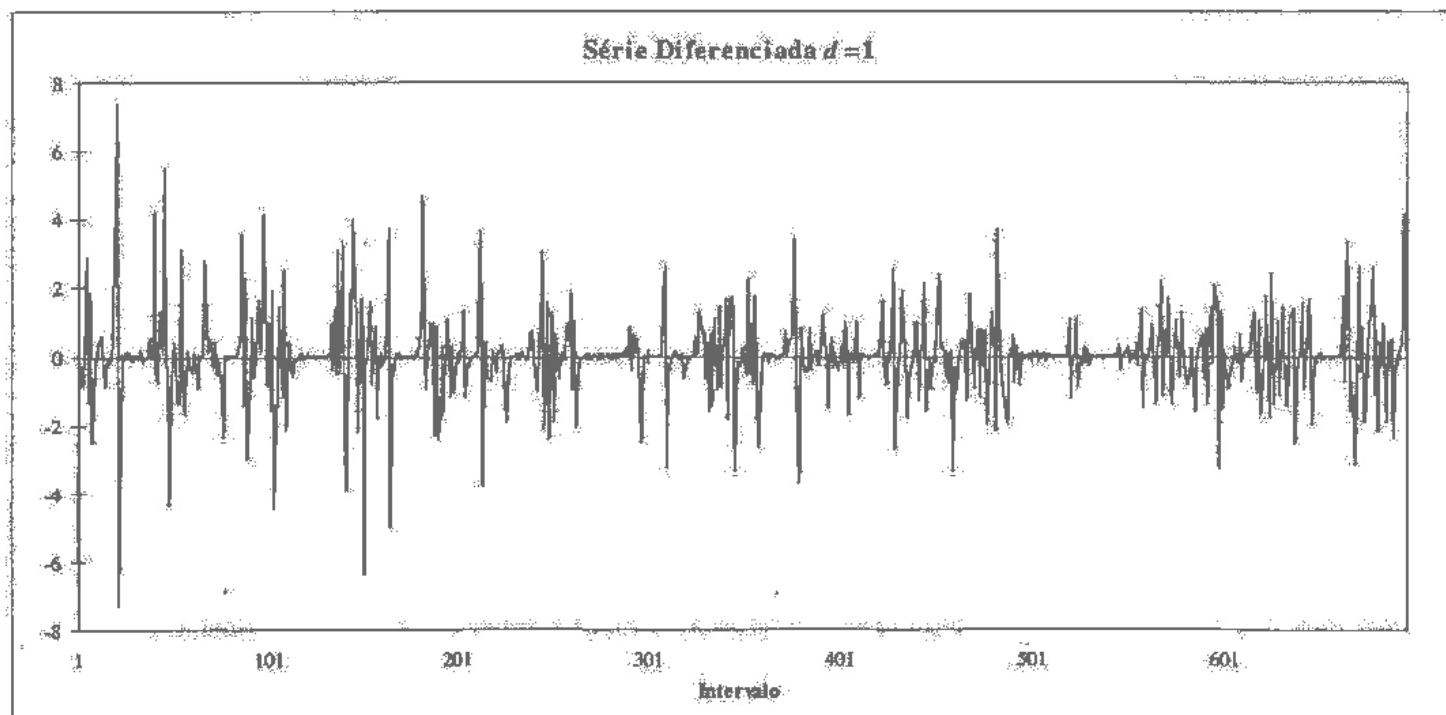


Figura 4- Série Lida, seu ACF e PACF



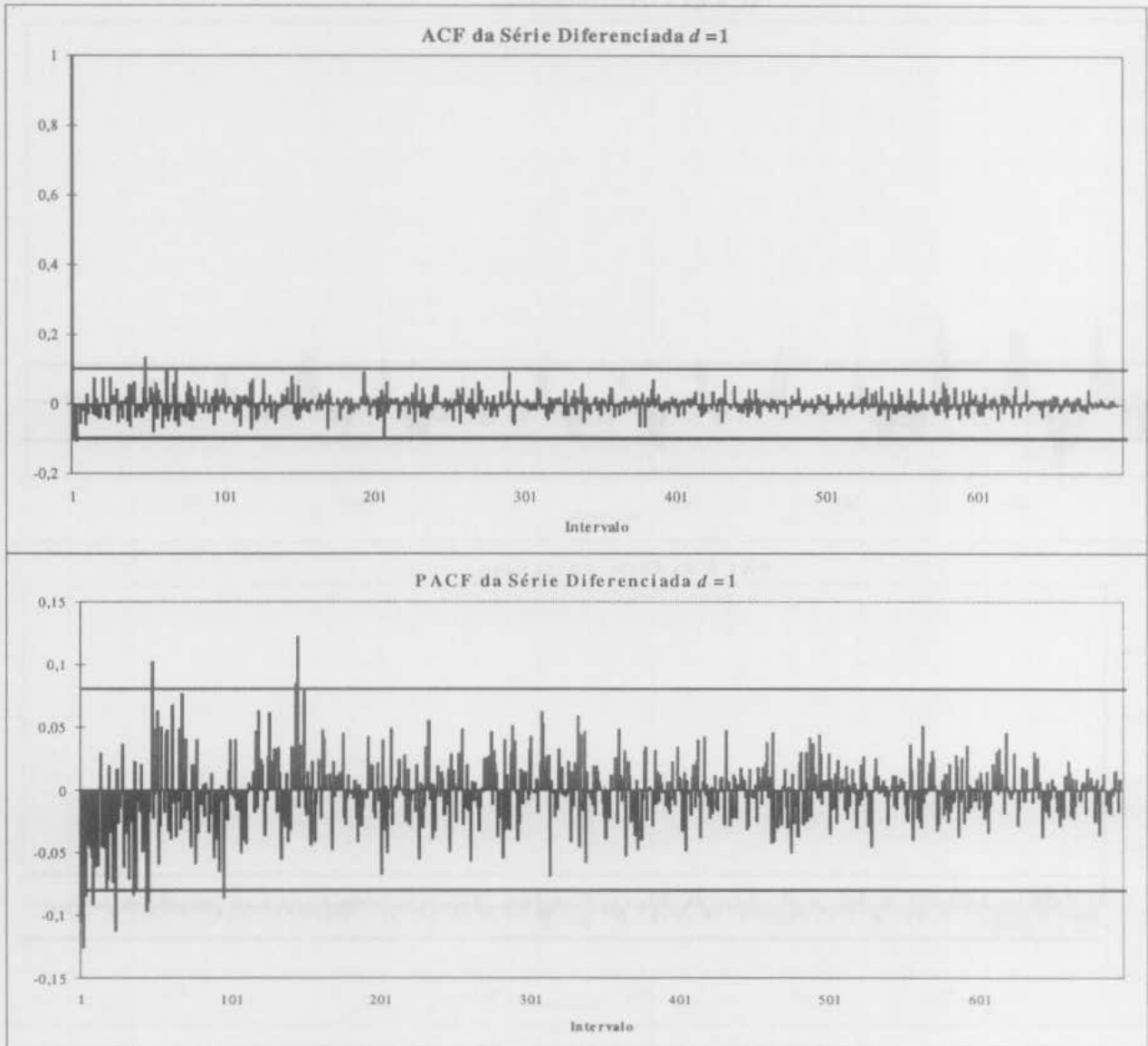
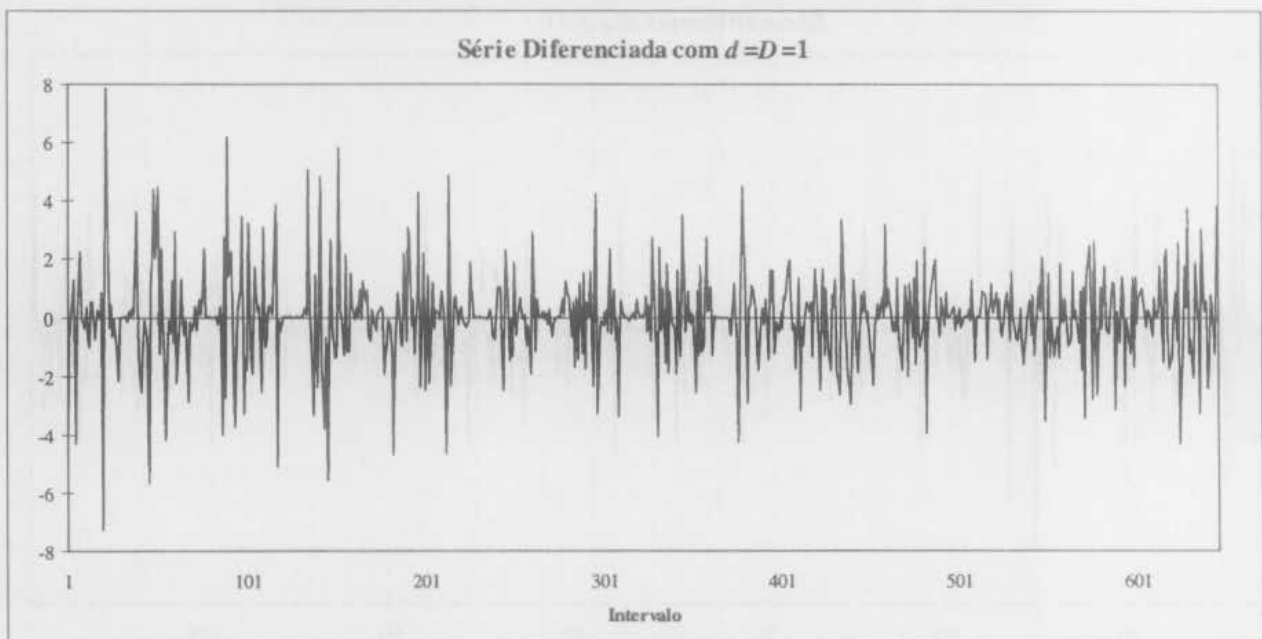


Figura 5 - Plotagem da Série com $d=1$, seu ACF e PACF



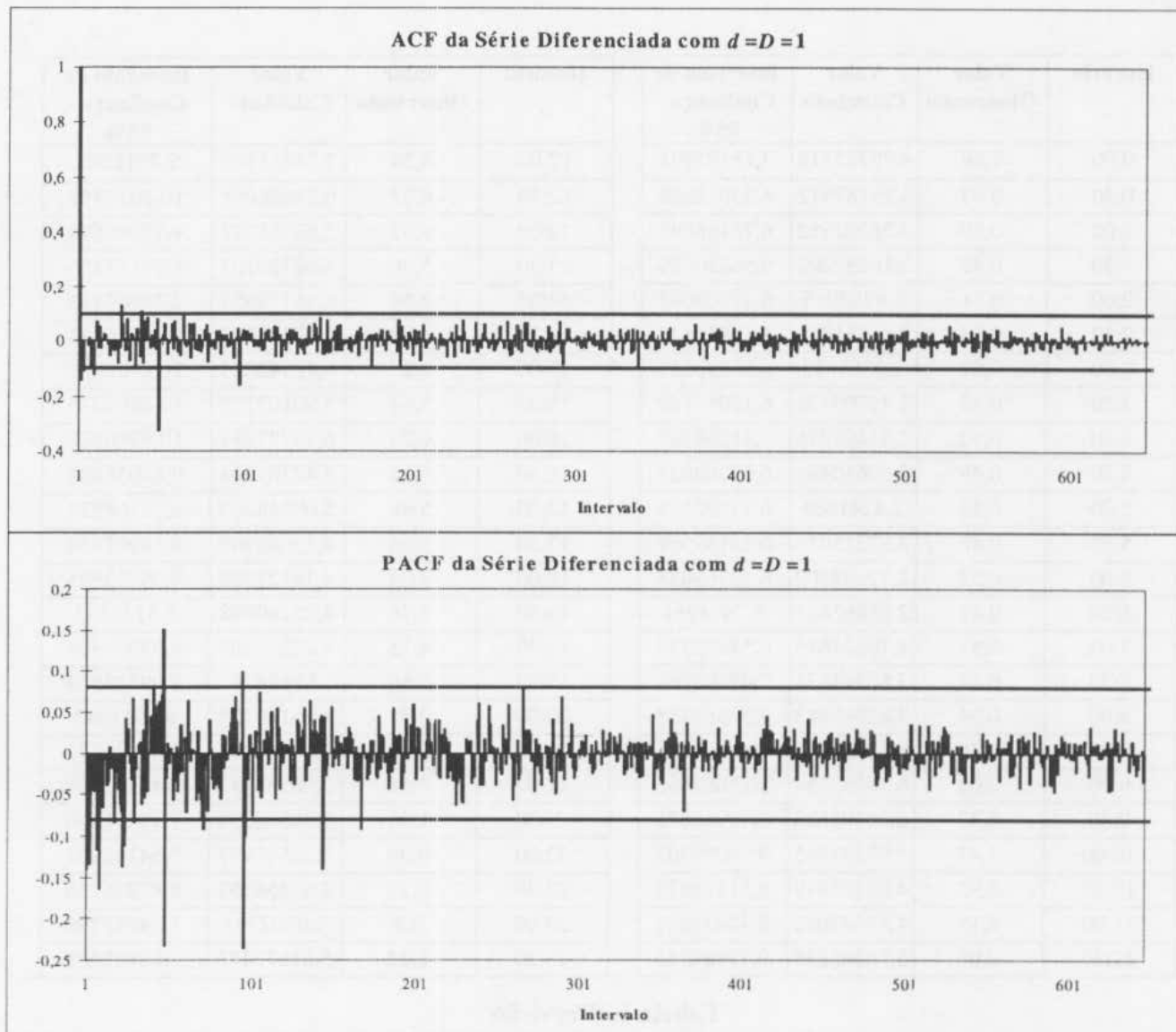


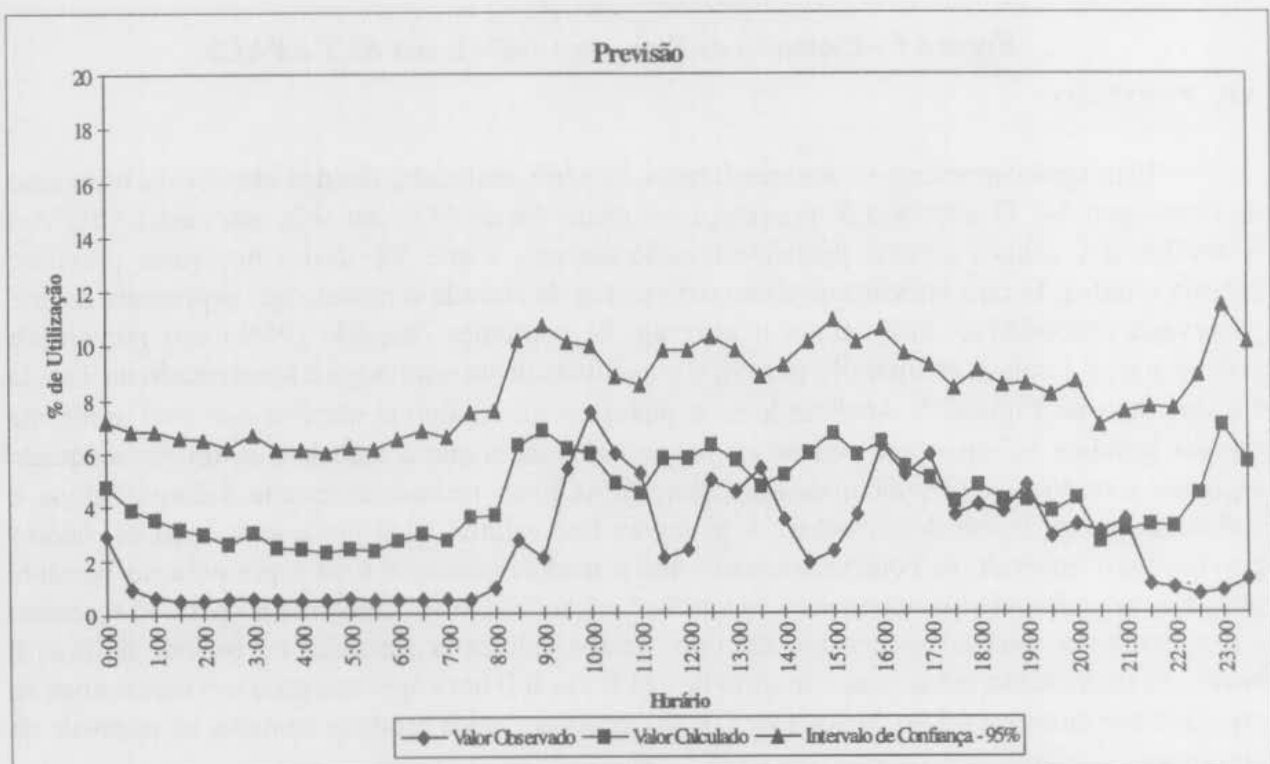
Figura 6 - Plotagem da Série com $d=D=1$, seu ACF e PACF

3.6. PREVISÃO

Esta seção apresenta os valores futuros da série analisada, obtidos através da utilização da linguagem S+. O intervalo de confiança escolhido foi de 95%, ou seja, em média 95% dos intervalos dos valores futuros previstos estarão corretos e que 5% destes intervalos previstos estarão errados. Foram informados como parâmetros de entrada o modelo que representa a série observada (ARIMA(1,1,1)(2,1,0)₄₈), o intervalo de confiança desejado (95%) e o número de pontos que se desejava estimar (48 pontos) e o resultado desta estimativa é apresentado na Tabela I e ilustrado na Figura 7. Analisando-se a plotagem do resíduo contata-se que este apresenta valores bastante baixos e que podem ser desprezados dado que o modelo mostrou-se adequado segundo a estatística Q^* , além disso, a função ACF do resíduo apresenta valores baixos o suficiente para serem desprezados. A plotagem dos valores lidos juntamente com os valores previstos e o intervalo de confiança mostra que o modelo consegue uma representação bastante próxima do padrão de funcionamento da variável em análise: nível crescente no período matutino e decrescente no período vespertino, além dos baixos valores encontrados no período da 0 às 8 horas. As estimativas feitas para o horário das 21 horas à 0 hora apresentaram um maior nível de erro. Merece destaque o fato de que todos os valores observados estavam contidos no intervalo de confiança calculado.

Horário	Valor Observado	Valor Calculado	Intervalo de Confiança - 95%
0:00	2,88	4,79725118	7,15397991
0:30	0,93	3,89187912	6,83033668
1:00	0,59	3,55292952	6,78168098
1:30	0,48	3,16735202	6,56220226
2:00	0,51	2,9783915	6,47766043
2:30	0,53	2,61841845	6,18889173
3:00	0,57	3,0361933	6,65893646
3:30	0,48	2,45709438	6,12081162
4:00	0,52	2,41498976	6,11268295
4:30	0,49	2,30650491	6,03369923
5:00	0,53	2,4241686	6,17792552
5:30	0,48	2,37325046	6,15159569
6:00	0,51	2,72605109	6,52763058
6:30	0,47	2,95862863	6,7824951
7:00	0,51	2,70424675	6,54972372
7:30	0,47	3,60944631	7,47603964
8:00	0,94	3,63949843	7,52683788
8:30	2,75	6,22632736	10,1341275
9:00	2,03	6,78053154	10,7085653
9:30	5,37	6,14308623	10,0911671
10:00	7,47	5,97283345	9,94080303
10:30	5,92	4,82403849	8,81175829
11:00	5,19	4,47697052	8,48431611
11:30	2,06	5,75203858	9,77889561
12:00	2,38	5,74555869	9,7918202
12:30	4,97	6,23668068	10,3022452
13:00	4,52	5,66763587	9,75240598
13:30	5,36	4,68729203	8,79117366
14:00	3,44	5,11770659	9,24060823
14:30	1,88	5,94385208	10,0856844
15:00	2,27	6,70983023	10,8705057
15:30	3,64	5,90107125	10,0805039
16:00	6,24	6,39792431	10,5960297
16:30	5,12	5,42362384	9,64031884
17:00	5,86	5,00848265	9,24368535
17:30	3,66	4,13632909	8,38995874
18:00	4,04	4,78127898	9,05325595
18:30	3,76	4,22280778	8,5130535
19:00	4,75	4,22211303	8,53054998
19:30	2,86	3,768874	8,09542563
20:00	3,31	4,26160375	8,6061945
20:30	3,02	2,65066247	7,01321773
21:00	3,48	3,08538413	7,46583021
21:30	1,08	3,30866095	7,70692506
22:00	0,95	3,22541937	7,64142961
22:30	0,71	4,43858252	8,87226786
23:00	0,8	7,0202393	11,4715295
23:30	1,22	5,61673472	10,0855605

Tabela I - Previsão



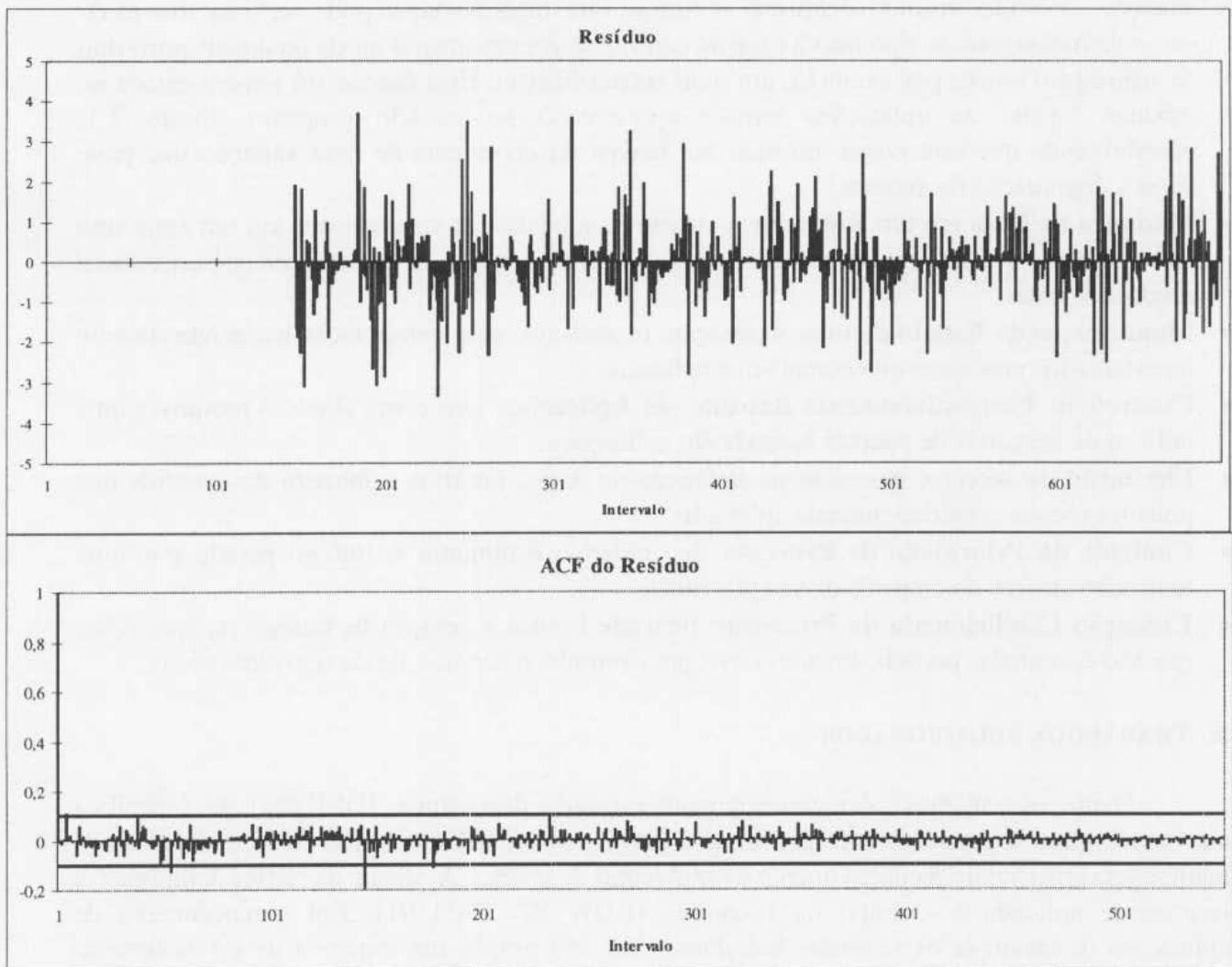


Figura 7 - Plotagem dos Valores Lidos e Previstos, Resíduos e ACF do Resíduo

3.7. CONSIDERAÇÕES SOBRE OS RESULTADOS OBTIDOS

O modelo apresenta uma boa estimativa de intervalos onde os valores futuros da série temporal poderão estar localizados. Devido ao considerável grau de aleatoriedade apresentado pela variável não apenas com relação aos horários mas também nos diversos dias da semana, a estimativa dos valores futuros não pode ser feita, porém é possível estimar um intervalo onde este valor poderá estar incluído, atendendo à forma de previsão estabelecida neste trabalho (vide seção 3). O período das 8 horas às 18 horas foi o que apresentou menor diferença entre os valores estimados e os observados, já o horário das 20 horas até as 6 horas possui um maior erro na estimativa devido a uma maior aleatoriedade da variável neste período.

4. FUNÇÕES DE GERENCIAMENTO PRÓ-ATIVO

Baseando-se no modelo de gerenciamento proposto e na aplicação dos estudos de séries temporais na análise de *baselines*, foram propostas uma série de funções de gerenciamento que fazem uso destas facilidades. Dentre elas, foi implementada a função de notificação ao Gerente da Rede. As funções são:

- **Notificação ao Gerente da Rede:** ela informa ao gerente da rede sobre a ocorrência de situações que podem levar a degradação do sistema e as ações pró-ativas tomadas. Tais

notificações funcionariam como um *log* do sistema, possibilitando ao gerente, uma análise das situações ocorridas visando identificar as causas. Sua implementação pode ser feita através do envio de mensagens do tipo *trap* à estação central de gerenciamento ou de qualquer outro tipo de mensagem como, por exemplo, um *mail* de notificação. Esta função foi implementada no módulo Agente de aplicações remoto apresentado no modelo proposto, Seção 2.1, possibilitando que este possa informar ao gerente da ocorrência de uma situação que pode levar a degradação do sistema.

- **Mudança de Rota em um Roteador:** consiste na mudança de rota baseada em um algoritmo de roteamento que levasse em consideração as informações sobre as aplicações que estivessem sendo utilizadas.
- **Monitoração do Estado de uma Aplicação:** responsável pela monitoração e o gerenciamento do estado dos processos que compõem a aplicação .
- **Controle de Congestionamento Baseado em Aplicações:** tem como objetivo promover uma política de descartes de pacotes baseada em aplicações.
- **Limitação de Acesso:** baseia-se na definição de datas, horários e número de usuários que podem executar uma determinada aplicação.
- **Controle de Prioridade de Processo:** seu objetivo é diminuir o tráfego gerado por uma aplicação através do controle da sua prioridade.
- **Execução Condicionada de Processos:** pretende limitar a geração de tráfego de aplicações que são executadas periodicamente como, por exemplo, o servidor de correio eletrônico.

5. TRABALHOS RELACIONADOS

Dentre os trabalhos sobre gerenciamento pró-ativo destacam-se [DEF 96a] que especifica um modelo para o gerenciamento pró-ativo e [ROC 96] e [SAE 96b] que utilizam técnicas de inteligência artificial no reconhecimento de problemas de redes. A teoria de séries temporais é largamente aplicada no campo da economia [BOW 87, PAN 91]. Em gerenciamento de aplicações destacam-se os seguintes trabalhos: [SIL 95] propõe um esquema de gerenciamento para o tráfego de informações TCP/IP, [PIC 95] reporta um esquema de gerenciamento da aplicação WWW, [CAR 94] propõe um modelo de gerenciamento do protocolo ftp e [NAG 84] propõe um algoritmo de controle de fluxo para o protocolo TCP. Este trabalho apresenta como contribuição a aplicação de séries temporais no campo de Gerenciamento de Redes.

6. CONCLUSÃO

Este trabalho contribuiu para o gerenciamento pró-ativo de três formas: a utilização da teoria de séries temporais na estimativa de possíveis degradações nos sistemas, a proposta de um modelo de gerenciamento pró-ativo, que trata de informações em nível de aplicação e novas funções de gerenciamento pró-ativas para o módulo Sistema de Correção Automática.

A aplicação da teoria de séries temporais a curto prazo permite uma compreensão bastante apurada do padrão de funcionamento do sistema possibilitando a identificação de mudanças no seu comportamento. A aplicação desta teoria a longo prazo permite a análise do impacto de mudanças no sistema, como por exemplo, a inclusão de novos equipamentos. Além disso, com a disponibilização de um volume maior de informações, a modelagem do sistema pode conter um elevado nível de detalhamento. A proposição de um conjunto de passos a serem seguidos quando da aplicação dos conceitos de séries temporais na monitoração de redes foi mais uma contribuição que merece registro.

A disponibilização de informações sobre as aplicações que estão sendo utilizadas no segmento na forma de *baseline* possibilitou um maior nível de detalhamento sobre o sistema monitorado, permitindo uma maior compreensão do seu comportamento.

A especificação de um conjunto de funções de gerenciamento pró-ativas que se utilizam das informações disponíveis sobre as aplicações contribui para o alargamento dos horizontes do gerenciamento pró-ativo, abrindo novas áreas de pesquisas. Além disso, esta atividade demonstrou quanto são fechadas as aplicações utilizadas no tocante a sua administração.

Quando da implementação do sistema de coleta de informações em nível de aplicação, constatou-se uma grande inadequação das aplicações à atividade de monitoração. Foi principalmente neste fase que verificou-se quão grande era a sobrecarga causada pelo programa *Beholder*, que implementa a RMON MIB, no equipamento em que ele estava sendo executado. A resolução de tal problema resultou na utilização de um equipamento dedicado para a tarefa de coleta de dados. Uma alternativa para a coleta de dados em nível de aplicação bastante eficiente é a utilização do programa desenvolvido em [SIL 95], solução esta não adotada neste trabalho porque foram utilizadas diversas facilidades disponibilizadas pelo programa *Beholder*. Com relação às amostras coletadas sobre as aplicações, constata-se a grande aleatoriedade das mesmas, dificultando sua análise e não permitindo sua modelagem através de séries temporais.

A utilização do programa *Beholder* levou a uma série de pesquisas sobre a sua configuração, funcionamento e utilização. Além disso o emprego das facilidades de filtragem e armazenamento de pacotes, normalmente pouco utilizadas devido à complexidade existente na sua configuração e principalmente na compreensão do seu funcionamento, mereceu uma considerável carga de esforço. Como resultado destas pesquisas, criou-se um conjunto de programas do tipo *batch* capazes de configurar os grupos da RMON 1 MIB, implementados no programa *Beholder*, utilizados neste trabalho.

Através da aplicação do modelo de gerenciamento pró-ativo proposto e dos estudos sobre séries temporais foi implementado um protótipo do sistema de verificação. Com a utilização da RMON MIB e de programas complementares, foi monitorado um segmento do Instituto de Computação com o objetivo de compor um perfil de funcionamento da rede, tanto até o nível de rede quanto a nível de aplicação. Visando aplicar os conceitos de séries temporais foi implementado um módulo estimador que busca estimar um intervalo para o parâmetro nível de utilização do segmento. Este intervalo mostrou-se bastante apropriado para o intervalo de tempo que foi modelado (30 minutos). Em intervalos menores (30 segundos), as estimativas de limite superior demonstraram ser muito baixas dada a variância que apresenta a variável modelada, duas seriam as soluções para este fato: a utilização de um intervalo de confiança menor, que elevaria o limite superior estimado ou o aumento da tolerância aos avisos de que o nível de utilização do segmento ultrapassou o limite superior.

Como continuidade das pesquisas, sugere-se o aprimoramento do sistema de verificação com o melhora da estimativa de situações que possam degradar a rede. O avanço nos estudos sobre a aplicação de séries temporais no campo do gerenciamento de redes poderia ser feito através de modelagens baseadas em um maior volume de informações e da caracterização de um número maior de sazonalidades tais como semanas de um mês e meses do ano.

Ainda no tocante a trabalhos futuros, propõe-se continuar os estudos sobre a especificação do sistema de correções automáticas e das funções de gerenciamento pró-ativas, tratando de problemas como a definição de um algoritmo de roteamento baseado em aplicações e a implementação da função de controle de congestionamento baseado em aplicações.

AGRADECIMENTOS

Agradecemos à Fapesp e ao CNPq pelo apoio financeiro.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- [BIE 97] BIERMAN, A., Iddon, R. "Remote Network Monitoring MIB Protocol Identifiers", RFC 2074, jan., 1997.
- [BOW 87] BOWERMAN, Bruce L., O'CONNEL, Richard T. "Time Series Forecasting - Unified Concepts and Computer Implementation". PWS Publishers, 1987.
- [BOX 70] BOX, GEORGE P., JENKINS, Gwilym M. "Time Series Analysis - forecasting and control". Holden-Day, 1970.
- [CAR 94] CARRILHO, J.A., MADEIRA, E.R.M. "A Scheme for FTP Management". INET'94 Annual Conference on the Internet Society, Praga, República Tcheca, Junho 1994, pp. 562-1/562-10
- [CHA 84] CHATFIELD, Christopher. "The Analysis of Time Series". Chapman and Hall, 1984.
- [DEF 96] DE FRANCESCHI, Analúcia Schiaffino Morales, KORMANN, Luiz Fernando, WESTPHALL, Carlos Becker. "A Performance Application for Proactive Network Management". Second IEEE Workshop on Management Systems, jun., 1996.
- [DEF 96a] DE FRANCESCHI, Analúcia Schiaffino Morales. "Aplicação de Desempenho para Validar a Gerência Pró-ativa de Redes", Tese de Mestrado, UFSC, 1996.
- [EVE 94] EVERITT, Brian S. "A Handbook of Statistical using S-plus". Chapman & Hall Publisher, 1994.
- [NAG 84] NAGLE, J. "Congestion Control in TCP/IP Internetworks", Computer Communications Review, vol. 14, pp. 11-17, oct. 1984.
- [PAN 91] PANKRATZ, Alan. "Forecasting with Dinamic Regression Models". Wiley-Interscience Publication, 1991.
- [PIC 95] PICOLO, Carlos; VEIGA, Pedro. "Management of a WWW Server using SNMP". <http://smurfland.cit.buffalo.edu/NetMan/Papers.html>, may, 1995.
- [ROC 96] ROCHA, Marco Antonio da., FERNANDEZ, Luiz Fernando Nunes, WESTPHALL, Carlos Becker. "Gerência Pró-ativa de Redes de Computadores usando Agentes e Técnicas de Inteligência Artificial". Anais do 14º Simpósio Brasileiro de Redes de Computadores, Fortaleza, 1996.
- [SAE 96a] SÁENS Artola, Esmilda. "Um Sistema Especialista para a Gerência Pró-ativa Remota". Anais do 14º Simpósio Brasileiro de Redes de Computadores, Fortaleza, 1996.
- [SAE 96b] SÁENS Artola, Esmilda. "Sistema Especialista para a Gerência Pró-ativa de Redes". Tese de Mestrado, UFRGS, 1996.
- [SIL 95] SILVEIRA, C. K., MADEIRA, E.R.M.. "Um esquema para o gerenciamento do tráfego de aplicações em redes TCP/IP". 13o Simpósio Brasileiro de Redes de Computadores - SBRC'95 - Belo Horizonte, MG, Maio 1995, pp. 567-585
- [SPE 94] SPECTOR, Phill. "An Introduction to s and S-plus". Duxburg Publisher, 1994.
- [WAL 91] WALDBUSSER, S. . "Remote Network Monitoring Management Information Base". RFC 1271, nov., 1991.
- [WAL 95a] WALDBUSSER, S. . "Remote Network Monitoring Management Information Base". RFC 1757, feb. 1995.
- [WAT 97] WATZKO Neto, Francisco. "Gerenciamento Pró-ativo: consideração do tráfego de aplicação e utilização de séries temporais" Tese de Mestrado, UNICAMP, 1997.