

Modelo de tráfego para conexão à Internet utilizando conexões discadas

Carlos Marcelo Pedroso¹, Mateus Cruz², Keiko Fonseca²

¹Pontifícia Universidade Católica do Paraná - PUCPR
R. Imaculada Conceição 1155, Curitiba, Paraná

²Universidade Tecnológica Federal do Paraná - UTFPR
Av. Sete de Setembro 3165, Curitiba, Paraná

c.pedroso@pucpr.br, mateus@cpgei.cefetpr.br, keiko@cpgei.cefetpr.br

Abstract. *Recent reports present a significant number of users that still access the Internet through dial-up connections. This kind of access should remain important until broadband access systems became accessible at affordable prices not only in Brasil but worldwide. Particularly in Brazil, dial-up traffic exchange between public switched telephone network (PSTN) providers represents an important revenue to them. Indeed, the knowledge of the dial-up traffic behavior is fundamental to the dimensioning of modem pool and preserve/improve service quality. This paper presents a traffic model for Internet access through dial-up connections. The model was established with recorded data from a national telecom switching operator related to call charges period. The resulted model updates the models found in the literature and allowed better resources dimensioning and management as well system capacity planning.*

Resumo. *Estudos recentes indicam que uma significativa parcela dos usuários da Internet ainda se utilizam de Modems e linhas discadas. Esta situação deve permanecer ainda por algum tempo até a completa difusão de sistemas de acesso de banda larga no Brasil ou mesmo no mundo, a preços acessíveis. Os acessos discados representam importante parcela das remunerações das operadoras do sistema de telefonia fixa comutada (STFC) no Brasil. Assim sendo, o conhecimento do comportamento do tráfego originado deste tipo de acesso é fundamental para o dimensionamento eficiente de recursos das operadoras do STFC. Este artigo apresenta um modelo de tráfego para um sistema de acesso à Internet utilizando linhas discadas estabelecido a partir de dados reais obtidos de uma operadora de telecomunicações. O modelo proposto a partir da análise detalhada desses dados em relação ao regime de tarifação telefônica no Brasil atualiza os modelos apresentados na literatura até então, permite estabelecer e testar políticas de gerência de recursos, bem como realizar um melhor planejamento de capacidade do sistema.*

1. Introdução

Conexão por linha discada (*dial-up*) é um tipo de acesso à Internet feito através das operadoras do Sistema Telefônico Fixo Comutado (STFC) na forma de uma ligação telefônica

destinada ao *Provedor de Acesso a Serviços Internet (PASI)*. O usuário origina uma chamada telefônica com um modem e uma linha telefônica para se ligar a um nó de uma rede de computadores de um provedor de acessos. A partir desse momento, o PASI encarrega-se de fazer o roteamento para a Internet. Este tipo de acesso perde adeptos devido à massificação de acessos de banda larga, como xDSL e ligações por cabo. No entanto, a utilização de linhas discadas está longe de ser extinta.

Um levantamento do Ibope/NetRatings revela que 4,4 milhões de brasileiros ou 36,3% dos usuários domiciliares acessaram a internet utilizando linhas discadas em janeiro de 2006 [Ibope 2006]. A tarifação realizada atualmente pelas operadoras de telefonia fixa no Brasil utiliza o pulso único, desonera os horários entre 0h e 6h durante a semana e a partir de 14h do sábado até 6h da segunda-feira. No entanto, existem várias operadoras oferecendo linhas telefônicas exclusivas para realização do acesso à internet na modalidade de linhas discadas, com tarifas diferenciadas e PASI próprio.

Também nos Estados Unidos da América o número de acessos residenciais com banda larga aumentou 40%, de 60 milhões para 84 milhões de março de 2005 a março de 2006. No entanto, o número de conexões discadas aumentou em 25 milhões neste mesmo período, o que revela que esta modalidade de acesso está longe de ser extinta [News 2006].

Uma das principais explicações para a longa vida das conexões discadas está no preço. Usuários americanos reportam taxas mensais de \$41 para acesso utilizando cabo, \$38 para serviços xDSL e \$18 para conexões discadas. No Brasil, várias operadoras de telecomunicações oferecem linhas telefônicas com contrato diferenciado para realização do acesso à internet através de linhas discadas. Neste caso, o preço do acesso fica em torno de R\$30,00 mensais contra R\$55,00 para o acesso mais barato em banda larga.

Este trabalho apresenta a caracterização do tráfego para os sistemas *dial-up*, realizado sobre dados de uma operadora de telecomunicações no Brasil (ver seção 8). Foram analisados *traces* de sistemas de acesso discado de grandes cidades brasileiras de população maior que 1,5 milhão de habitantes. Exceto pelo trabalho de [Iversen et al. 2000], não é de nosso conhecimento trabalhos recentes que reportem caracterização deste tipo de tráfego a partir de dados reais de operadoras de telecomunicações mas sim de ambientes universitários, o que limita o escopo destes modelos. Esta constatação também é notada em [Novak et al. 2003] que atribui ao fato dos resultados serem proprietários e se reportarem ao tráfego após o banco de modems, sem levarem em conta as perdas de conexão ali ocorridas.

A identificação de um modelo que caracterize o tráfego oferecido ao acesso *dial-up* é fundamental para o estudo de alternativas de planejamento de capacidade do sistema bem como para o estabelecimento de políticas de gerência de recursos disponíveis [Vanucci et al. 2003] [Patek and Campos-Nanez 2000] [Schikora and Godfrey 2006]. Os resultados obtidos do modelo estabelecido a partir da análise de dados de uma rede real são discutidos com os da literatura estudada em termos de planejamento da capacidade do sistema telefônicos e redes de dados. A nossa contribuição está na caracterização mais precisa do intervalo entre chegadas de conexão discada e do tempo de conexão aceita e fez uso do conhecimento do regime de tarifação das conexões em função do horário no Brasil. Os resultados obtidos do modelo aqui proposto determinaram uma maior precisão no cômputo da probabilidade de bloqueio do sistema.

O trabalho está estruturado da seguinte forma: a seção 2 apresenta o sistema em estudo e a seção 3 identifica o problema de tráfego do ponto de vista da operadora de telecomunicações. A metodologia utilizada para caracterizar o tráfego de acesso discado é abordada na seção 4 e a seção 5 apresenta a modelagem do sistema a partir da análise dos dados coletados. A seção 6 discute e avalia o modelo e a sua aplicação. Finalmente, a conclusão é apresentada na seção 7.

2. Acesso discado à Internet

Nos sistemas de telefonia a rede de acesso é definida como sendo a conexão entre os usuários e as centrais telefônicas. Anteriormente as redes de acesso eram normalmente construídas utilizando cabos de fios metálicos em que um par era dedicado a cada aparelho telefônico conectado até a central telefônica. Este par, juntamente com os recursos da central telefônica dedicados ao usuário é conhecido como acesso ou linha telefônica.

As linhas telefônicas de vários assinantes chegam às centrais telefônicas e são conectadas entre si quando um usuário denominado (A) deseja falar com outro usuário (B). Convencionou-se chamar de A o usuário que origina a chamada e de B aquele que recebe a chamada, sendo comutação o termo usado para indicar a conexão estabelecida entre os usuários. A central de comutação estabelece circuitos temporários entre os usuários permitindo o compartilhamento de meios e promovendo uma otimização dos recursos disponíveis.

No acesso discado à Internet, o usuário se relaciona com duas entidades: a operadora de telecomunicações do Sistema de Telefonia Fixo Comutado (STFC) e o Provedor de Acesso a Serviços Internet, considerado um provedor de serviço de valor adicionado. Esta conexão é feita, na maior parte dos casos, através das operadoras de telefonia fixa, na forma de uma ligação telefônica. O usuário (A) origina uma chamada telefônica destinada ao PASI (B), utilizando o modem do seu computador e a sua linha telefônica.

A central de comutação ao perceber que uma chamada tem destino ao PASI, realiza a entrega desta ao mesmo, que ao recebê-la, inicia a troca de informações com o computador e estabelece uma conexão para realizar a autenticação do usuário. Este tipo de conexão é chamada de acesso discado à internet, e permite a comunicação via modem entre o computador do usuário, linha telefônica e o PASI, enquanto durar a ligação telefônica. Normalmente, esta comunicação está limitada a taxas de 56kbit/s.

O custo de acesso discado à Internet possui dois componentes:

- Custo da ligação telefônica, cobrado por pulsos ou minutos, conforme o produto ou plano de serviço da operadora de serviço telefônico.
- Custo do provedor de acesso, em geral um valor fixo independente da utilização ou igual a zero no caso de provedores de Internet Grátis.

Dependendo da utilização, o custo da ligação telefônica pode se tornar o item preponderante no custo total de acesso à Internet, pois se tratam de ligações telefônicas de elevada duração. Assim sendo, algumas operadoras de telecomunicações do STFC já encaminham o tráfego destinado aos PASI de modo diferente das ligações telefônicas de voz. Este tráfego é desviado para redes IP que se conectam aos PASI.

Existe uma situação particular, durante o horário de tarifa reduzida, em que a operadora de telecomunicações do STFC do usuário paga mais de interconexão para a

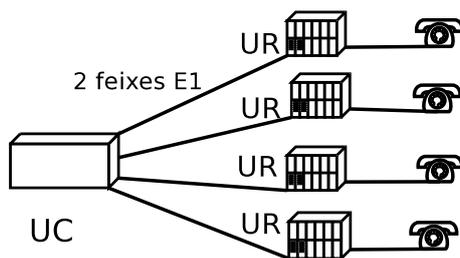


Figura 1. Topologia típica

operadora à qual está conectada o PASI, do que o que recebe do assinante. Isto ocorre porque enquanto o assinante paga apenas o valor de um pulso ou minuto por conexão, independentemente da duração desta, a sua operadora continua remunerando a outra com tarifas de uso de rede associadas ao tempo de sua utilização. Desta forma, é de particular interesse das operadoras que fornecem o serviço de acesso discado que este seja oferecido com qualidade pois este impacta positivamente na sua remuneração.

O modelo atual de acesso discado apresenta ainda uma série de pontos que limitam, na prática, um acesso mais efetivo à Internet. O custo da conexão telefônica é, em muitos casos elevado para a população em geral, fazendo com que o tempo de conexão seja reduzido ou deslocado para a madrugada e fins de semana nos horários de tarifa reduzida. Esta nova característica de utilização da linha telefônica para acesso à internet e voz, embora não simultâneos, veio determinar um perfil de tráfego diferenciado dos modelos tradicionalmente utilizados para tráfego de voz.

A revisão da literatura apresenta análises de tele-tráfego de acesso discado mas que se referem a dados pouco recentes [Naldi 2002], [Färber et al. 1998b], [Färber et al. 1998a] ou obtidos em redes universitárias envolvendo um número pequeno de usuários [Novak et al. 2003], a exceção do trabalho de [Iversen et al. 2000]. A contribuição do nosso trabalho é particular e inédita por realizar a modelagem do processo de atendimento e chegada de conexões a partir do tráfego real de uma operadora brasileira de telecomunicações, cujo perfil de usuário é grandemente afetado pelo regime atual de tarifação, tal como reportado em [Vannucci et al. 2003]. Os dados de uma rede real nos permitiu gerar um modelo mais realista de tráfego e obter resultados de cálculo de probabilidade de bloqueio de requisições no banco de modems de acesso *dial-up* onde outros trabalhos o fizeram com limitações [Novak et al. 2003].

3. Identificação do problema

Uma topologia típica comumente encontrada nas operadoras de telecomunicações do STFC é apresentada na Figura 1. Descrita como tipo *mestre-escravo*, na figura, o mestre é denominado de Unidade Central (UC) e o escravo de Unidade Remota (UR).

Nesta topologia o conjunto de equipamentos denominados UC e UR estão conectados respectivamente à central telefônica local e à rede de acesso. Considera-se neste estudo um sistema operando com a tecnologia PDH (Plesiochronous Digital Hierarchy) [ITU-T 1972]¹.

¹Nenhuma restrição é imposta ao uso da tecnologia SDH (Synchronous Digital Hierarchy) [ITU-T 1993]

A UC é conectada com feixes E1 (2M bits/s, 30 canais) à central telefônica local. A UR é responsável por fornecer o acesso final aos clientes e se encontra conectada à UC. Quando um usuário inicia uma chamada telefônica, ele irá reservar um dos canais disponíveis do enlace (ou *time slot*) ao longo do tempo de duração da chamada. Este procedimento corresponde à reserva realizada normalmente para um canal de voz. O número de canais a serem disponibilizados para o serviço é o objeto do estudo realizado.

Usualmente, o planejamento de capacidade da UC determina a quantidade de canais a serem reservados baseado em modelos aplicados ao tráfego de voz [Ferrari 2005] [Iversen 2005]. O dimensionamento estabelece a taxa de ocupação que minimize o bloqueio por falta de recursos.

Quando este dimensionamento se faz baseado em modelos de tráfego de voz, as especificações desejadas de bloqueio para uso da linha telefônica para acesso a Internet não são atendidas e o número de canais entre UC e UR deve ser redimensionado baseado no novo perfil de tráfego.

A topologia da Figura 1 é um exemplo que descreve uma Central telefônica ligada a UC através de 4 feixes E1, a qual disponibiliza *time slots* de acordo com o solicitado pelas UR's; a UC é conectada a até 4 UR's com 2 feixes(estáticos) E1 por UR, somando 8 feixes E1 entre UC e UR's; uma UR pode estar conectada a 480 clientes (linhas telefônicas) sendo expansível a 960, porém disponibiliza o máximo de 2 feixes E1, ou seja 60 *time slots*, que são ocupados de acordo com a demanda de tráfego.

No exemplo da Figura 1, na Hora de Maior Movimento, é possível a ocorrência de uma grande quantidade de perda de chamadas, cuja origem está entre UC e UR1. Estas perdas referem-se à quantidade de solicitações para chamadas originadas com destino à UR1 que ultrapassam o limiar de canais disponíveis, em especial quando o sistema é dimensionado baseando-se no perfil de tráfego de voz. Este problema pode ser resumido em um problema um sistema de filas com vários servidores e perda de chamadas onde cada *time slot* é equivalente a um servidor disponível.

A situação de congestionamento é agravada com o crescimento da planta da operadora e da ocupação dos equipamentos UC e URs devido aos novos produtos ou serviços oferecidos, principalmente serviços de acesso *dial-up*. Este fatos impactam no uso das linhas telefônicas e modificam o perfil de uso do sistema, criando assim a necessidade de estudar os novos perfis de tráfego e metodologias para a realização do planejamento de capacidade.

4. Metodologia

O controle e gerenciamento do tráfego real da rede da operadora em geral se baseia na realização de medições ativas ou passivas e posterior análise de dados obtidos. Estes dados são analisados estatisticamente e o tráfego, uma vez caracterizado, pode determinar um modelo adequado a ser aplicado para o dimensionamento do sistema.

4.1. Coleta de dados

Os dados necessários para o desenvolvimento do modelo foram obtidos a partir de uma ferramenta desenvolvida particularmente para coleta automática de tráfego para equipamentos com topologia "mestre-escravo" utilizada na operadora de telecomunicações em

Tabela 1. Conjunto de dados em estudo

Data	Local
31/03/2006, sexta-feira	Cidade1
31/03/2006, sexta-feira	Cidade2
01/04/2006, sábado	Cidade1
01/04/2006, sábado	Cidade2
02/04/2006, domingo	Cidade1
02/04/2006, domingo	Cidade2
03/04/2006, segunda-feira	Cidade1
03/04/2006, segunda-feira	Cidade2
04/04/2006, terça-feira	Cidade1
04/04/2006, terça-feira	Cidade2

estudo. A coleta de dados foi realizada diretamente nas URs, tornando possível inclusive determinar o número de ligações não completadas por falta de capacidade no sistema. Os principais benefícios do desenvolvimento desta ferramenta foram:

- Disponibilidade de verificação de tráfego em data e horários de interesse específicos;
- Ferramenta auxiliar no planejamento e controle da qualidade de rede;
- Criação de histórico (banco de dados) do tráfego por elemento instalado na planta da operadora;
- Tomada de decisão embasada em necessidade real de recursos (melhor alocação de recursos e investimentos).

Através da ferramenta, dados reais de tráfego gerado por acesso discado em grandes cidades brasileiras foram coletados. Um exemplo do conjunto de dados em estudo é mostrado na Tabela 1, onde a Cidade1 possui uma população atual de aproximadamente 1,5 milhão de habitantes e na Cidade2 vivem 2,3 milhões de habitantes.

Os dados são registrados em termos de chamada individual, instante de chegada no sistema, instante de término, origem, destino², identidade do modem-pool, entre outros de interesse limitado para este trabalho. A precisão dos dados é da ordem de segundos para as chegadas de conexões e de décimos de segundo para o tempo de conexão.

A modelagem de tráfego objeto da próxima seção resulta das métricas de desempenho extraídas com a ferramenta citada.

5. Modelagem do sistema

Os passos seguidos para o estabelecimento do modelo abrangeram a construção dos histogramas dos dados coletados, análise de intervalos de observação em função do regime de tarifação, verificação de estacionariedade das séries temporais obtidas, realização dos testes de aderência, caracterização dos processos em termos de conjuntos de dados estatísticos e variáveis de tráfego.

O sistema foi modelado observando-se as seguintes métricas:

²As identidades pessoais dos assinantes foram convertidas em valores numéricos de forma a proteger a privacidade dos usuários da operadora

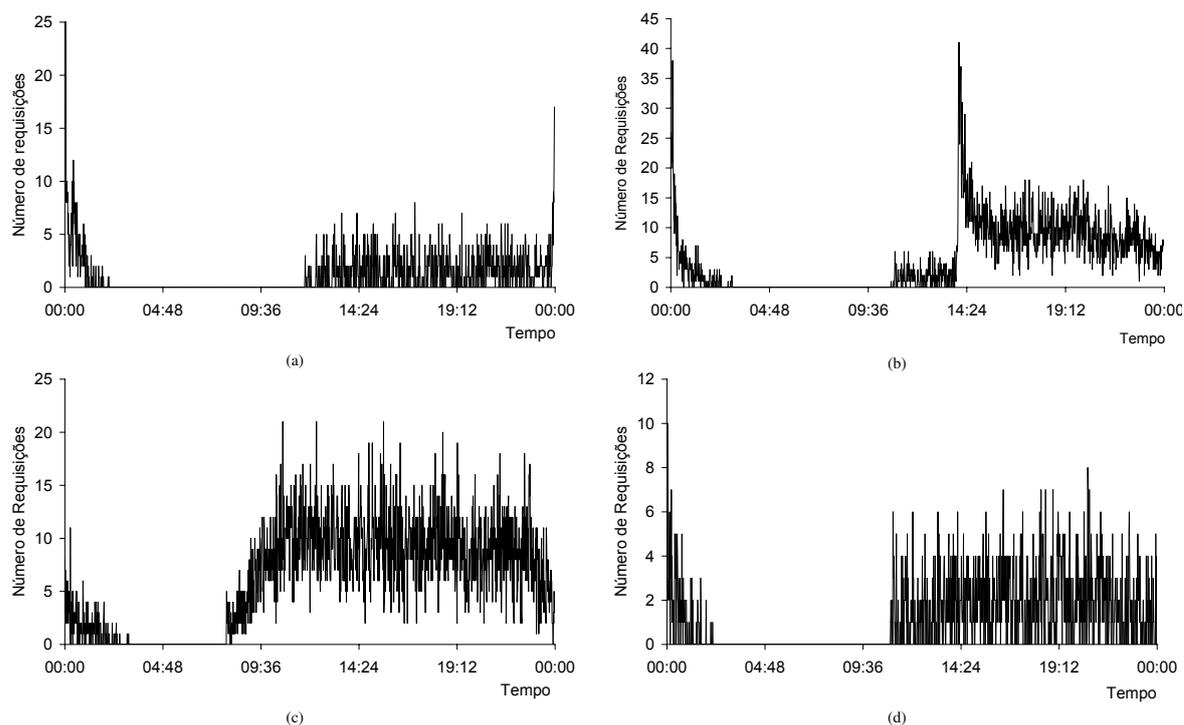


Figura 2. Número de chegadas ao longo do dia para Cidade2 em (a) 31/03/2006, (b) 01/04/2006, (c) 02/04/2006 e (d) 03/04/2006

1. Tempo de conexão de um dado cliente (ou duração da conexão, caracterizando o processo de atendimento);
2. Intervalo entre chegadas (caracterizando o processo de chegadas).

A série temporal típica de acessos ao sistema é mostrada na Figura 2 e reflete os regimes de tarifação atuais para o sistema de telefonia. De acordo com a regra em vigor para tarifação telefônica, de segunda a sexta-feira, entre 0 e 6 horas; aos sábados, da 0 às 6 e das 14 às 24 horas; aos domingos e feriados nacionais as ligações são tarifadas por completamento. Nos demais horários, as ligações são tarifadas por pulsos.

As solicitações para conexão nos dias comerciais possuem comportamento indicado na Figura 2(a) e (d), que mostra o número de chegadas para sexta e segunda feira, respectivamente. Nestes dias, a partir de 00:00h ocorre um aumento do número de novas conexões devido ao menor custo de ligações, configurando vários regimes de operação durante o dia. A análise da série com as chegadas de requisições de conexão mostra que o comportamento do usuário varia conforme o período do dia, sendo influenciado principalmente pelo regime de tarifação. Também fica claro que o processo de chegadas não pode ser estudado em conjunto, pois o sistema é periódico e não estacionário. Assim sendo, a investigação se concentrou em determinadas faixas horárias onde o sistema apresenta estacionariedade. Para tanto observou-se a série de chegadas de diversos dias, por exemplo, em dias de semana nos intervalos de horários de 9:00h à 12:00h, de 15:00h a 18:00h ou de 00:00h a 01:00h.

A Figura 3 mostra a função de auto-correlação dos intervalos entre chegadas tomados como um todo em comparação com a função de auto-correlação do intervalo entre 19:00h e 23:00h. Percebe-se que o decaimento da função de auto correlação é mais

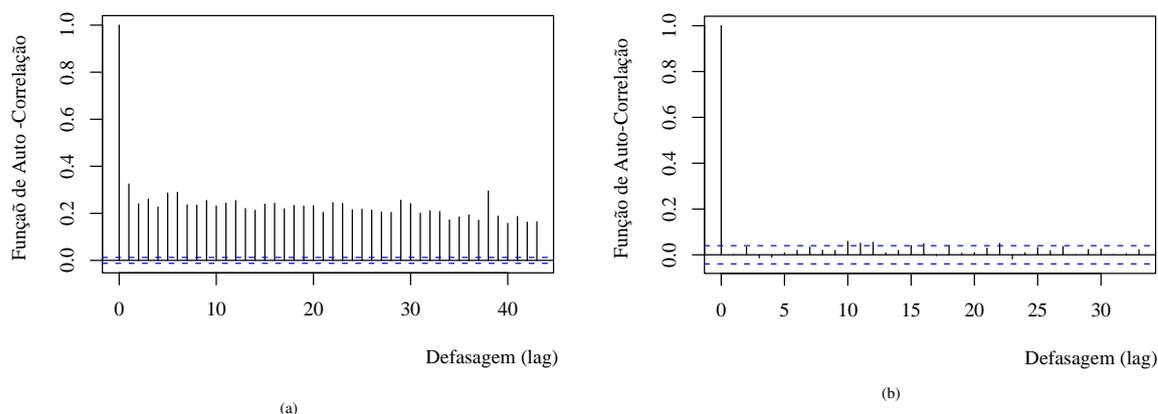


Figura 3. Função de auto-correlação para (a) os intervalos entre chegadas tomadas como um todo e (b) para o horário entre 19:00h e 23:00h

acentuado para os dados tomados entre 19:00h e 23:00h, indicando a estacionariedade da série neste período. Foram observados diversos períodos de tempo durante o dia com característica semelhante, e também comportamentos diferenciados para dias de semana, sábados e domingos - devido à características particulares do regime de tarifação. Nos dias de semana observa-se um particular aumento de tráfego no intervalo entre 00:00h e 01:00h, também devido à redução na tarifa telefônica neste horário. De fato, este é o período de maior quantidade de requisições de conexão observado.

Um horário particularmente congestionado é de 00:00h à 01:00h e representa o período onde o ritmo de chegadas de requisições de conexão é maior. O estudo do intervalo entre chegadas procurou comparar a distribuição de probabilidade empírica encontrada no sistema com uma distribuição de probabilidade teórica conhecida. A Figura 4 mostra o histograma, *box-plot* e *QQ-plot* para o intervalo entre chegadas observado no dia 01/04/2006. Estes gráficos apresentam um forte indicativo de que os dados podem ser aderentes a uma distribuição exponencial³.

Para comprovar esta hipótese, foi realizado o teste de aderência da distribuição exponencial aos dados para os diversos dias em estudo no horário entre 00:00h e 01:00h. O resultado é mostrado na Figura 5, que mostra a comparação entre a distribuição empírica (linha contínua) e a distribuição exponencial (linha tracejada), onde pode ser observada a boa aderência aos dados.

O teste de aderência foi realizado também sobre outros intervalos de horários e a distribuição que se ajusta melhor aos dados foi a distribuição exponencial, conforme mostra a Tabela 2.

Outra variável importante em estudo é a duração da conexão. A série temporal representando o tempo de duração da conexão ao longo de todos os dias em estudo para as conexões realizadas na Cidade2 é mostrada na Figura 6. Observa-se que esta série apresenta um bom nível de estacionariedade, indicando que ela pode ser tratada como um todo.

O resultado é apresentado na Figura 7. Observou-se que a distribuição do tempo

³Função de densidade de probabilidade: $f(x) = \lambda e^{-\lambda x}$, onde λ representa a taxa de chegadas

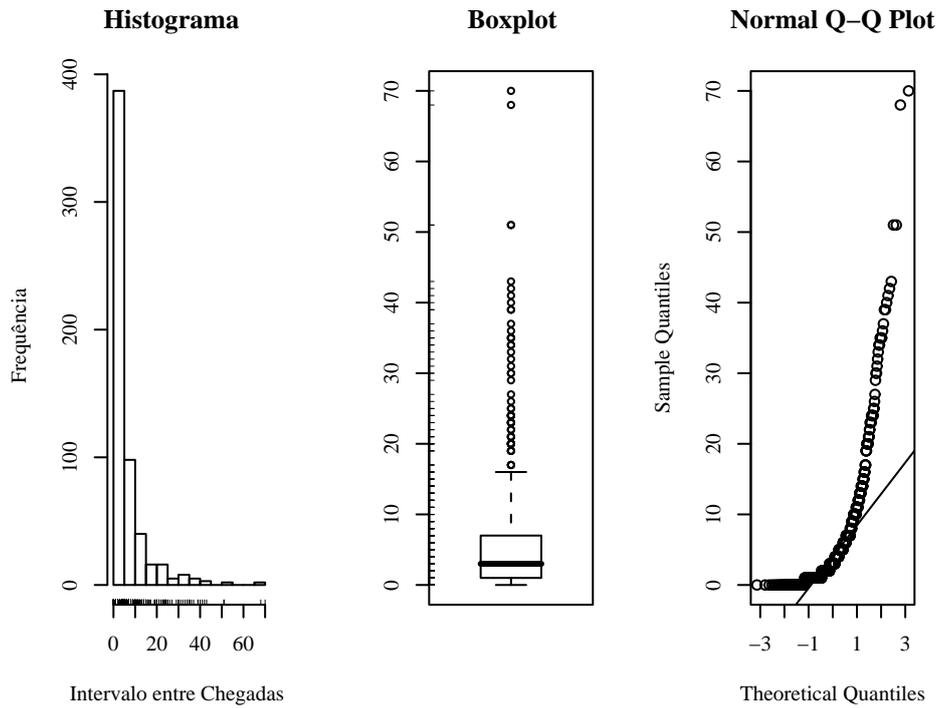


Figura 4. Histograma, diagrama *boxplot* e *QQPlot* para o intervalo entre chegadas no intervalo 00:00h a 01:00h

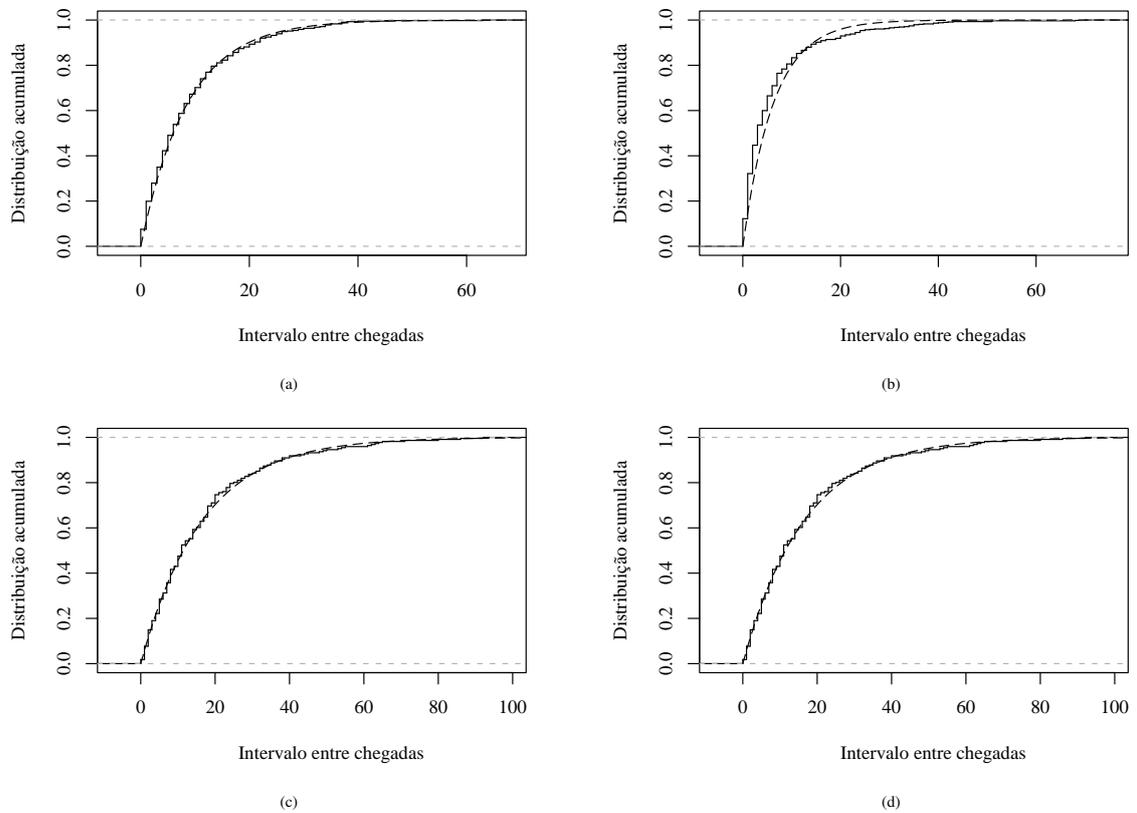


Figura 5. Distribuição de probabilidade do intervalo entre chegadas de requisições (Cidade2) entre 00:00h e 01:00h em 31/03/2006 (a), 01/04/2006(b), 02/04/2006(c) e 03/04/2006(d)

Tabela 2. Resultados dos testes de Aderência

Intervalo da observação	Local	Modelo
00:00h a 01:00h	Cidade1	Exponencial
09:00h a 12:00h	Cidade1	Exponencial
14:00h a 18:00h	Cidade1	Exponencial
19:00h a 23:00h	Cidade1	Exponencial
00:00h a 01:00h	Cidade2	Exponencial
09:00h a 12:00h	Cidade2	Exponencial
14:00h a 18:00h	Cidade2	Exponencial
19:00h a 23:00h	Cidade2	Exponencial

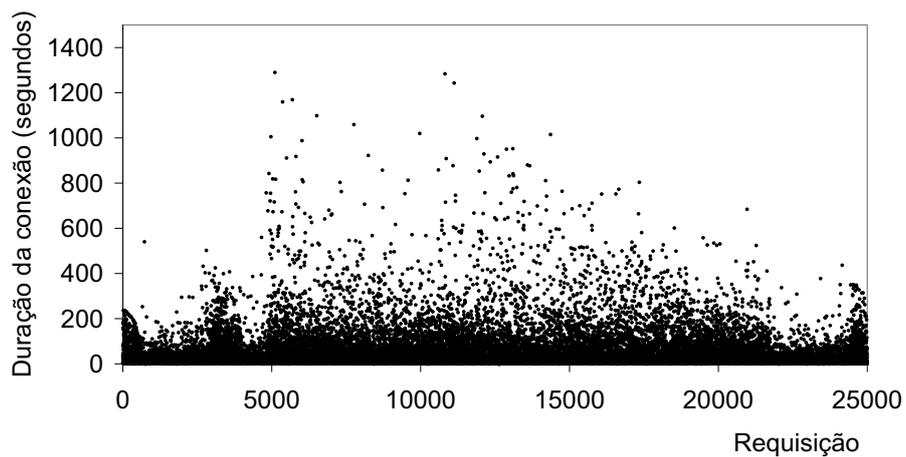


Figura 6. Duração da conexão ao longo de todos os dias em estudo para a Cidade2

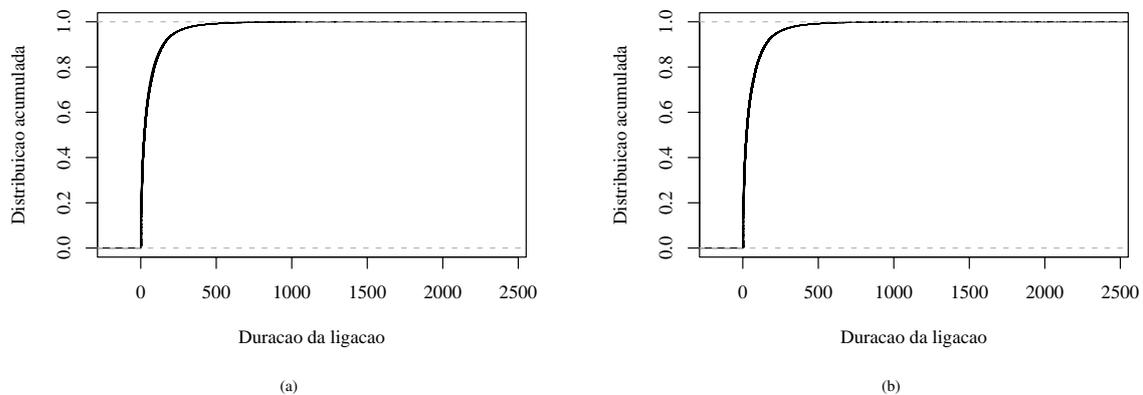


Figura 7. Distribuição acumulada para o tempo de conexão comparado com a distribuição de Weibull (linha pontilhada) em (a) Cidade1 e (b) Cidade2.

de conexão foi sempre aderente à distribuição de Weibull⁴, mesmo se observado em intervalos de tempo pequenos (horas) ou grandes (dias).

A análise dos dados indica que o comportamento do tráfego observado dos dados coletados pode ser caracterizado pela distribuição exponencial para o intervalo entre conexões e pela distribuição de Weibull para tempo de conexão.

A caracterização do tempo de conexão é importante, mas para o estudo de desempenho do sistema a característica predominante é o fato do intervalo entre conexões poder ser modelado por uma distribuição exponencial, o que torna possível o uso do modelo M/G/m com perdas - ou seja, a probabilidade de perda de ligação do sistema depende do tempo de conexão apenas através de sua média e não através de sua distribuição de probabilidade.

6. Análise e Aplicação de Resultados

O estudo do intervalo entre chegadas e do tempo de conexão indica que é possível utilizar o modelo de filas M/G/m com perdas (ou *Erlang loss*) [Jain 1991], sendo possível utilizá-lo em todos os períodos, inclusive o de pico. Este resultado contrasta com os estudos realizados anteriormente por [Färber et al. 1998b] e [Färber et al. 1998a], onde o intervalo entre chegadas foi modelado utilizando-se distribuições de cauda longa, impedindo a aplicação do modelo *Erlang loss*. No entanto, estes estudos foram realizados com bases de dados relativamente pequenas e refletiram o comportamento de um grupo bastante específico de usuários (acesso via linhas discadas em universidades) [Färber et al. 1998b], [Novak et al. 2003] e [Schikora and Godfrey 2006]. Em [Novak et al. 2003], o autor percebeu que 6% dos usuários foram responsáveis por 50% dos acessos, o que possivelmente reflete uma característica bastante específica de um público universitário. Nossos resultados de processo de chegada se aproximam daqueles obtidos em [Iversen et al. 2000] na operadora dinamarquesa Tele Danmark, pois consideram uma base realmente grande de usuários, que permitem avaliar melhor grandes sistemas de usuários. No

⁴Função de densidade de probabilidade: $f(x) = \frac{bx^{b-1}}{a^b} e^{-(x/a)^b}$

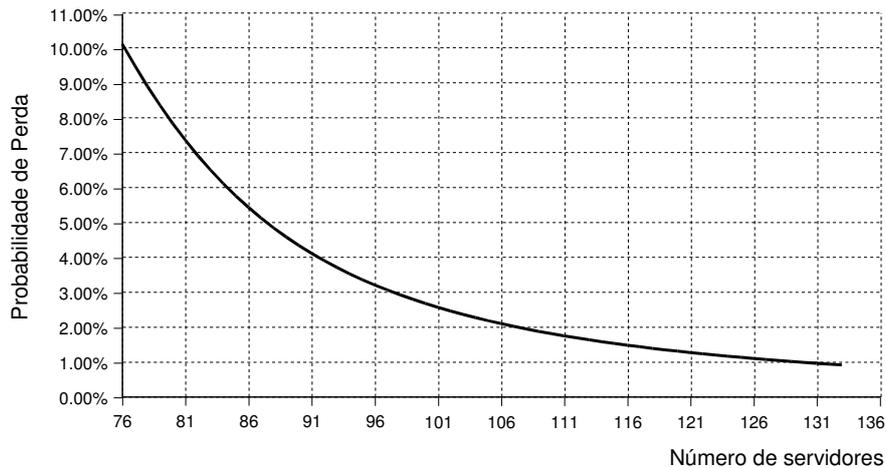


Figura 8. Probabilidade de perda em função do número de canais para o horário entre 19:00h e 23:00 (Cidade2)

entanto, os resultados dinamarqueses consideram acessos mistos, discados e ISDN (ainda significativos na Europa mas não no Brasil).

Para um sistema de filas M/G/m, sem espaço em buffer, o modelo desenvolvido por Erlang indica que a probabilidade de perda é dada pela Equação 1 a seguir:

$$p_m = \frac{(m\rho)^m / m!}{\sum_{j=1}^m [(m\rho)^j / j!]} \quad (1)$$

onde p_m representa a probabilidade de perda, m é o número de servidores, ρ é a intensidade de tráfego, dada por $\rho = \lambda / (m\mu)$, com μ representado a taxa de serviço e λ representando a taxa de chegadas.

Quando o número de servidores é muito grande é interessante o uso de expressões mais simples ao invés da expressão exata dada pela Equação 1. Neste trabalho foi utilizado o limite superior para p_m desenvolvido por [Harel 1988]:

$$UP1 = \frac{m(1 - \rho)^2 + 2\rho - (1 - \rho)\sqrt{4m\rho + m^2(1 - \rho)^2}}{-m\rho(1 - \rho) + 2\rho + \rho\sqrt{4m\rho + m^2(1 - \rho)^2}} \quad (2)$$

Como exemplo de aplicação do modelo proposto, o dimensionamento do sistema para o horário entre 00:00h a 01:00h para a Cidade2 indica que o tempo médio de conexão é de 61,98 minutos e o intervalo entre chegadas é de 6,25 segundos. Neste intervalo, para obter-se uma probabilidade de perda de no máximo 1% o sistema deverá reservar no mínimo 595 canais para este serviço (resultados que explicam a quantidade de perda de ligações observada neste horário para um dimensionamento baseado em voz [Cruz 2005]). Para a faixa de horário entre 19:00h e 23:00, observa-se um intervalo médio entre chegadas de 19,71 segundos e um tempo médio de conexão de 24,9 minutos, com o sistema necessitando de no mínimo 132 canais reservados. A Figura 8 ilustra o limite para probabilidade de perda obtido para o sistema no horário entre 19:00h e 23:00h em função do número de canais reservados. As perdas previstas pelo modelo estão de acordo com

as perdas observadas no sistema real - no entanto estes números não podem ser revelados porque a operadora em estudo não permite a sua divulgação.

7. Conclusões

Neste artigo foi realizada a modelagem do comportamento do usuário de conexões utilizando linhas discadas à Internet. Esta modalidade de acesso permanece importante e vem apresentando crescimento no número de usuários, a despeito do crescimento da modalidade de acesso em banda larga.

Foram analisados dados obtidos de uma operadora de telecomunicações de alcance nacional com tráfego de duas grandes cidades brasileiras. Os resultados mostram que o intervalo entre chegadas de requisições de conexão pode ser modelado utilizando-se a distribuição exponencial para determinadas faixas de horário e dias de semana especialmente selecionadas considerando-se como critério a estabilidade da série no intervalo. Mostrou-se também que a duração da conexão é melhor modelada pela distribuição de Weibull. Como consequência, é possível utilizar o modelo M/G/m com perdas (Erlang Loss) para dimensionar o sistema.

Como o sistema *dial-up* possui características diversas da ligação para conferência de voz em termos de sua duração média, recomenda-se que o dimensionamento seja feito em separado, resultando em um determinado número de canais para suportar o sistema *dial-up* que deve ser somado ao número de canais necessários para o sistema de comutação de voz.

Em [Novak et al. 2003] mostrou-se que um melhor dimensionamento do número de modems é mais eficiente do que a implementação de regras de negócio (por exemplo, limitação de tempo de sessões) para minimizar perdas por bloqueio. Logo, a importância de modelos mais adequados para dimensionamento de capacidade para sistemas de conexão por linha discada justifica plenamente o estudo aqui apresentado. Outra aplicação importante do modelo apresentado é a investigação de formação de preços para estes sistemas, uma questão bastante atual [Cheng et al. 2006] e [Patek and Campos-Nanez 2000].

Segundo a Agência Nacional de Telecomunicações, até 31 de julho de 2007, os usuários poderão optar por realizar a tarifação das chamadas locais por minutos de utilização. No entanto, não são alterados os horários de tarifa reduzida (de segunda a sexta-feira, entre 0 e 6 horas; aos sábados, da 0 às 6 e das 14 às 24 horas; aos domingos e feriados nacionais), quando a tarifação será realizada por completamento da chamada e não por tempo de utilização. Desta forma, as características do sistema de conexão discada identificadas por este trabalho continuam válidas e os resultados apresentados poderão ser utilizados para o dimensionamento de sistemas semelhantes no futuro.

8. Agradecimentos

Os autores agradecem a GVT pela disponibilização dos dados utilizados no estudo realizado. Com telefonia convencional a GVT está presente nas maiores cidades das regiões Sul, Centro-Oeste e Norte do país. Na Grande São Paulo, Rio de Janeiro e em Belo Horizonte atende exclusivamente com soluções corporativas. Seu provedor de internet grátis, o POP, está entre os provedores mais utilizados no país.

Referências

- Cheng, H. K., Dogan, K., and Elnicki, R. A. (2006). Pricing and capacity decisions for non-profit internet service providers. *Inf. Tech. and Management*, 7(2):91–107.
- Cruz, M. (2005). Technical Report - CPGEI/UTFPR.
- Ferrari, A. M. (2005). *Telecomunicações - Evolução e Revolução*. Editora Érica, 9 edition.
- Färber, J., Bodamer, S., and Charzinski, J. (1998a). Evaluation of dial-up behaviour of internet users. In *Proceedings of VDE-Kongress*, Stuttgart, German.
- Färber, J., Bodamer, S., and Charzinski, J. (1998b). Measurement and modelling of internet traffic at access networks. In *Proceedings of the EUNICE '98 Open European Summer School on Network Management and Operation*, pages 196–203, Munich, German.
- Harel, A. (1988). Sharp bounds and simple approximations for the erlang delay and loss formulas. *Manage. Sci.*, 34(8):959–972.
- Ibope (2006). Ibope/NetRatings. <http://www.ibope.com.br>, 30/Nov/2006.
- ITU-T (1972). Recommendation G.703. Physical/electrical characteristics of hierarchical digital interfaces. ITU 1972.
- ITU-T (1993). Recommendation G.803. Architectures of transport networks based on the synchronous digital hierarchy. ITU 1993.
- Iversen, V. B. (2005). *Teletraffic engineering handbook*.
- Iversen, V. B., Glenstrup, A. J., and Rasmussen, J. (2000). Internet dial-up traffic modelling. In *Proceedings of the NTS-15: 15th Nordic Teletraffic Seminar*, pages 196–203, Lund, Sweden.
- Jain, R. (1991). *The art of computer systems performance analysis: techniques for experimental design, measurement, simulation and modeling*. John Wiley & Sons.
- Naldi, M. (2002). Internet access traffic sharing in a multi-operator environment. *Elsevier Computer Networks*, 38:809–824.
- News, F. (2006). Reports of death of dial-up internet greatly exaggerated. <http://www.foxnews.com/story/0,2933,200606,00.html>, 30/Nov/2006].
- Novak, D. C., Rowland, D., and DaSilva, L. (2003). Modeling dialup internet access: An examination of user-to-modem ratios, blocking probability, and capacity planning in a modem pool. *Computers & Operations Research*, 30(13):1959–1976.
- Patek, S. and Campos-Nanez, E. (2000). Pricing of dialup services: an example of congestion-dependent. In *Proceedings of the 39th IEEE Conference on Decision and Control*, volume 3, pages 2296–2301.
- Schikora, P. F. and Godfrey, M. R. (2006). Connect time limits and customer service levels in dial-up modem pools. *Journal of Network and Systems Management*, 14(2).
- Vannucci, D., Kennedy, I., and Barker, M. (2003). Impact of tariff on dial-up internet traffic: modelling the subscriber response as a dynamic system. In *Proceedings of the ITC18 Workshop for Developing Countries*, Berlin.