



LARISSA PIOVEZANI

**“SIMULAÇÃO DE SERVIÇOS MULTIMÍDIA
EM REDES DE ACESSO DE BANDA LARGA MÓVEL”**

***“MULTIMEDIA SERVICES SIMULATION ON MOBILE
BROADBAND NETWORKS”***

Limeira, 2013



**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE TECNOLOGIA
MESTRADO EM TECNOLOGIA**

LARISSA PIOVEZANI

**“SIMULAÇÃO DE SERVIÇOS MULTIMÍDIA
EM REDES DE ACESSO DE BANDA LARGA MÓVEL”**

***“MULTIMEDIA SERVICES SIMULATION ON MOBILE
BROADBAND NETWORKS”***

Dissertação apresentada à Faculdade de Tecnologia de Limeira, da Universidade Estadual de Campinas, para obtenção do Título de Mestra em Tecnologia.

Área de Concentração: Tecnologia e Inovação

Orientador: Prof. Dr. Varese Salvador Timóteo.

Limeira, 2013

FICHA CATALOGRÁFICA ELABORADA POR VANESSA EVELYN COSTA CRB-8/8295
BIBLIOTECA UNIFICADA FT/CTL
UNICAMP

Piovezani, Larissa, 1987-
P659s Simulação de serviços multimídia em redes de acesso
de banda larga móvel / Larissa Piovezani. -
Limeira, SP : [s.n.], 2013.

Orientador: Varese Salvador Timóteo.
Dissertação (mestrado) – Universidade Estadual de
Campinas, Faculdade de Tecnologia.

1. Telecomunicações – tráfego. 2. Sistemas multimídia.
3. Simulação (Computadores). I. Timóteo, Varese Salvador.
II. Universidade Estadual de Campinas. Faculdade de
Tecnologia. III. Título.

Informações para Biblioteca Digital

Título em inglês: Simulation of multimedia services network access mobile broadband.

Palavras-chave em inglês (Keywords):

- 1- Telecommunication - traffic
- 2- Multimedia systems
- 3- Simulation (Computers)

Área de concentração: Tecnologia e Inovação

Titulação: Mestra em Tecnologia

Banca examinadora: Edson Luiz Ursini, Luiz Henrique Bonani do Nascimento.

Data da Defesa: 27-02-2013

Programa de Pós-Graduação em Tecnologia

DISSERTAÇÃO DE MESTRADO EM TECNOLOGIA
ÁREA DE CONCENTRAÇÃO: TECNOLOGIA E INOVAÇÃO

Simulação de serviços multimídia em redes de acesso de banda larga móvel

Autor: Larissa Piovezani

A Banca Examinadora composta pelos membros abaixo aprovou esta Dissertação:



Prof. Dr. Varesé Salvador Timóteo - Presidente
FT/UNICAMP



Prof. Dr. Edson Luiz Ursini
FT/UNICAMP



Prof. Dr. Luiz Henrique Bonari
CESC/UFABC

À minha querida Lara, que independentemente de onde, estaremos sempre juntas.

AGRADECIMENTOS

Primeiramente ao meu orientador, Professor Doutor Varese Salvador Timóteo, cujo apoio foi de enorme importância nas decisões e realizações dos percursos desta dissertação.

Ao meu noivo e eterno companheiro, Marco M. Chiara, pela paciência, apoio e persistência.

À toda minha família, por ter me dado a oportunidade de chegar até aqui. À minha mãe pelo apoio nas escolhas profissionais.

À Milene Rosa de Almeida, pela catalogação e sugestões.

À CAPES, pela bolsa de estudo, por proporcionar condições para que essa dissertação fosse realizada e à UNICAMP, que sem ela nada seria possível.

RESUMO

Este projeto consiste no desenvolvimento de um modelo de simulação para serviços multimídia em redes móveis de terceira geração. Inicialmente foram simulados serviços de voz com diferentes cenários de tráfego e posteriormente serviços de dados e vídeo. Para o trabalho foi considerado que o núcleo da rede atende à demanda de acesso, de modo que é necessário modelar apenas a interação entre os terminais móveis e a rede de acesso por rádio. A validação do modelo de simulação foi feita tentando descrever o acesso em redes reais, por exemplo, usando como referência valores estimados para o número de tentativas de chamadas no cenário de pior caso (*Busy Hour Call Attempts*). As simulações foram realizadas por eventos discretos, baseadas em variáveis aleatórias e métodos de Monte Carlo. O software ARENA, projetado exatamente para este tipo de modelo, foi utilizado nas simulações, já que o sistema em questão tem sua operação baseada em serviços cujos intervalos de requisição e durações podem ser representadas por distribuições conhecidas.

Palavras-chave: Telecomunicações - tráfego, Sistemas multimídia, Simulação (Computadores).

ABSTRACT

This project develops a simulation model for multimedia services of the third generation mobile network. Initially it was simulated voice services with different traffic scenarios and later video and data services. For this assignment was considered that the core of the network achieves the access demands, thus it is only necessary to model the interaction between the mobile terminal and the radio access network. The validation of the simulation model was done in order to describe the real network access, e.g., estimated values were used as reference for the number of call attempts in the worst case scenario (Busy Hour Call Attempts). The simulations were carried out by discrete events based on random variables and Monte Carlo methods. The ARENA software, designed exactly for this type of model, was used in the simulations because this system has its operation based on services whose request intervals and durations of the system can be represented by known distributions.

Keywords: Telecommunication - traffic, Multimedia systems, Simulation (Computers).

Lista de Figuras

Figura 1 - Representação simplificada de uma rede 3G.....	12
Figura 2 - Representação pictórica da evolução das redes móveis.....	14
Figura 3 - Representação simplificada do planejamento de frequências em macro células e micro células de uma rede celular.....	15
Figura 4 - Representação das transferências entre células do mesmo tipo (horizontal) e de tipos diferentes (vertical).....	16
Figura 5 - Evolução dos protocolos para redes móveis. Fonte: TELECO.	25
Figura 6 - Tipos e classes de tráfego, serviços e aplicações.....	27
Figura 7 - Representação do perfil típico de um tráfego de blocos.....	30
Figura 8 - Representação do perfil típico de um tráfego de transação.	31
Figura 9 - Representação do perfil típico de um tráfego <i>Stream</i>	32
Figura 10 - Quantidade de videoconferências em simulações feitas com o cenário considerado em função do intervalo entre as requisições de serviço para uma duração média fixa de 900s (15min).	41
Figura 11 - Quantidade de videoconferências em simulações feitas com o cenário considerado em função da duração média do serviço para um intervalo médio entre as requisições de serviço fixo de 3600s.....	42
Figura 12 - Quantidade de videoconferências em simulações feitas com o cenário considerado em função do número máximo de serviços simultâneos para uma duração média fixa de 30.000s e um intervalo médio entre as requisições fixo de 3600s.....	43
Figura 13 - Quantidade de serviços de voz em função do intervalo médio entre as tentativas de chamada para uma duração média de 900 s e um máximo de 10 serviços simultâneos.	45
Figura 14 - Quantidade de serviços de dados em função do intervalo médio entre as tentativas de envio de dados para uma duração média de 900 s e um máximo de 10 serviços simultâneos.....	46
Figura 15 - Quantidade de serviços de videoconferências em função do intervalo médio entre as tentativas de videoconferências para uma duração média de 900 s e um máximo de 10 serviços simultâneos.	46
Figura 16 - Quantidade de serviços de voz em função da duração média das chamadas de voz para um intervalo médio de 3600 s.....	47

Figura 17 - Quantidade de serviços de dados em função da duração do serviço para um intervalo médio de 3600 s.....	48
Figura 18 - Quantidade de videoconferências em função da duração do serviço.	48
Figura 19 - Quantidade de serviços de voz em função do número máximo de serviços simultâneos.	49
Figura 20 - Quantidade de serviços de dados em função do número máximo de serviços simultâneos.....	50
Figura 21 - Quantidade de videoconferências em função do número máximo de serviços simultâneos.....	50
Figura 22 - Quantidade de serviços de voz em simulações considerando a tarifação por nº de chamadas.	51
Figura 23 - Quantidade de serviços de dados em simulações considerando a tarifação por nº de chamadas.	51
Figura 24 - Quantidade de videoconferências em simulações considerando a tarifação por nº de chamadas.	52
Figura 25 - Quantidade de serviços de voz em simulações com intervalo de requisição fixa considerando a tarifação por nº de chamadas.....	52
Figura 26 - Quantidade de serviços de dados em simulações com intervalo de requisição fixa considerando a tarifação por nº de chamadas.	53
Figura 27 - Quantidade de videoconferências em simulações com intervalo de requisição fixa considerando a tarifação por nº de chamadas.	53
Figura 28 - Quantidade de serviços de voz em simulações com intervalo de requisição e duração média fixa considerando a tarifação por nº de chamadas.	54
Figura 29 - Quantidade de serviços de dados em simulações com intervalo de requisição e duração média fixa considerando a tarifação por nº de chamadas.	54
Figura 30 - Quantidade de videoconferências em simulações com intervalo de requisição e duração média fixa considerando a tarifação por nº de chamadas.	55
Figura 31 - Quantidade de serviços de voz em simulações considerando a tarifação fixa.....	56
Figura 32 - Quantidade de serviços de dados em simulações considerando a tarifação fixa.	57
Figura 33 - Quantidade de videoconferências em simulações considerando a tarifação fixa.	57

Figura 34 - Quantidade de serviços de voz em simulações com intervalo de requisição fixa considerando a tarifação fixa.....	58
Figura 35 - Quantidade de serviços de dados em simulações com intervalo de requisição fixa considerando a tarifação fixa.....	58
Figura 36 - Quantidade de videoconferências em simulações com intervalo de requisição fixa considerando a tarifação fixa.....	59
Figura 37 - Quantidade de serviços de voz em simulações com intervalo de requisição e duração média fixa considerando a tarifação fixa.....	59
Figura 38 - Quantidade de serviços de dados em simulações com intervalo de requisição e duração média fixa considerando a tarifação fixa.	60
Figura 39 - Quantidade de videoconferências em simulações com intervalo de requisição e duração média fixa considerando a tarifação fixa.	60
Figura 40 - Quantidade de videoconferências em simulações com intervalo de duração variável, <i>jitter</i> e latência.	62
Figura 41 - Quantidade de videoconferências em simulações com duração de serviço variável, <i>jitter</i> e latência.	63
Figura 42 - Quantidade de videoconferências em simulações variação do número máximo de serviços, <i>jitter</i> e latência.	63
Figura 43 - Quantidade de videoconferências em simulações considerando a tarifação por nº de chamadas, <i>jitter</i> e latência.	64
Figura 44 - Quantidade de videoconferências em simulações com intervalo de requisição fixa considerando a tarifação por nº de chamadas, <i>jitter</i> e latência.	64
Figura 45 - Quantidade de videoconferências em simulações com intervalo de requisição e duração média fixa considerando a tarifação por nº de chamadas, <i>jitter</i> e latência.	65
Figura 46 - Quantidade de videoconferências em simulações considerando a tarifação fixa, <i>jitter</i> e latência.....	65
Figura 47 - Quantidade de videoconferências em simulações com intervalo de requisição fixa considerando a tarifação fixa, <i>jitter</i> e latência.....	66
Figura 48 - Quantidade de videoconferências em simulações com intervalo de requisição e duração média fixa considerando a tarifação fixa, <i>jitter</i> e latência.	66

Lista de Tabelas

Tabela 1 - Modulação e codificação adaptativa.	20
Tabela 2 - Características do HSPA.	21
Tabela 3 - Taxas de pico de vazão do HSUPA.....	22
Tabela 4 - Serviços multimídia e suas respectivas classes de tráfego.	32
Tabela 5 - Resultado para dez rodadas da simulação para um canal de 4 Mbps, tempo total de simulação de 100.000 segundos e tempo de <i>warm-up</i> de 2000 segundos.....	40
Tabela 6 - Melhores taxas de tempo para 1 única perda de serviço.	61

Lista de Abreviaturas

BHCA – Busy Hour Call Attempts

CN – Core Network

GGSN – Gateway GPRS Support Node

GMSC – GPRS Mobile Switching Center

GSM – Global System for Mobile communications

HSDPA – High-Speed Downlink Packet Access

HSS – Home Subscriber Service

IS-95 - Interim Standard 95 (Primeiro padrão digital baseado em CDMA)

MGW – Media Gateway

MSC – Mobile-Switching Center

OTS - One Tunnel Solution (Encapsulamento)

OVSF - Orthogonal variable spreading factor

RAN – Radio Access Network

RBS – Radio Base Station

RNC – Radio Network Controller

SGSN – Serving GPRS Support Node

SMP – Personal Service Mobile

UE – User Equipment

UMTS – Universal Mobile Telecommunications System

WCDMA – Wide-band Code Division Multiple Access

Sumário

Lista de Figuras	viii
Lista de Tabelas	xi
Lista de Abreviaturas.....	xii
1 INTRODUÇÃO.....	10
2 REDES MÓVEIS	14
2.1 REDES DE 1ª GERAÇÃO (1G)	16
2.2 REDES DE 2ª GERAÇÃO (2G)	17
2.3 REDES DE 3ª GERAÇÃO (3G)	19
2.4 REDES DE 4ª GERAÇÃO (4G)	24
3 MODELOS DE TRÁFEGO	27
3.1 MODELAGENS DE COMPONENTES DE TRÁFEGO MULTIMÍDIA	29
3.1.1 Tráfego de Bloco	30
3.1.2 Tráfego de Transação	30
3.1.3 Tráfego <i>Stream</i>	31
3.1.4 Identificação da categoria de um tráfego multimídia	32
3.2 Funções de distribuição	33
3.2.1 Distribuição exponencial	34
3.2.2 Distribuição de Weibull (Cauda Longa).....	34
3.2.3 Distribuição de Pareto	35
3.2.4 Distribuição Lognormal.....	36
3.2.5 Distribuição discreta de Poisson.....	36
3.3 CONCEITOS BÁSICOS DA MODELAGEM COM O ARENA	37
4 RESULTADOS	44

5	CONCLUSÕES E PERSPECTIVAS.....	67
	REFERÊNCIAS	69

1 INTRODUÇÃO

O surgimento de sistemas de comunicação de banda larga móvel e a diminuição no preço de componentes para computação têm permitido a utilização de aplicações multimídia cada vez mais complexas. Os padrões de tráfego originados de aplicações multimídia tendem a ser mais complicados do que os padrões gerados por fontes somente de voz e dados. Assim, modelos de tráfego são muito importantes para uma avaliação precisa do desempenho de algoritmos de alocação de recursos em redes de acesso por rádio.

Quando o tráfego de dados estava emergindo em sistemas de comunicações móveis, a modelagem de fontes de tráfego para dados em redes sem fio já era importante [1]. Mais recentemente, diversos trabalhos têm abordado o problema de modelagem de tráfego multimídia. Golaup and Aghvami propuseram um ambiente de modelagem de tráfego multimídia para realizar estudos de avaliação de desempenho de redes WCDMA baseados em simulações [2]. Dentro do esquema proposto, o tráfego gerado por uma aplicação multimídia é dissociado em um conjunto de componentes básicos e um modelo de tráfego é deduzido para cada componente. Para ilustrar a utilidade do ambiente proposto, foi realizado um estudo de avaliação do desempenho de algoritmos de agendamento para o acesso de pacotes em alta velocidade no canal direto (HSDPA) em redes de terceira geração (3G).

Outra abordagem baseada em simulação para a modelagem de tráfego multimídia foi proposta para investigar o problema de habilitar a qualidade de serviço (QoS) para o tráfego multimídia na porta de entrada de comutadores de alta performance do tipo *input-queued* [3]. Um método alternativo para a modelagem de tráfego multimídia também foi utilizado para propor um controle de admissão de chamadas baseado em QoS para sistemas de comunicação sem fio baseados em acesso múltiplo por divisão de código (CDMA), que suporta a transmissão de diversos tipos de informação [4].

Um novo modelo para o tráfego de vídeo MPEG-4 multiplexado originado por *streams* de videoconferência foi proposto por Lazaris et al. [5], em que foi investigada a possibilidade desse

tipo de tráfego de vídeo ser modelado usando distribuições bem conhecidas. Nesse caso, eles mostraram que uma distribuição Pearson do tipo V é o melhor ajuste entre as distribuições consideradas [5].

Nos últimos anos houve um aumento enorme no número de redes móveis de banda larga disponíveis, criando um ambiente perfeito para o oferecimento de serviços multimídia por meio de aplicações para dispositivos móveis. Além disso, os planos atrativos de voz e dados oferecidos pelas operadoras de redes móveis de terceira geração estão fazendo o número de assinantes aumentarem. Como consequência, o perfil do tráfego nas redes, que antes era apenas oriundo dos serviços de voz, é alterado significativamente com a introdução de outros serviços com características e demandas diferentes.

Um modelo de simulação pode ser muito útil para fazer previsões para o tráfego da rede em determinadas condições para guiar o processo de dimensionamento da rede de acesso e para manter o desempenho dentro dos níveis estabelecidos pelas agências reguladoras de serviços de telecomunicações. A motivação para analisar diferentes cenários de tráfego é a influência do tipo de tarifação no tráfego e no desempenho da rede. Três tipos de tarifação mais comuns utilizados por operadoras de redes móveis serão considerados: tarifação pelo tempo de utilização do canal de rádio (conexão), tarifação pelo número de chamadas (independentemente da duração de cada conexão) e tarifação fixa (independentemente da duração e do número de conexões).

No primeiro caso, os intervalos entre as requisições de serviço tendem a ser maiores e as durações dos serviços menores. No segundo caso, as durações aumentam substancialmente e os intervalos entre as requisições devem diminuir em relação ao primeiro. No terceiro, os intervalos entre as requisições devem diminuir em relação ao segundo caso e diminuir ainda mais em relação ao primeiro, diminuindo também em relação ao segundo. As durações tendem a manter uma proporção similar. O modelo de simulação deve ser capaz de mostrar o colapso no desempenho da rede ao alterar o tráfego para o mesmo dimensionamento, e de mostrar como a rede deve ser redimensionada para que o desempenho seja mantido mesmo com o tráfego alterado. As redes de terceira geração foram concebidas principalmente a partir do requisito de altas taxas de transmissão. O caminho predominante escolhido pela maior parte das empresas que

desenvolvem infraestrutura de redes foi o de utilizar o WCDMA na interface aérea e melhorar o núcleo das redes GSM de segunda geração.

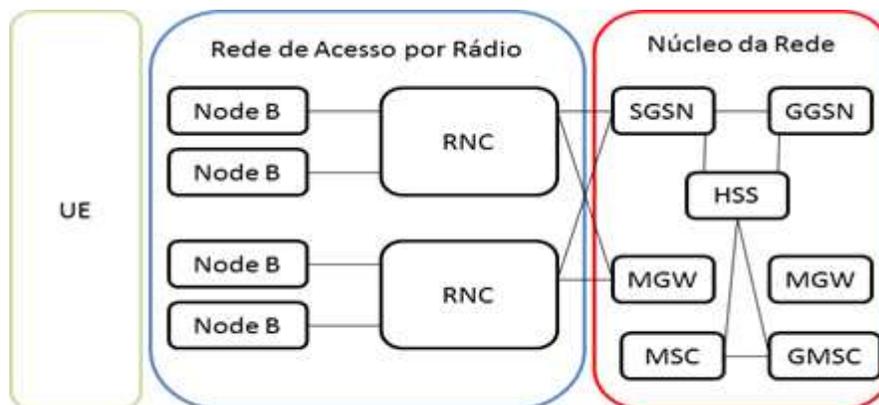


Figura 1 - Representação simplificada de uma rede 3G.

A Figura 1 mostra um esboço da arquitetura de uma rede móvel de terceira geração UMTS [6], em que a maioria dos protocolos de comunicação entre os diversos subsistemas são padronizados. Uma exceção é a comunicação entre as controladoras (RNC) e os nós B (RBS) conectadas a ela, que normalmente é feito utilizando protocolos proprietários.

O planejamento de uma rede desse tipo é tipicamente composto de duas partes: o planejamento do núcleo da rede e o planejamento da rede de acesso por rádio. Idealmente, os dimensionamentos do núcleo e do acesso devem ser feitos de tal forma que não haja nenhum gargalo no tráfego de informação. Além disso, deve-se considerar um cenário de tráfego estimado pelo pior caso, ou seja, no período de maior movimento. Desta forma o indicador BHCA pode ser utilizado para fornecer as tentativas de acesso à rede em um período de alto tráfego.

Neste trabalho considera-se que o núcleo da rede seja superdimensionado, ou seja, toda demanda da rede de acesso é atendida. A ausência de recusa de serviços por parte do núcleo da rede é uma hipótese que simplificará o modelo de simulação, que pode incluir apenas a interação entre a rede de acesso por rádio e os dispositivos móveis.

No segundo Capítulo, uma introdução sobre redes móveis é apresentada e sua evolução técnica e funcionamento são discutidos. O Capítulo 3 trata dos modelos de tráfego, das distribuições mais utilizadas e uma introdução sobre o *software* ARENA. Os resultados obtidos são apresentados no Capítulo 4. As possíveis extensões do trabalho e discussão dos resultados são concluídas no Capítulo 5, seguido das referências bibliográficas.

2 REDES MÓVEIS

Este capítulo tem como referências principais [11], [12] e [22] e apresenta a evolução das redes móveis, que pode ser representada pictoricamente como na Figura 2.

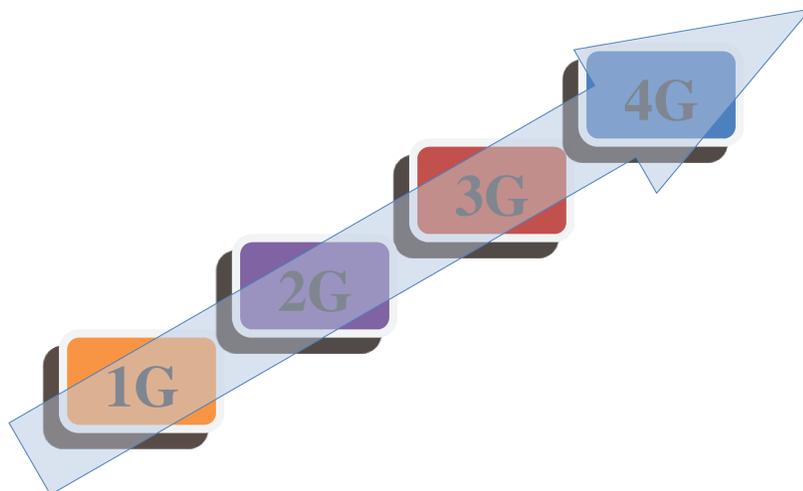


Figura 2 - Representação pictórica da evolução das redes móveis.

Um sistema de comunicação móvel tem como objetivo principal prover serviços de comunicação a usuários que podem se deslocar na área de cobertura da rede. Essa área é normalmente dividida em células para que as frequências possam ser repetidas, assim aumentando sua capacidade. A comunicação aos terminais móveis (assinantes) é assegurada por uma estação base em cada célula, que em geral é dividida em três setores (para aumentar a capacidade). Alguns problemas comuns em redes celulares são a interferência co-canal e a transferência de ligações de uma célula para a outra, sem que haja interrupção no serviço, enquanto o assinante se desloca dentro da rede e passa da área de cobertura de uma célula para outra (*handover*).

Dividindo a área em células é possível reutilizar as frequências em células afastadas e a criação de micro células permite aumentar a capacidade da rede. É possível definir as células com uma dimensão variável ajustando a potência de transmissão e os tipos de cobertura podem se

adequar aos diferentes requisitos de tráfego. A Figura 3 ilustra simplificada o planejamento de frequências em uma macro célula e em uma micro célula.

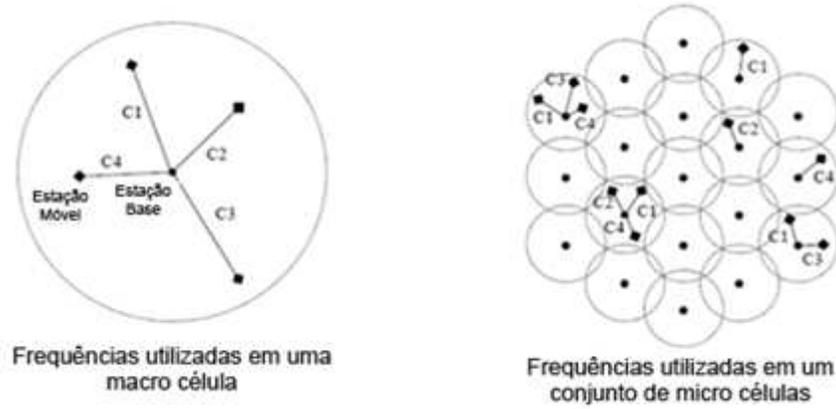


Figura 3 - Representação simplificada do planejamento de frequências em macro células e micro células de uma rede celular.

Uma rede celular é composta por unidades móveis (UM, no caso 3G é chamada UE), Estações Rádio Base (ERBs), transmissores, receptores, torre e antenas, e unidades de controle (BSCs), que realizam todo o controle da interface aérea - como controle de potência, controle de transferências (*handover*) e controle dos enlaces de rádio.

A Central de Comutação e Controle (CCC) é uma central telefônica para redes celulares que tem como funções os serviços de processamento de chamadas, normalmente executadas por uma central telefônica comum, o registro das unidades móveis, o gerenciamento e coordenação dos processos de *handoff*, mudança de células, controle geral das chamadas e a geração dos registros das chamadas, fundamental para a tarifação, e o *paging*, que é o processo de atualização de informações de localização. A Figura 4 mostra que pode haver transferências horizontais entre células de um mesmo tipo de rede (de uma célula 3G para outra célula 3G, por exemplo) e transferências verticais entre células de redes diferentes (de uma célula 3G para uma célula 2G, por exemplo).

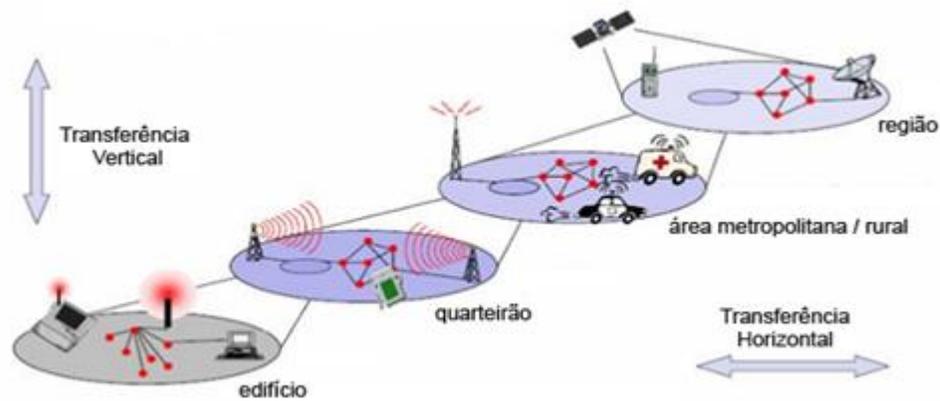


Figura 4 - Representação das transferências entre células do mesmo tipo (horizontal) e de tipos diferentes (vertical).

2.1 REDES DE 1ª GERAÇÃO (1G)

Desenvolvida no início dos anos 80 [13], as redes móveis de 1º geração eram baseadas em tecnologia analógica, como FDMA/FDD e FM analógica. Esse sistema permitia que um assinante pudesse se movimentar e manter um tráfego de voz em qualquer lugar dentro da área de serviço do operador da rede móvel. Essa cobertura contínua fez com que houvesse necessidade de criar um mecanismo que permitisse transferir uma chamada entre áreas de cobertura (células) pertencentes a estações base diferentes, mecanismo esse denominado de *handover* (ou *handoff*). Era um serviço de voz sem fios baseado em macro células.

Nas redes móveis 1G o único serviço existente era o de voz e cada país criou o seu próprio sistema, levando à existência de sistemas que só funcionavam dentro dos limites de cada país. Esses sistemas foram substituídos pelos sistemas de segunda geração durante os anos 90.

Os primeiros sistemas possuíam células grandes providas por transmissores de alta potência omnidirecionais e tinham sua capacidade e segurança bem limitadas [19]. Essas redes foram baseadas principalmente nos seguintes padrões:

- AMPS (*Advanced Mobile Phone System*) surgiu em 1976 nos Estados Unidos e constitui o primeiro padrão de rede celular. Utilizado principalmente na Rússia e na Ásia, essa rede analógica de primeira geração possuía fracos mecanismos de segurança que tornavam possível o acesso à informação que trafegava nas redes.
- TAC (*Total Access Communication System*) é a versão europeia do modelo AMPS. Utilizando a banda de frequência de 900 MHz, esse sistema foi utilizado principalmente na Inglaterra e em seguida na Ásia (Hong-Kong e Japão).
- ETACS (*Extended Total Access Communication System*) é uma versão aprimorada do padrão TAC desenvolvida no Reino Unido que utiliza um número maior de canais de comunicação.

As redes celulares de primeira geração tornaram-se obsoletas com o aparecimento de uma segunda geração inteiramente digital.

2.2 REDES DE 2ª GERAÇÃO (2G)

A segunda geração de redes móveis (2G) foi marcada pela ruptura com a primeira geração de telefones celulares devido à passagem da transmissão analógica para a transmissão digital. Os principais padrões para a comunicação em redes móveis 2G são os seguintes:

- TDMA (*Time Division Multiple Access*) - utilizando uma técnica de divisão temporal dos canais de comunicação, a fim de aumentar o número de usuários que a rede pode comportar simultaneamente. A tecnologia TDMA é principalmente utilizada no continente americano, na Nova Zelândia e na Ásia do Pacífico. Cada canal de 30 kHz é dividido em três compartilhamentos.

- GSM (*Global System for Mobile communications*) - o padrão mais utilizado na Europa no fim do século XX e também utilizado nos Estados Unidos. Esse padrão utiliza as bandas de frequências 900 MHz e 1800 MHz na Europa. Nos Estados Unidos, por outro lado, a banda de frequência utilizada é a banda 1900 MHz, já que a banda de 1800 MHz é reservada para uso militar. Assim, os terminais móveis que podem funcionar na Europa e nos Estados Unidos devem ser do tipo tri-banda. Na verdade, o GSM é um TDMA de oito sub-canais em um canal de 200 kHz.
- CDMA (*Code Division Multiple Access*) - utilizando uma técnica de espalhamento espectral, os assinantes são separados por códigos pseudo-ortogonais. Cada canal de 1,25 MHz no sistema CDMA é obtido a partir de 42 canais de 30 kHz.

Com as redes 2G é possível transmitir voz e um pequeno volume de dados, por exemplo, mensagens de textos (SMS – *Short Message Service*) ou mensagens multimídia (MMS – *Multimedia Message Service*). A norma GSM permite uma taxa máxima de 9,6 kbps.

Extensões da norma GSM foram criadas a fim de melhorar a taxa de transmissão de dados enquanto as redes de terceira geração eram desenvolvidas. É principalmente o caso do padrão GPRS (*General Packet Radio System*), que permite obter taxas teóricas de aproximadamente 114 kbit/s (mais próximas de 40 kbit/s na prática). Essa tecnologia foi denominada 2.5G.

A norma EDGE (*Enhanced Data Rates for GSM Evolution*), multiplica por quatro a taxa obtida no GPRS, chegando a uma taxa teórica de 384 Kbps, abrindo assim a porta para as aplicações multimídia. Essa norma permite atingir taxas máximas teóricas de 473 kbit/s, mas foi limitada a fim de adequar-se às especificações IMT-2000 (*Internacional Mobile Telecommunications - 2000*) do ITU (*International Telecommunications Union*).

2.3 REDES DE 3ª GERAÇÃO (3G)

As redes 3G foram inseridas no Brasil, pelas operadoras de SMP em 2008 [22]. Essa geração utiliza uma tecnologia conhecida como *High-Speed Downlink Packet Access* (HSDPA), cujo objetivo é uma banda mais rápida no *dowlink*. O UMTS (*Universal Mobile Telecommunication System*) utiliza o dispositivo *Rake Receiver*, que já existia nas primeiras versões do CDMA para que o *soft handoff* (conexão simultânea com duas ou mais células ou setores) pudesse ser usado para melhorar a qualidade de voz. Contudo, existem algumas diferenças entre os padrões que devem ser observadas:

- A necessidade de sincronismo para interoperação entre as estações base, ERBS, no IS-95 é opcional no UMTS. Com isso, os nós-B dos sistemas UMTS não utilizam GPS como é necessário nas ERBS dos sistemas baseados em IS-95.
- Para a separação dos assinantes no IS-95 são utilizados os códigos de Walsh, enquanto no UMTS os códigos utilizados são *Orthogonal Variable Spreading Factor* (OVSF). Esses códigos têm algumas características, tais como:
 - O sistema 3GPP acomoda canais com valores diferentes de *throughput* utilizando códigos OVSF com *Spreading Factors* (SF) diferentes,
 - Canais para taxas altas devem usar SFs curtos, enquanto que canais para taxas baixas devem usar SFs mais longos.

O WCDMA é um sistema *Direct Sequence Spread Spectrum* (DSSS) que utiliza a técnica de múltiplos códigos para acesso ao meio (códigos OVSF combinados com *Scrambling Codes*). O *Transmission Time Interval* (TTI), que representa o tamanho do quadro do sistema é de 10, 20, 40 ou 80 ms e latência, ou seja, a diferença de tempo entre o início de um evento e o momento em que seus efeitos tornam-se perceptíveis, para TTI de 10 ms, é de aproximadamente 150 ms. A

modulação utilizada é a *Quadrature Phase-Shift Keying* (QPSK) no *Downlink* e *Binary Phase-Shift Keying* (BPSK) no *Uplink*.

O HSDPA consiste em um conjunto de melhorias para permitir uma taxa de dados maior e atrasos menores e atinge um *throughput* médio de 800 kbit/s até 1,5 Mbit/s, com pico máximo de 3,6 Mbit/s até 14,4 Mbit/s. Algumas das implementações para construir o HSDPA são:

- Modulação e codificação adaptativa (*Adaptive Modulation and Code – AMC*), conforme a Tabela 1;
- TTI de dois milissegundos no *downlink*;
- Utilização de *Spread Factor* fixo e igual a 16 (SF=16);
- Algoritmo de agendamento executado pelo nó B;
- Utilização de modulação de ordem alta (16 QAM);
- Utilização apenas de *hard handover* (não ocorre *soft* ou *softer handover*);
- As retransmissões, caso sejam necessárias, agora podem ser realizadas pelo nó B, não necessitando mais da intervenção da RNC (*Radio Network Controller*).
- Utilização do HARQ (*Hybrid Automatic Request*);
- Latência de cerca de 100 milissegundos.

Tabela 1 - Modulação e codificação adaptativa.

Modulação	Taxa de Código Efetiva	Máxima Vazão (Mbit/s)
QPSK	$\frac{1}{4}$	1,8
QPSK	$\frac{2}{4}$	3,6
QPSK	$\frac{3}{4}$	5,3
16QAM	$\frac{2}{4}$	7,2

16QAM	$\frac{3}{4}$	10,7
-------	---------------	------

Fonte: HOLMA, H.; TOSAKALA, A. (Ed.), 2004. p. 321.

O *High-Speed Uplink Packet Access* (HSUPA) foi desenvolvido para aumentar a taxa de transmissão também no *Uplink*. A combinação do HSDPA e do HSUPA é chamada simplesmente de *High Speed Packet Access* (HSPA).

Uma escolha foi observar as técnicas aplicadas ao HSDPA e, se possível, aplicá-las também ao *uplink*. As técnicas a seguir são responsáveis pela velocidade de pico de 5,7 Mbit/s atingida no *uplink* e pela redução da latência (a Tabela 2 resume algumas das principais características e se são usadas pelo HSDPA e/ou HSUPA):

- Criação de um novo canal de transporte para o *uplink* chamado *Enhanced Dedicated Physical Channel* (E-DCH);
- Utilização do TTI de 2 ou 10 ms no *uplink*;
- *Fast Scheduling* executado no *Nó B*;
- Latência de aproximadamente 70 milissegundos.

Tabela 2 - Características do HSPA.

Funcionalidade	DCH	HSDPA (HS-DSCH)	HSUPA (E-DCH)
Fator de espalhamento variável	Sim	Não	Sim
<i>Fast Power control</i>	Sim	Não	Sim
Modulação Adaptada	Não	Sim	Não

<i>BTS based scheduling</i>	Não	Sim	Sim
Fast L1 HARQ	Não	Sim	Sim
<i>Soft handover</i>	Sim	Não	Sim
TTI [ms]	80,40,20,10	2	10 e 2

Fonte: HOLMA, H.; TOSAKALA, A. (Ed.), 2004, p. 63.

No HSUPA, diferentemente do que acontece no HSDPA, utiliza-se o *soft handover* e não se usa um fator de espalhamento fixo para atingir taxas de dados variadas (ver Tabela 3). Também não se usa modulação adaptativa porque nenhum esquema de modulação de ordem alta é suportado, sendo a modulação BPSK o único tipo suportado pelo HSUPA.

Tabela 3 - Taxas de pico de vazão do HSUPA.

HSUPA Categoria	<i>Codes x Spreading</i>	TTI	<i>Transport Block Size</i>	Taxa de Dados
1	1 x SF4	10	7296	0,73 Mbit/s
2	2 x SF4	10	14592	1,46 Mbit/s
2	2 x SF4	2	2919	1,46 Mbit/s
3	2 x SF4	10	14592	1,46 Mbit/s
4	2 x SF2	10	20000	2 Mbit/s
4	2 x SF2	2	5837	2,9 Mbit/s
5	2 x SF2	10	20000	2 Mbit/s

6	2 x SF2 + 2 x SF4	10	20000	2 Mbit/s
6	2 x SF2 + 2 x SF4	2	11520	5,76 Mbit/s

Fonte: EDGE, HSPA and LTE: The Mobile Broadband Advantage, 2007, p. 74.

Sistemas que utilizam *Orthogonal Frequency Division Multiple Access* – OFDMA têm atraído considerável atenção por meio de tecnologias tais como 3GPP LTE, WiMax e UMB. Contudo, o desempenho de sistemas CDMA pode equivaler a sistemas OFDMA com reduzidas larguras de banda. O objetivo do HSPA+ é explorar as tecnologias de rádio disponíveis para maximizar o desempenho dos sistemas baseados no CDMA.

Taxas de transmissão de até 42 Mbit/s podem ser suportadas pelo HSPA+. Essas taxas podem ser atingidas pela combinação da nova tecnologia de modulação de ordem alta (64-QAM no *downlink* e 16-QAM no *uplink*), em conjunto com a tecnologia de antenas 2X2 MIMO (*Multiple Input Multiple Output*).

Uma importante melhoria é o avanço dos receptores que podem implementar as seguintes esquemas:

- Diversidade na recepção – combinação de sinais provenientes de duas antenas diferentes no lado do receptor;
- Equalização de canal LMMSE (*Linear Minimum Mean Square Error*) – trabalha em conjunto com a diversidade de recepção para proporcionar um ganho ao sinal recebido pelo dispositivo móvel;
- Cancelamento de interferência – estimativa da interferência produzida pelo acesso múltiplo e pelo multipercurso e posterior subtração dessa interferência estimada.

Outra melhoria implementada é a Conectividade Contínua de Pacote. Ela é responsável por introduzir a transmissão descontínua (DTX) e recepção descontínua (DRX). O DTX permite ao UE (*User Equipment*) desligar os canais de controle reduzindo a interferência total no nó B, aumentando a capacidade do sistema e reduzindo o consumo de energia do móvel. Da mesma forma o DRX permite que o UE desligue o receptor no momento em que o nó B não está transmitindo. Isso permite redução do consumo de energia do dispositivo móvel.

A arquitetura da rede também é otimizada com o objetivo de reduzir o número de nós do sistema. A latência da rede pode ser minimizada com a introdução da solução de encapsulamento.

Foi desenvolvida, também, uma tecnologia chamada MIMO (*Multiple-Input Multiple-Output*), que implementa múltiplas antenas transmissoras e múltiplas antenas receptoras, geralmente combinadas com múltiplos rádios e múltiplos fluxos de dados paralelos. O uso mais comum do termo MIMO se aplica à multiplexação espacial. O transmissor envia diferentes fluxos de dados em cada antena. O MIMO explora o percurso múltiplo, reforçando os sinais que atravessam diferentes caminhos. Isso resulta em diversos caminhos de dados efetivamente operando em paralelo e, por meio de uma decodificação apropriada, em uma multiplicação do ganho e do *throughput*.

Com relação a sistemas convencionais de antenas 1x1, os sistemas MIMO 2x2 podem oferecer ganhos de *throughput* nas células de até 80%. O sistema HSPA+ utilizando MIMO 2x2, 64QAM no *downlink* e 16QAM no *uplink*, poderá atingir uma vazão de até 42 Mbit/s no *downlink* e até 11,5 Mbit/s no *uplink*. O HSPA+ também reduzirá a latência para abaixo dos 25 ms.

2.4 REDES DE 4ª GERAÇÃO (4G)

O dois padrão de 4G hoje é o LTE *Advanced (Long Term Evolution)*. Esta tecnologia, desenvolvida pela *Third Generation Partnership Project (3GPP)*, é uma evolução do HSPA e usa

novas técnicas de modulação, por isso consegue alcançar taxas de transmissão maiores no *downlink*. Na Figura 5 é mostrada a evolução das redes digitais, do 2G até o 4G. O LTE usa como modulação OFDM e antenas MIMO.

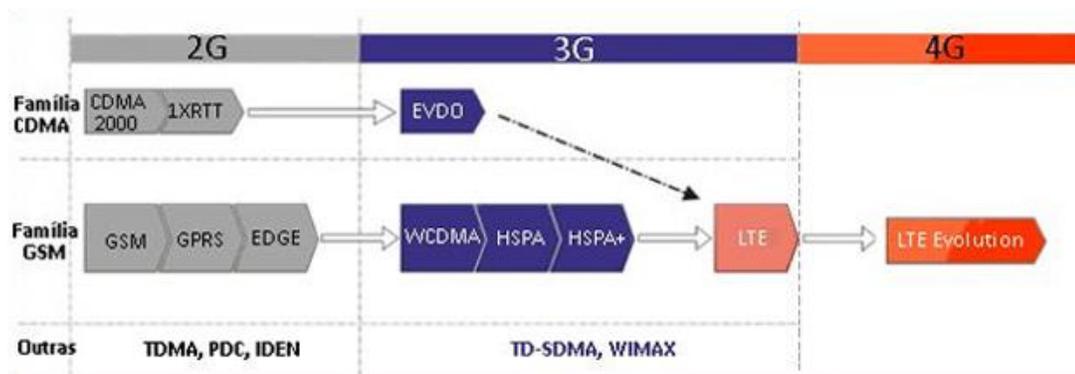


Figura 5 - Evolução dos protocolos para redes móveis. Fonte: TELECO.

A tecnologia LTE é uma evolução das redes 3G atuais. Poderemos usar TV interativa, *blogs* de vídeo móveis, jogos avançados e serviços profissionais, além de muitos outros recursos.

O LTE vai oferecer benefícios não apenas às operadoras, mas também aos consumidores. Em um primeiro momento, o preço do serviço será alto, pois a demanda da tecnologia será pequena. Conforme as pessoas forem adquirindo aparelhos compatíveis com a tecnologia, a tendência é que o preço diminua, assim como ocorre hoje com o 3G.

O 4G vai se aproximar da internet cabeada com relação às taxas de transmissão. A tecnologia LTE permite que taxas superiores a 200 Mbit/s sejam atingidas. A empresa Ericsson, por exemplo, conseguiu taxas acima de 150 Mbit/s. Essas taxas já atendem aos requisitos das redes de quarta geração exigidos pela ITU.

As tecnologias para redes de comunicações têm evoluído em uma significativa velocidade nos últimos anos. Na vanguarda dessa tendência estão situadas as redes móveis que, com a implementação plena dos requisitos das redes de terceira geração, podem passar a ser chamada também de redes de banda larga móvel.

A implantação efetiva de três gerações de sistemas celulares e a preparação para o surgimento de uma quarta geração suscitam a questão de como se dará a convivência entre as tecnologias mais novas e o legado deixado pelas tecnologias mais antigas. Um exemplo de como isso pode ocorrer são as redes das operadoras que possuem a tecnologia GSM e UMTS funcionando simultaneamente, nas quais é possível até mesmo o *handover* vertical entre os dois sistemas.

Outro aspecto interessante a ser observado na escala evolutiva do 3GPP é a futura transição do CDMA para o OFDMA, que ocorrerá quando com a implementação do LTE. Conforme já mencionado, o OFDMA possui melhor desempenho para bandas acima de 5 MHz. A redução da latência também é um aspecto importante que deve ser alcançada principalmente com a aplicação encapsulada, que implica na redução do número de nós das redes.

Analisando com cuidado as técnicas apresentadas na escala evolutiva do 3GPP, observamos a preocupação com dois aspectos complementares: de um lado a tentativa de aumentar a eficiência espectral com a adoção técnicas de modulação mais eficientes do ponto de vista do uso da largura de banda disponível (como o 64-QAM), e de outro a tentativa de minimizar a interferência com a adoção de técnicas como o MIMO, a diversidade de recepção e o cancelamento sucessivo de interferência. As primeiras tendem a aumentar a vulnerabilidade do receptor no caso de ocorrência de interferências e as segundas tentam minimizar o efeito da interferência percebida, viabilizando o uso mais eficiente do espectro de frequências disponível.

3 MODELOS DE TRÁFEGO

Quando surgiram as redes de computadores, o tráfego de informação era constituído basicamente de troca de pequenos arquivos e mensagens eletrônicas, com as características do tráfego elástico, vide a seguir. O crescimento extremamente rápido das redes de dados e o avanço das tecnologias na fabricação de processadores e outros componentes para computação impulsionou o surgimento de novos serviços e aplicações que incluíam diversos tipos de informação (multimídia). O tráfego predominante passou a ser de voz, vídeo e áudio *streaming*, constituindo o que chamamos de tráfego *stream*. Esse tipo de tráfego também é conhecido como tráfego inelástico, pois não se adapta à variações de atrasos (elasticidade), que é a principal característica do tráfego do tipo elástico. A denominação *Streaming* é proveniente da característica do fluxo de dados presentes no tipo de transmissão dos pacotes.

Tráfego elástico é a denominação utilizada para aplicações não interativas (como navegação *Web*, transferência de arquivos e *e-mail*), onde o ponto crítico está na integridade da informação e não no atraso da entrega dos pacotes, demandando apenas requisitos de vazão e de perda de pacotes. No entanto, com o surgimento de novos serviços que necessitam de prioridade na transmissão (voz sobre IP ou quando ambos competem pelo mesmo canal, por exemplo), pode ocorrer um atraso muito grande no envio de pacotes provenientes do tráfego elástico.

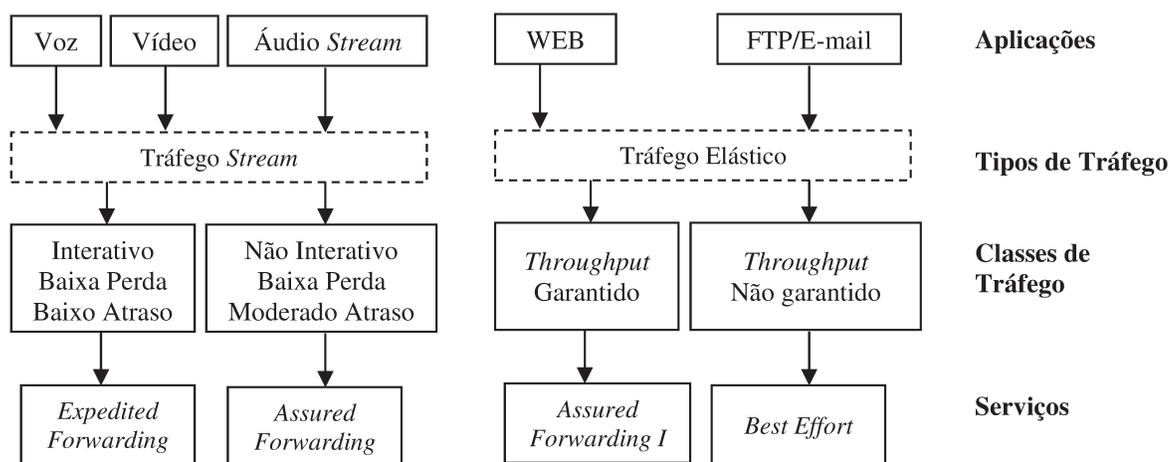


Figura 6 - Tipos e classes de tráfego, serviços e aplicações.

O tráfego *Stream* é uma denominação usual para aplicações interativas que são exigentes quanto ao atraso (aplicações de tempo real). Essas aplicações demandam garantia de vazão e atraso. É levada em consideração para seu dimensionamento a probabilidade de perda de pacotes e probabilidade de bloqueio. A Figura 6 ilustra os tipos de tráfego e exemplos de aplicações correspondentes. A principal característica do tráfego *stream* é a sensibilidade ao atraso, enquanto o tráfego elástico é sensível à perda.

As tecnologias de transmissão em redes multimídia têm proporcionado uma integração de vídeo, voz e dados. Grosso modo, os serviços podem ser divididos em dois tipos: *stream*, como áudio e vídeo e, os que não possuem necessidade de serem em tempo real e os denominados de elásticos, que são os serviços de dados em geral.

No contexto do protocolo IP, as aplicações em tempo real frequentemente são associadas com o protocolo de transporte UDP e as aplicações que não requerem tempo real são associadas com o protocolo TCP. Em aplicações de tempo real, a sequência dos pacotes é definida pela fonte. Os pacotes chegam aos seus destinos respeitando a sequência adotada pela fonte transmissora. Caso isso não ocorra cabe ao receptor ordenar novamente os pacotes. Algumas variações nos valores de *delay* e *jitter* são permitidas sem que haja grandes perdas se os limites são respeitados.

Em aplicações que não são em tempo real, as fontes não possuem uma taxa de pacotes definida. A transferência de dados tolera variação do atraso e do *jitter* desde que não haja perda de pacotes, que compromete completamente o serviço de dados. Com isso não temos uma qualidade de serviço em termos de comparação com aplicações em tempo real que não toleram atraso. É necessário suprir uma taxa de bits nominal. Esse tipo de aplicação baseada em TCP gera tráfego Elástico com as seguintes características:

- A taxa da fonte de tráfego varia de um pacote para outro de acordo com a janela de congestionamento do TCP;

- A taxa de *bits* da fonte de tráfego muda quando a fonte detecta a perda de um pacote e faz a sua retransmissão.

A maior diferença entre estas duas categorias é o comportamento da aplicação sobre as condições da rede. A taxa de *bits* dos aplicativos UDP é independente da rede e determinada pela fonte, enquanto que a taxa de *bits* das aplicações que usam TCP variam de uma rede para outra. Isso ocorre porque normalmente o protocolo UDP utiliza valores fixos, exigindo que haja uma taxa de banda disponível no mínimo igual ou superior à do codec que o serviço utiliza. Caso essa condição não seja satisfeita ocorrerá degradação da qualidade do serviço, mau funcionamento ou, dependendo da situação, inviabiliza completamente o serviço.

3.1 MODELAGENS DE COMPONENTES DE TRÁFEGO MULTIMÍDIA

Uma aplicação multimídia consiste na integração de dois ou mais tipos de informação em uma mesma sessão [4]. Um ambiente simples como a projeção de uma aula pode conter aplicações de áudio, vídeo e dados, sendo que cada aplicação utilize um tipo de tratamento diferente para a informação. Mesmo se duas aplicações têm os mesmos tipos de informação, a quantidade de dados gerada e as estatísticas do processo de geração de dados podem ser completamente diferentes. Por consequência, em vez de deduzir um modelo de tráfego para cada aplicação, é mais adequado identificar padrões de tráfego comuns a um grande número de aplicações multimídia.

O primeiro passo na modelagem de tráfego multimídia são os algoritmos que são necessários para gerar os padrões de tráfego de cada componente, tais como Bloco, Transação e *Stream*. Os algoritmos apenas fornecem os procedimentos para a geração de tráfego. A fim de gerar um padrão de tráfego específico de uma aplicação multimídia, os parâmetros do modelo devem ser especificados.

3.1.1 Tráfego de Bloco

Para definir uma fonte de tráfego de blocos é necessário especificar a função de distribuição e os parâmetros correspondentes para cada uma das seguintes variáveis:

- Número de blocos por sessão.
- Tamanho do bloco.
- Tempo de chegada dos blocos.

A fim de especificar o número de blocos por sessão, uma variação discreta é necessária. Isso pode ser feito usando a distribuição de Poisson. O perfil típico de um tráfego de blocos é mostrado na Figura 7.

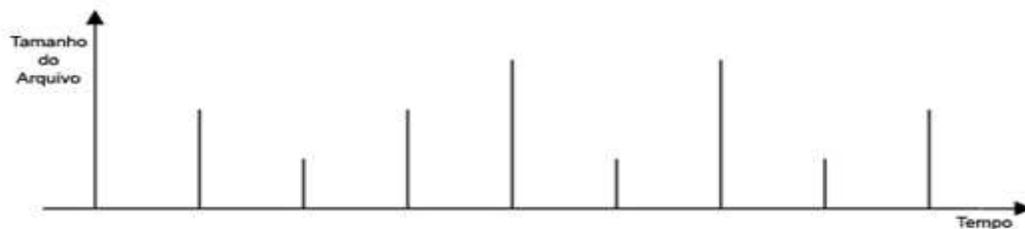


Figura 7 - Representação do perfil típico de um tráfego de blocos.

3.1.2 Tráfego de Transação

O tráfego de transação consiste em alterações “*ON*” e “*OFF*”. Durante o período “*ON*” pacotes são gerados continuamente. As seguintes variáveis são necessárias para definir uma sessão:

- Duração da sessão.
- Duração dos períodos *ON*.
- Duração dos períodos *OFF*.

- Tamanho do pacote.
- Intervalo de chegada dos pacotes durante o período *ON*.

Para cada variável, uma função de distribuição pode ser usada e os parâmetros correspondentes para as distribuições devem ser especificados. O perfil típico de um tráfego de transação é mostrado na Figura 8.

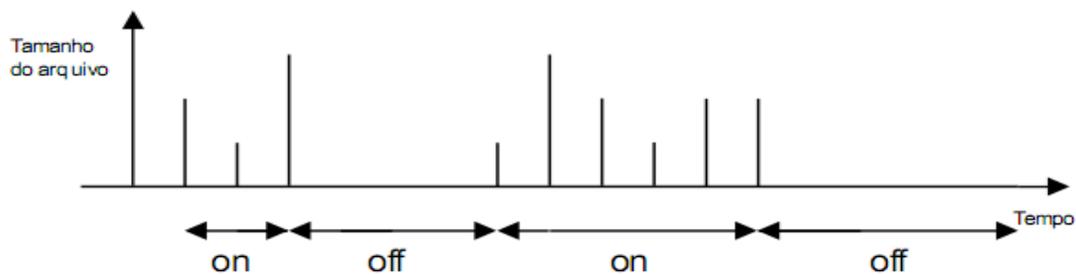


Figura 8 - Representação do perfil típico de um tráfego de transação.

Tanto o tráfego de blocos quanto o de transação correspondem a tráfegos elásticos. Para efeito de modelagem eles foram separados nessas duas categorias.

3.1.3 Tráfego *Stream*

A fonte do tráfego *stream* emite unidades de dados em períodos sem silêncio significativo. Uma representação esquemática pode ser vista na Figura 9. Além disso, é provável que se mantenham rajadas em diferentes escalas de tempo, devido aos chamados efeitos de auto similaridade [5]. A característica de um tráfego de vídeo *streaming* também depende da compressão adotada. O padrão MPEG-4 e H.263 são *codecs* candidatos potenciais para prover comunicação multimídia sobre redes sem fio, por exemplo [6]. O quadro pode ser modelado como:

- Uma fonte de taxa constante de *bits*, especificando a taxa de *bits*;
- E uma fonte de taxa de *bits* variável com o modelo auto regressivo [8];

- Uma fonte de taxa de *bits* variável utilizando o modelo F-ARIMA [9];
- Uma fonte de taxa de *bits* variável com o modelo Wavelet [10].

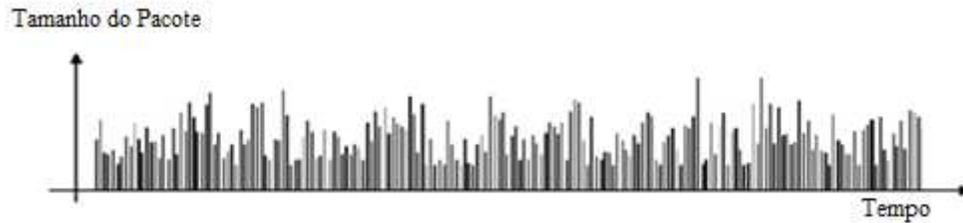


Figura 9 - Representação do perfil típico de um tráfego *Stream*.

3.1.4 Identificação da categoria de um tráfego multimídia

Para modelar uma aplicação multimídia utilizando um ou mais dos componentes básicos descritos acima, primeiramente deve-se fazer a sua identificação. Aplicações multimídia normalmente geram caminhos de tráfego, em que os componentes podem ser identificados a partir de suas fontes. Mesmo se dois componentes se originam da mesma fonte, os pacotes geralmente podem ser identificados por marcações adequadas. No *header* do TCP/IP, essas marcações podem ser identificadas no campo *Traffic Class* (IPv6) ou no ToS (IPv4) no cabeçalho dos pacotes enviados ou recebidos. Sempre que chega ou se transmite um novo pacote é especificado nesse campo qual o tipo de informação que está sendo transmitida. Na Tabela 4, são destacados tipos anais de aplicações e seus respectivos modelos.

Tabela 4 - Serviços multimídia e suas respectivas classes de tráfego.

Serviços	Bloco	Transação	<i>Streaming</i>
Orientação de Localização	X	X	
Pesquisa de Vão		X	
<i>Empresas – Negócios</i>			

Entrada de pedido		X	
Mensagem	X	X	
Palestras em sala	X	X	X
Mensagens de imagem	X		
Mensagens Multimídia	X	X	X
Video Mensagem	X	X	X
PDA Virtual	X	X	
<i>Entretenimento Móvel</i>			
Filmes Pequenos ou curtos			X
Músicas			X
Áudio <i>Stream</i>			X
Jogos Móveis		X	X
Telefone com vídeo		X	X
Chat Móvel		X	
Comércio Móvel	X	X	X
Máquinas de venda automática		X	
Controle Remoto		X	
Medição Remota		X	

3.2 Funções de distribuição

As distribuições de probabilidade mais utilizadas na prática, mesmo em desempenho de redes de comunicações e problemas de confiabilidade, são as distribuições exponencial e normal. O mesmo esquema pode ser usado para vários outros tipos de distribuição de probabilidade, contínuas ou discretas. O comportamento de cada serviço pode diferir de uma rede para outra, correspondendo a diferentes distribuições. A seguir são descritas algumas das distribuições mais

utilizadas em modelos de tráfego em redes de telecomunicações. Será feita uma breve descrição das distribuições e situações nas quais sua utilização é adequada.

Neste trabalho, vamos considerar a distribuição exponencial para o intervalo entre as requisições de serviço e para a duração do serviço. A distribuição exponencial é adequada para o serviço de voz em redes comutadas por circuito. Aqui, vamos utilizá-la como uma aproximação para os serviços de Voz, Vídeo e Dados, já que a ideia é observar o comportamento qualitativo da rede de pacotes. Assim, quando nos referimos ao intervalo médio entre as requisições de serviço ou à duração média do serviço, está implícito que se trata de uma distribuição exponencial.

3.2.1 Distribuição exponencial

A distribuição exponencial é utilizada quando se deseja analisar o espaço ou intervalo de acontecimentos de um evento, utilizada para eventos contínuos. É comumente utilizado para intervalos entre chegadas. Possui a propriedade da falta de memória. A função densidade de probabilidade (PDF) da distribuição exponencial é mostrada abaixo na equação 1:

$$f(x; \lambda) = \lambda e^{-\lambda x}, \quad (1)$$

onde $x \geq 0$ e $\lambda > 0$ são os parâmetros da distribuição.

3.2.2 Distribuição de Weibull (Cauda Longa)

A distribuição de Weibull, nomeada pelo seu criador Waloddi Weibull, é uma distribuição de probabilidade contínua. A distribuição de Weibull é utilizada para modelar o tempo entre falhas de equipamentos. Em redes de comunicações, um exemplo é a sua utilização na modelagem de transações onde pode representar o período que se situa um dos períodos (*on* ou *off*). A função densidade de probabilidade de Weibull para variável aleatória x é expressa por:

$$f(x; \lambda, k) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, \quad (2)$$

onde $x \geq 0$, $k > 0$ e $\lambda > 0$ é o parâmetro de escala da distribuição.

Sua função de distribuição acumulativa é uma extensão da função exponencial. A distribuição Weibull está relacionada a uma série de outras distribuições de probabilidade, em particular, pode descrever uma interpolação entre a distribuição exponencial ($k = 1$) e a distribuição normal ($k = 2$).

3.2.3 Distribuição de Pareto

A distribuição de Pareto foi introduzida primeiramente pelo economista italiano Vilfredo Pareto em 1897. Desde então muitos pesquisadores a utilizaram em diversas áreas. É uma distribuição de probabilidade que modela fatos dos campos social, científico, industrial e muitos outros tipos de fenômenos observáveis. Fora do campo da economia, é às vezes referida como a distribuição de Bradford. A função densidade de probabilidade da distribuição Pareto é:

$$f(x; \alpha) = \alpha \frac{x_m^\alpha}{x^{(\alpha+1)}}, \quad (3)$$

onde $x > x_m > 0$, e x_m é o valor mínimo de x . A família de distribuições de Pareto é parametrizada por, x_m e α , onde α é o índice de Pareto. Dentre as variáveis que assumem o comportamento da distribuição de Pareto, e que estão ligadas a modelos de tráfego em redes de comunicações, podem ser citados:

- Número de rajadas em uma conexão durante uma seção FTP;
- Tamanho das rajadas em *bytes*;
- Tamanho dos arquivos;

- Tempo de CPU consumido por processos;
- Período ocioso (quando não há tráfego de informação);
- Duração de uma conexão HTTP.

3.2.4 Distribuição Lognormal

Uma variável pode ser modelada como *log-normal* se puder ser considerada como o produto de muitos fatores pequenos independentes. Um exemplo típico é o retorno em longo prazo do investimento em uma ação. Pode-se considerar como o produto dos retornos diários.

Essa distribuição é também muito utilizada para estimativa de tempo de reparo de uma falha, ou seja, o tempo entre ocorrer a falha e o tempo para que ela seja corrigida. A função densidade de probabilidade da distribuição lognormal é dada por:

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{\ln(x-\mu)^2}{2\sigma^2}}, \quad (4)$$

onde μ e σ são a média e o desvio padrão respectivamente e $x > 0$.

3.2.5 Distribuição discreta de Poisson

A distribuição de Poisson é utilizada para variáveis aleatórias discretas. Ela determina a quantidade de ocorrências em um determinado intervalo de tempo. Os eventos devem ocorrer em certo intervalo de tempo ou de espaço. É utilizada em chamadas telefônicas por unidade de tempo, acidentes por unidade de tempo e chegada de clientes por unidade de tempo, por exemplo. Sua função densidade de probabilidade é dada por:

$$p(x; \lambda, t) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, x = 0, 1, 2, \dots \quad (5)$$

onde λ é a taxa média do processo e t é o intervalo de tempo ou espaço.

A distribuição de Poisson possui as seguintes características:

- A taxa média do processo λ é constante durante o período de tempo ou espaço considerado;
- Quanto menor o segmento de tempo ou espaço, menor a probabilidade de ocorrer mais de um evento naquele segmento. A probabilidade de ocorrência de dois ou mais eventos tende a zero quando o tamanho do segmento tende a zero.

3.3 CONCEITOS BÁSICOS DA MODELAGEM COM O ARENA¹

O *software* ARENA tem uma visão bastante peculiar, baseada na modelagem por eventos discretos. Os sistemas são reduzidos a um conjunto de estações de trabalho que prestam serviços a clientes (entidades). Cada bloco ou módulo é interconectado a outros e as entidades se locomovem entre eles seguindo uma sequência lógica.

Os modelos de simulação no *software* ARENA são construídos por meio de blocos que realizam funções pré-determinadas. Esses módulos são interconectados e formam uma rede de entidades por onde as informações são processadas. Eles descrevem a lógica do processo da maneira mais visual possível.

Cada entidade tem características próprias que as definem que são os atributos. As filas são um reservatório para entidades que não podem continuar seguindo normalmente no fluxo por

¹ As Referências [7] e [20] contém detalhes sobre a construção de modelos de simulação.

algum motivo, seja ele um recurso ocupado ou a espera de uma autorização. Existem dois tipos de fila: as individuais e as internas.

- As filas individuais possuem nome, modo de seleção e uma maneira específica de ordenar as entidades. Esse tipo de fila pode ser mostrado no modo de animação;
- As filas internas são *First-In, First-Out* (FIFO) e não podem ser mostradas no modo de animação. O número de entidades esperando em uma fila pode ser acessado por meio da expressão NQ (*Number in Queue*).

Para executar qualquer simulação nesse *software* é preciso descrever as estações de trabalho, o fluxo dentro do sistema, as durações, as distâncias, as velocidades, etc. O *software* utiliza um modelo de programação visual como técnica de modelagem, ou seja, é criado um diagrama de blocos que corresponde à ocorrência de eventos para um cliente genérico que flui pelo sistema sendo modelado. Portanto, cada bloco representa a função no sistema e trata de maneira adequada as entidades conforme o fluxo de tráfego e a ocorrência dos eventos.

O trabalho consiste em desenvolver um modelo de simulação por eventos discretos para serviços multimídia em redes de acesso para banda larga móvel. O ARENA é um *software* de simulação de sistemas que fornece uma interface gráfica que permite a elaboração de um modelo de simulação baseado na linguagem SIMAN e foi lançado em 1993 pela empresa Rockwell [20].

A simulação de sistemas é uma das ferramentas para estudo de processos em sistemas que pode ser feita por meio do uso de protótipos ou pela elaboração de modelos matemáticos. É uma ferramenta eficaz no auxílio à tomada de decisões dentro de uma empresa de logística, por exemplo. Com isso é possível criar e testar modelos dos diversos sistemas de uma corporação [20, 21].

Testando novas ideias e projetos em um simulador, é possível prever o que acontecerá, sem ter que efetivamente alterar qualquer coisa no ambiente físico. Os principais passos de uma modelagem utilizando simulação por eventos discretos são:

- Criar um modelo básico;
- Verificar o modelo por meio de simulações limitadas;
- Validar o modelo usando alguma referência;
- Executar as simulações para vários valores dos parâmetros;
- Analisar os resultados da simulação (voltar ao passo anterior, se necessário);
- Selecionar o melhor conjunto de parâmetros.

O modelo de simulação deste trabalho foi criado no *software* ARENA [7]. Para descrever a construção de um modelo de simulação por eventos discretos neste *software* um modelo similar foi utilizado como exemplo, construído recentemente, que descreve o comportamento do tráfego multimídia em uma rede IP [8]. Outro modelo, para simular o tráfego de dados em terminais móveis *always-on*, pode ser utilizado para escolher a melhor configuração de uma rede 3G com objetivo de aperfeiçoar o consumo de energia dos terminais móveis conectados permanentemente à rede [9, 10].

O modelo usado em [8] consiste em três tipos de tráfego: videoconferência, vídeo sob demanda e vídeos. Nesse caso, o tráfego tem característica de *stream* e pode ser representado por distribuições exponenciais negativas para os intervalos entre requisições de serviço e para a duração das conexões. É considerado também um controle de admissão primitivo, que estabelece um limite máximo de serviços simultâneos para cada tipo de serviço. O *software* gera os eventos de acordo com a distribuição escolhida e usa variáveis aleatórias e o método de Monte Carlo. Assim, é importante que as simulações para os diversos serviços sejam executadas diversas vezes, aumentando o número de eventos, até que os resultados mostrem convergência, ou seja, não mudem quando o número de eventos aumenta.

Os parâmetros básicos do modelo para o serviço de videoconferência, por exemplo, considerando uma distribuição exponencial negativa, são os seguintes:

- Intervalo médio entre requisições de serviço;
- Duração média do serviço;
- Número máximo de serviços simultâneos.

A Tabela 5 mostra os resultados obtidos para dez execuções de 100.000 segundos com 2.000 segundos de *warm-up* considerando um canal de 4 Mbps. As dez execuções mostram como a simulação depende bastante das sementes utilizadas nas variáveis aleatórias para os sorteios de Monte Carlo. Por isso é tão importante que se façam várias replicações. As perdas se devem à duração dos serviços e ao limite de serviços simultâneos no controle de admissão. Não houve perdas para o serviço de videoclipes. A perda média é de 2% para o serviço de Videoconferência e próxima a 9% para o serviço de *Video On Demand*.

Os intervalos de confiança foram calculados com a suposição de que a distribuição é Gaussiana com 95% de confiança. Para obter valores mais precisos basta aumentar o tempo de simulação suficientemente, por exemplo, uma ou duas ordens de grandeza.

Tabela 5 - Resultado para dez rodadas da simulação para um canal de 4 Mbps, tempo total de simulação de 100.000 segundos e tempo de *warm-up* de 2000 segundos.

Rodada	1	2	3	4	5	6	7	8	9	10	Média	L_inf	L_sup
Número de Videoconferências	17	32	38	20	23	27	27	23	30	18	25,5	12,5	38,5
Videoconferências Completadas	18(*)	29	35	20	23	27	28	23	30	18	25,1	14,1	36,1
Videoconferências Perdidas	0	1(**)	3	0	0	0	0(*)	0	1(*)	0	0,5	0,0	2,4
Número de <i>Videos On Demand</i>	7	10	4	7	6	15	10	8	4	7	7,8	1,4	14,2
<i>Videos On Demand</i> Completados	7	10	6	7	6	13	8	8	4	7	7,6	2,8	12,4
<i>Videos On Demand</i> Perdidos	0	1(*)	0(*)	0	0	3(*)	2	1(*)	0	0	0,7	0,0	2,8
Número de <i>Video Clips</i>	109	108	120	120	128	119	106	103	109	114	113,6	98,1	129,1
<i>Video Clips</i> Completados	109	108	120	120	128	119	106	103	109	114	113,6	98,1	129,1
<i>Video Clips</i> Perdidos	0	0	0	0	0	0	0	0	0	0	0,0	0,0	0,0

(*) – O número de solicitações de serviço somado ao número de serviços perdidos não é igual ao número de serviços completados devido ao fato de que a simulação desconsidera o cálculo de 2000 segundos (tempo de *warm-up*).

(**) – Significa que há serviços em processamento quando do fim da simulação.

A Figura 10 mostra as quantidades totais de videoconferências e a perda em função do intervalo médio entre requisições de serviço para uma duração média do serviço fixa. À medida que o intervalo aumenta, diminuem os serviços e as perdas como esperado.

Na Figura 11, observam-se as quantidades de videoconferências completadas e as perdas em função da duração média para um intervalo médio fixo entre requisições. Quando a duração do serviço aumenta, com o intervalo médio entre requisições e número máximo de serviços simultâneos fixos, os serviços completados diminuem e as perdas aumentam, mostrando o colapso da rede.

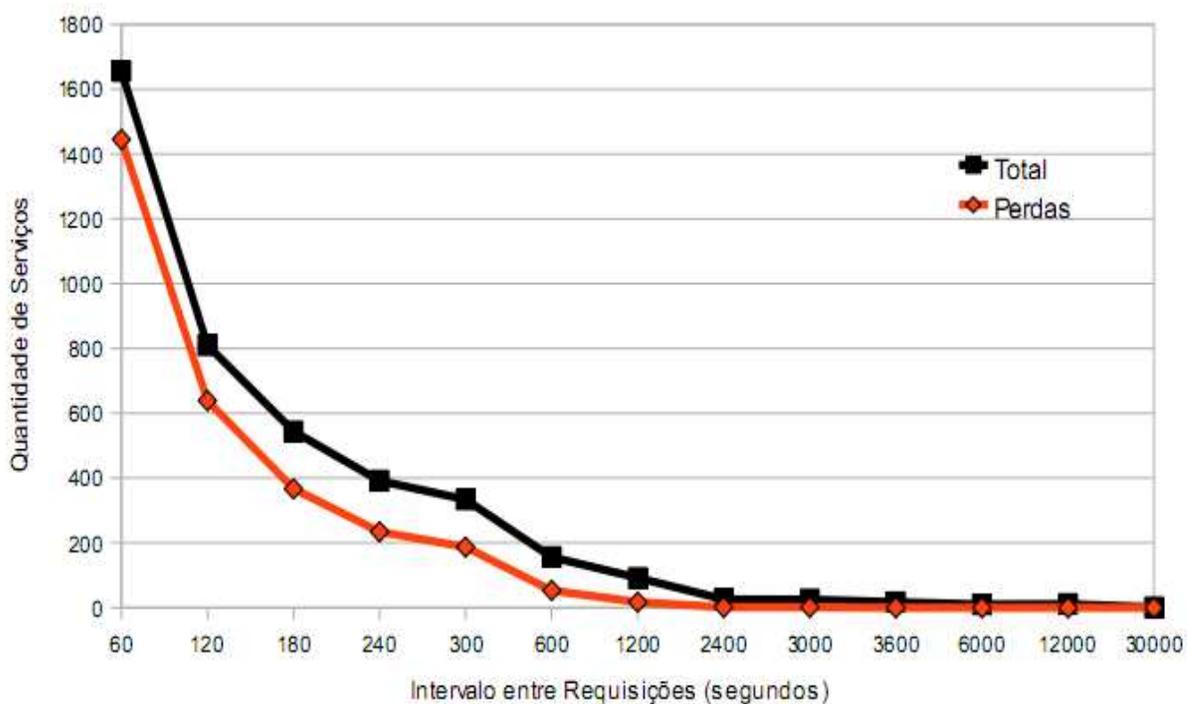


Figura 10 - Quantidade de videoconferências em simulações feitas com o cenário considerado em função do intervalo entre as requisições de serviço para uma duração média fixa de 900s (15min).

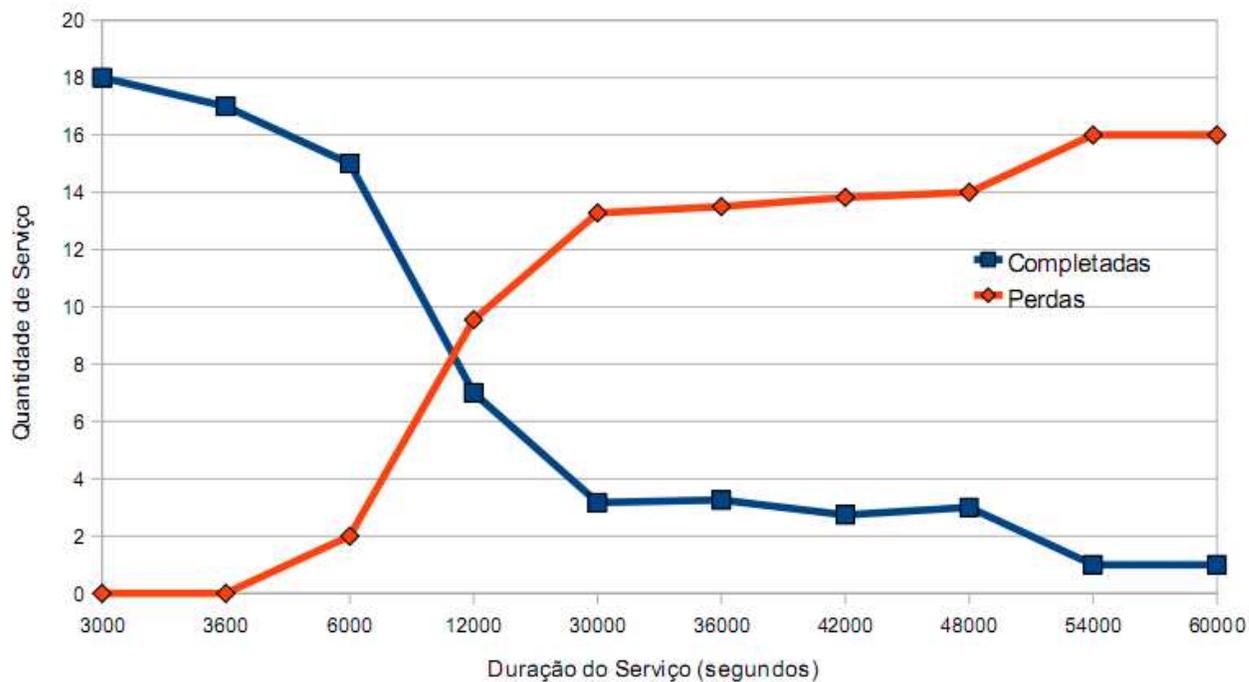


Figura 11 - Quantidade de videoconferências em simulações feitas com o cenário considerado em função da duração média do serviço para um intervalo médio entre as requisições de serviço fixo de 3600s.

É possível manter tanto o intervalo médio de requisição de serviço quanto à duração média fixa e variar o número máximo de serviços simultâneos. O resultado pode ser visto na Figura 12. Enquanto o número total de serviços se mantém constante, à medida que o limite de serviços simultâneos aumenta, o número de serviços completados cresce e as perdas diminuem. Note que foi preciso aumentar exageradamente a duração média do serviço para que fosse possível observar uma variação substancial da quantidade de serviços para uma variação pequena no limite de serviços simultâneos.

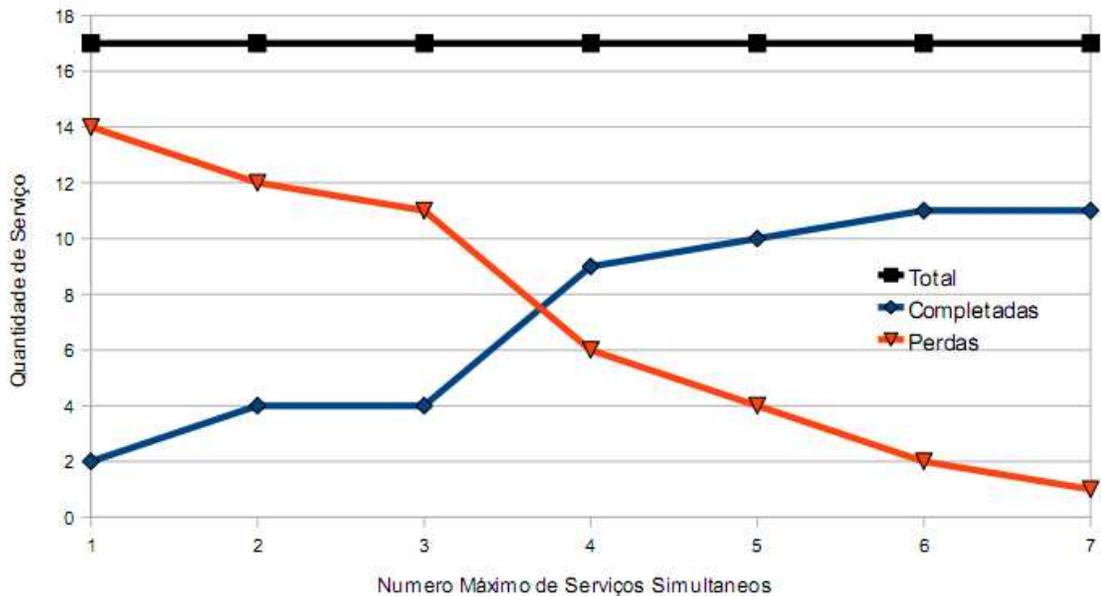


Figura 12 - Quantidade de videoconferências em simulações feitas com o cenário considerado em função do número máximo de serviços simultâneos para uma duração média fixa de 30.000s e um intervalo médio entre as requisições fixo de 3600s.

O modelo para a rede de acesso de banda larga móvel foi criado baseado nas seguintes propriedades:

- A rede de acesso é *single-node*, ou seja, consideramos apenas um sub-sistema de estação base (RNC+RBSs) interagindo com os dispositivos móveis;
- O núcleo da rede atende toda a demanda da rede de acesso, de modo que seus elementos não precisam ser incluídos no modelo de simulação;
- Os intervalos de requisição e as durações seguem distribuições exponenciais.

Trata-se de um modelo de rede multimídia em que, por exemplo, os serviços *stream* são estressados e o serviço elástico (representado pelo serviço de dados) tem o correspondente impacto. Outro aspecto fundamental é que a metodologia estabelece uma metodologia que é mais qualitativa que quantitativa, muito embora apresente valores numéricos. Para cada caso real, os valores observados em campo podem ser inseridos no modelo (ou em uma adaptação dele) e assim ele pode ser uma ferramenta efetiva de planejamento ou mostrar uma ideia de dimensionamento.

4 RESULTADOS

Neste capítulo são apresentados os resultados obtidos com as simulações da rede de acesso considerando os três tipos de tarifação diferentes: tarifação pelo tempo de utilização do canal de rádio (tempo de conexão), tarifação pelo número de conexões (independentemente da duração de cada chamada) e tarifação fixa (independentemente do número de chamadas e da duração de cada conexão). Além disso, consideramos dois modelos de simulação: o primeiro modelo não inclui atributos como *jitter* e latência e o segundo modelo possui esses atributos incluídos. Tais atributos têm influência no tempo de entrega do quadro ao seu destino, podendo causar problemas como degradação da qualidade ou bloqueio nos serviços *stream*.

Os modelos foram analisados considerando um estudo da demanda das requisições dos usuários da rede de acesso. Isso foi feito observando as quantidades de serviços (completadas, perdas e totais) em função dos três parâmetros fundamentais dos modelos de simulação considerados: o intervalo médio entre as requisições de serviço, a duração média do serviço e o número máximo de serviços simultâneos. Lembrando que as distribuições de chegada e de serviço são sempre exponenciais.

É importante notar a relação entre os parâmetros do modelo e os diferentes tipos de tarifação considerados. A tarifação pelo tempo de utilização do canal tende a gerar serviços com duração média menor, já que a duração da chamada influi diretamente no seu custo. Assim, em uma rede com tarifação baseada no tempo esperamos serviços com duração média menor. Quando a tarifação é feita pelo número de conexões, ou seja, paga-se um valor fixo por chamada independentemente de sua duração, espera-se serviços com duração média maior, já que o tempo de conexão não influi no valor da chamada. Além disso, espera-se um intervalo médio entre as requisições de serviços um pouco menor do que no caso da tarifação baseada no tempo pois as chamadas têm um preço fixo, o que leva a um maior número de ligações por parte dos usuários da rede. No caso da tarifação fixa, espera-se que o intervalo médio entre as requisições seja ainda menor do que no caso da tarifação por chamada e que as durações médias sejam tão grandes

quanto no caso da tarifação por chamada, já que nem o número de chamadas nem a duração das chamadas influem na tarifação (tarifação fixa).

Considerando a tarifação pelo tempo de utilização do canal de rádio, primeiramente fixamos a duração média do serviço e variamos o intervalo médio entre as requisições. As Figuras 13-15 mostram as quantidades de serviço em função do intervalo médio entre as requisições para uma duração média fixa de 900s (15 min) e no máximo dez (10) serviços simultâneos.

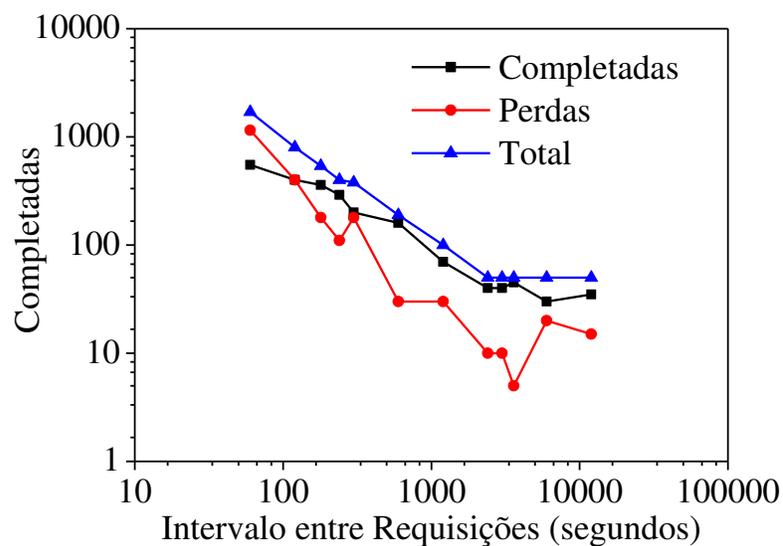


Figura 13 - Quantidade de serviços de voz em função do intervalo médio entre as tentativas de chamada para uma duração média de 900 s e um máximo de 10 serviços simultâneos.

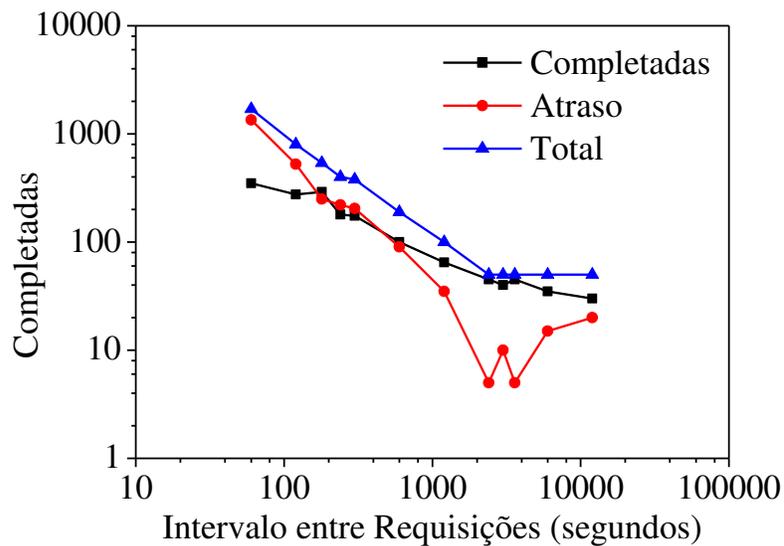


Figura 14 - Quantidade de serviços de dados em função do intervalo médio entre as tentativas de envio de dados para uma duração média de 900 s e um máximo de 10 serviços simultâneos.

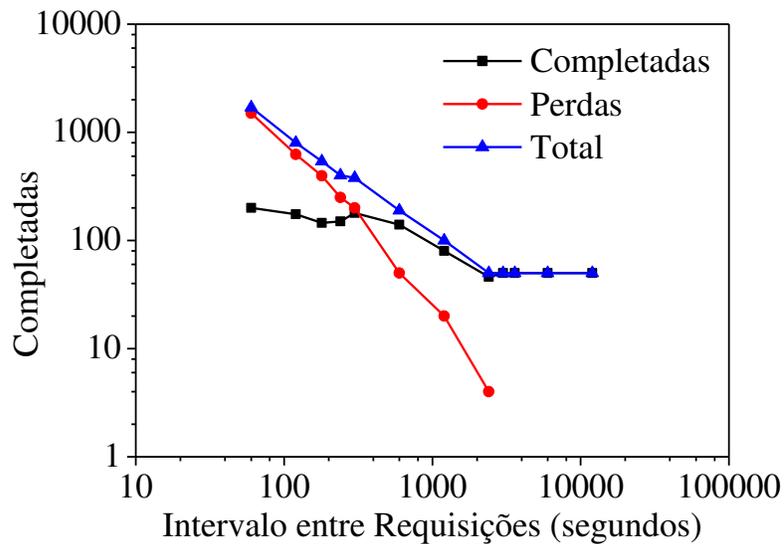


Figura 15 - Quantidade de serviços de videoconferências em função do intervalo médio entre as tentativas de videoconferências para uma duração média de 900 s e um máximo de 10 serviços simultâneos.

Podemos notar pelo comportamento linear no gráfico log-log, que ao aumentar o intervalo entre as requisições, o número de serviços diminui aproximadamente como uma lei de potência.

No caso do serviço de vídeo é possível ver claramente o número de serviços completados chegando ao número total quando não se tem mais perdas.

Em seguida, fixamos o intervalo médio entre as requisições e variamos a duração média dos serviços. Nas Figuras 16-18 as simulações foram realizadas considerando um intervalo médio entre as requisições de serviço fixo de 3600 segundos (1 hora) e no máximo dez (10) serviços simultâneos.

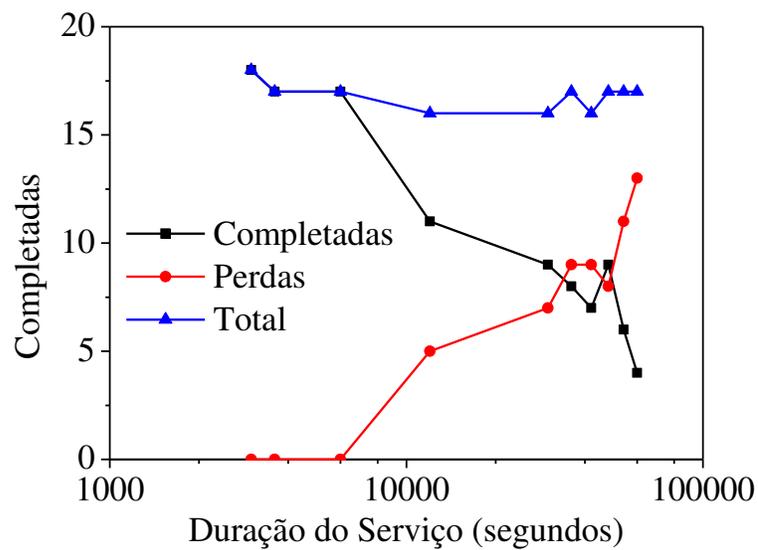


Figura 16 - Quantidade de serviços de voz em função da duração média das chamadas de voz para um intervalo médio de 3600 s.

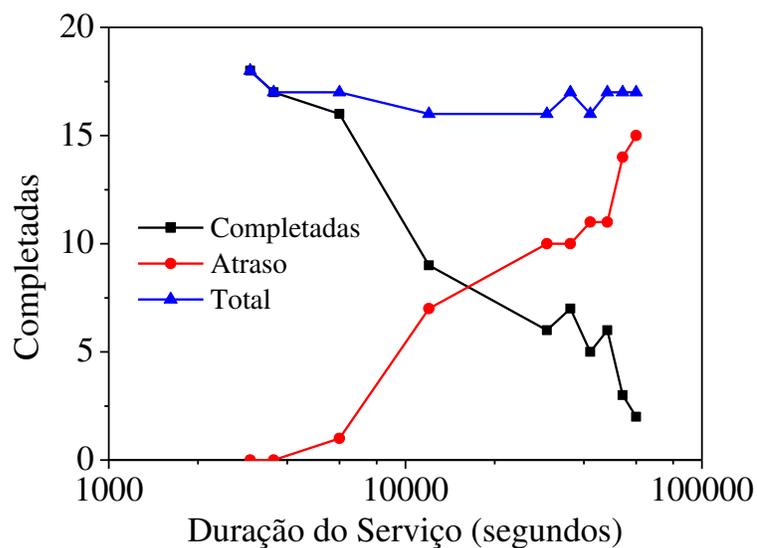


Figura 17 - Quantidade de serviços de dados em função da duração do serviço para um intervalo médio de 3600 s.

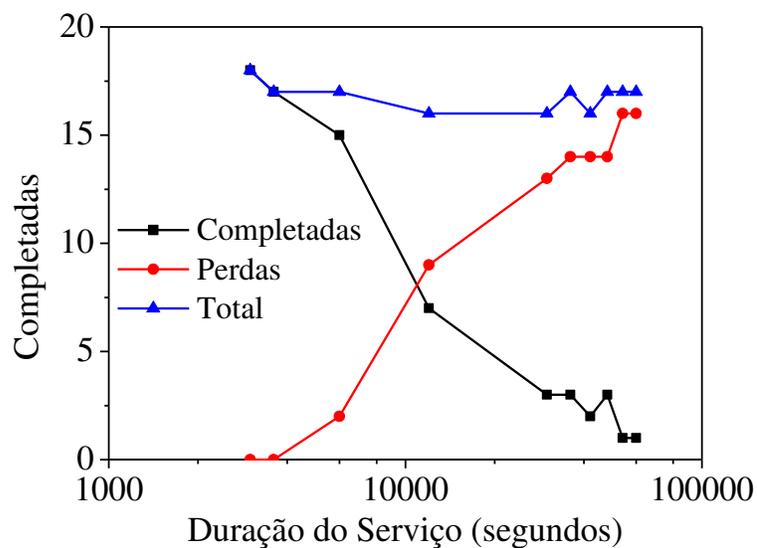


Figura 18 - Quantidade de videoconferências em função da duração do serviço.

Com o intervalo entre as requisições fixo, podemos ver claramente o colapso da rede quando a duração do serviço aumenta. O número de serviços completados vai a zero e as perdas chegam ao total de serviços.

As Figuras de 19 a 21 mostram os resultados obtidos considerando uma duração média fixa de 30.000 segundos e um intervalo médio entre as requisições fixo de 3600 segundos. Nota-se que quanto menos tráfego o serviço exige (como no caso do serviço de voz), ocorre menos perda. Neste caso os intervalos entre as requisições de serviço tendem a serem maiores e as durações dos serviços menores.

Nos casos dos serviços de dados e de vídeo nota-se claramente a restauração dos serviços na rede à medida que se aumenta o número máximo de serviços simultâneos. Na prática, isso significa que a rede de acesso necessitava mais canais para operar sem perdas.

Em seguida consideramos a tarifação pelo número de chamadas, independente da duração de cada conexão. As Figuras 22-24 representam graficamente os resultados da simulação, com o cenário considerado em função do intervalo entre as requisições de serviço para uma duração média fixa de 900s (15min) e o número máximo dez (10) de serviços simultâneos considerando a tarifação por chamadas.

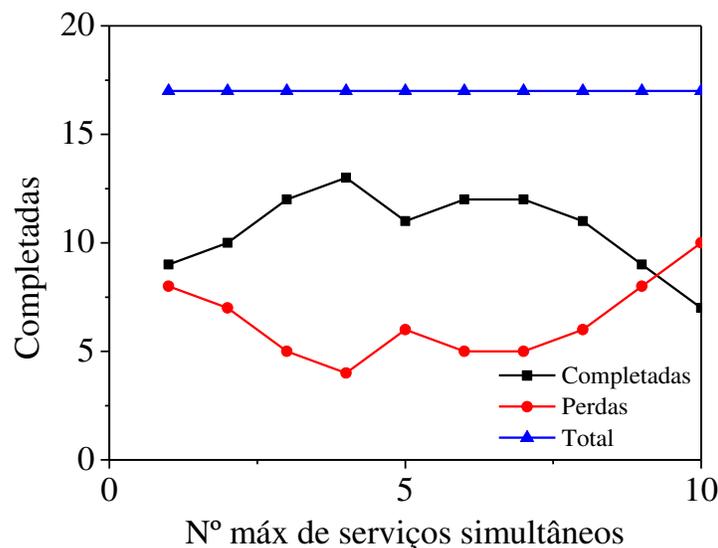


Figura 19 - Quantidade de serviços de voz em função do número máximo de serviços simultâneos.

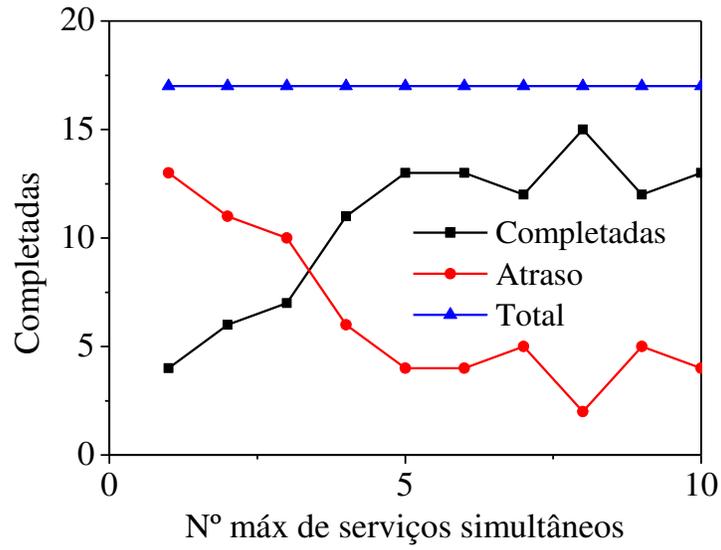


Figura 20 - Quantidade de serviços de dados em função do número máximo de serviços simultâneos.

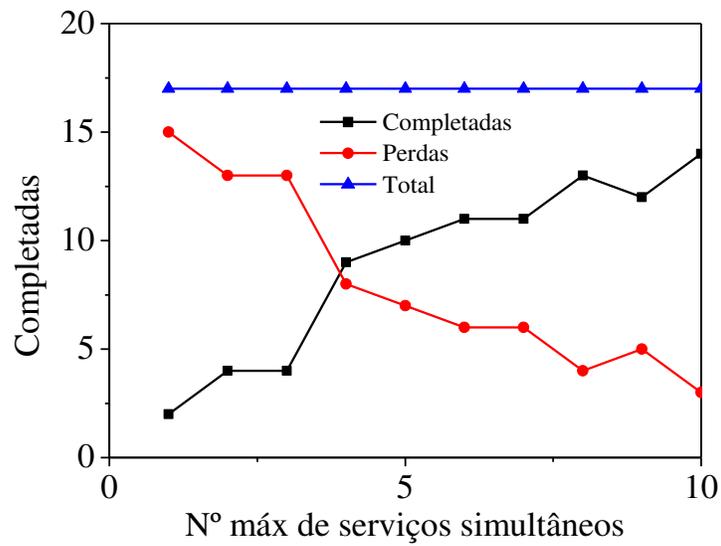


Figura 21 - Quantidade de videoconferências em função do número máximo de serviços simultâneos.

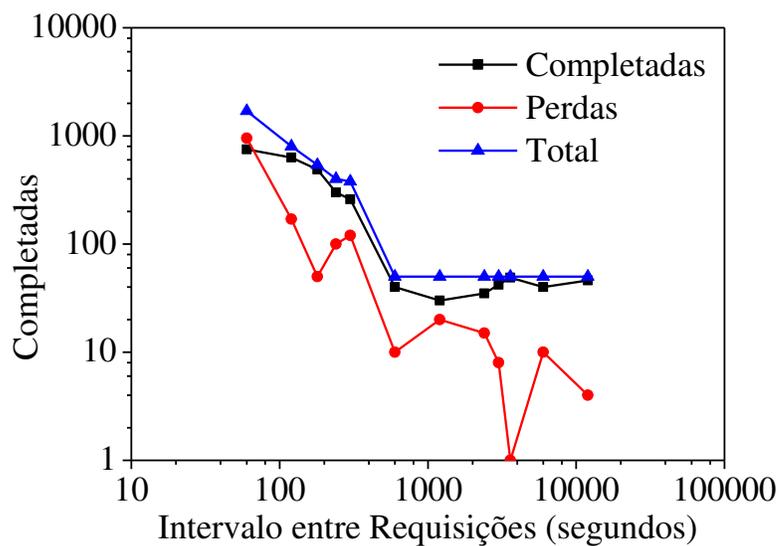


Figura 22 - Quantidade de serviços de voz em simulações considerando a tarifação por nº de chamadas.

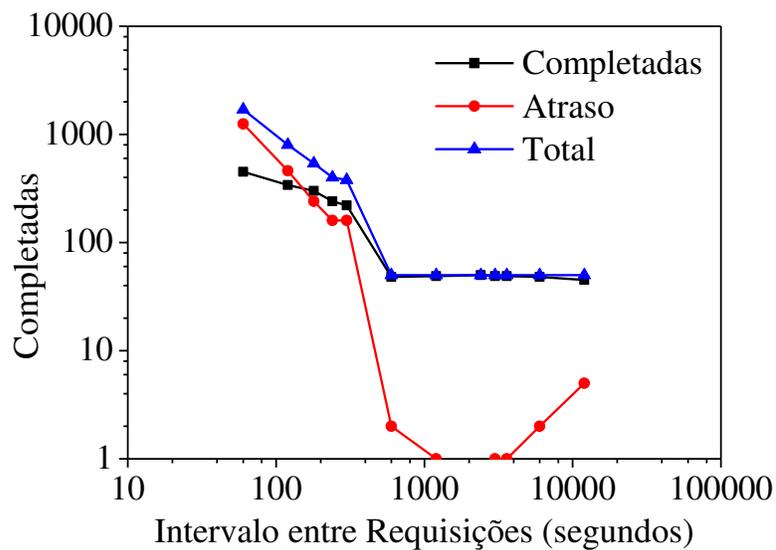


Figura 23 - Quantidade de serviços de dados em simulações considerando a tarifação por nº de chamadas.

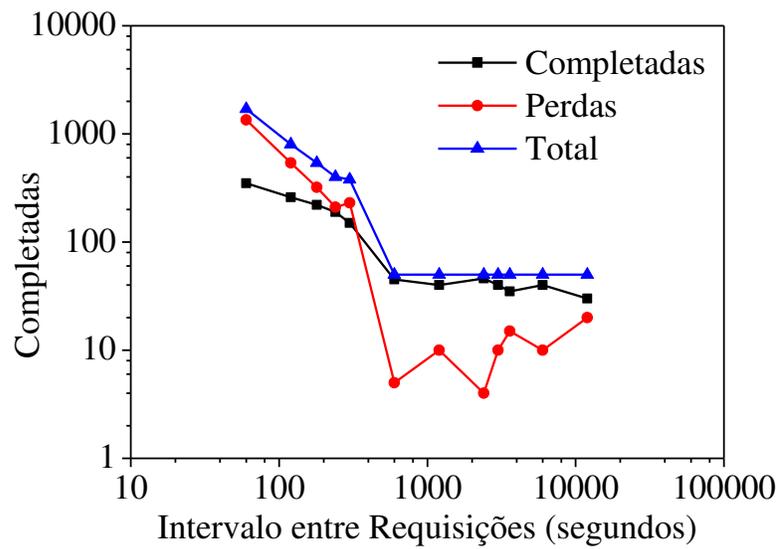


Figura 24 - Quantidade de videoconferências em simulações considerando a tarifação por nº de chamadas.

Nas Figuras 25-27 as simulações foram realizadas considerando um intervalo médio entre as requisições de serviço fixo de 3600 segundos.

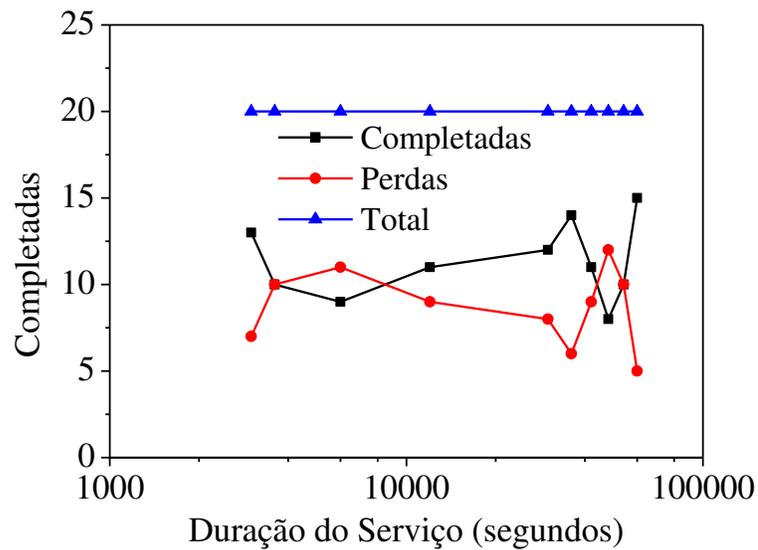


Figura 25 - Quantidade de serviços de voz em simulações com intervalo de requisição fixa considerando a tarifação por nº de chamadas.

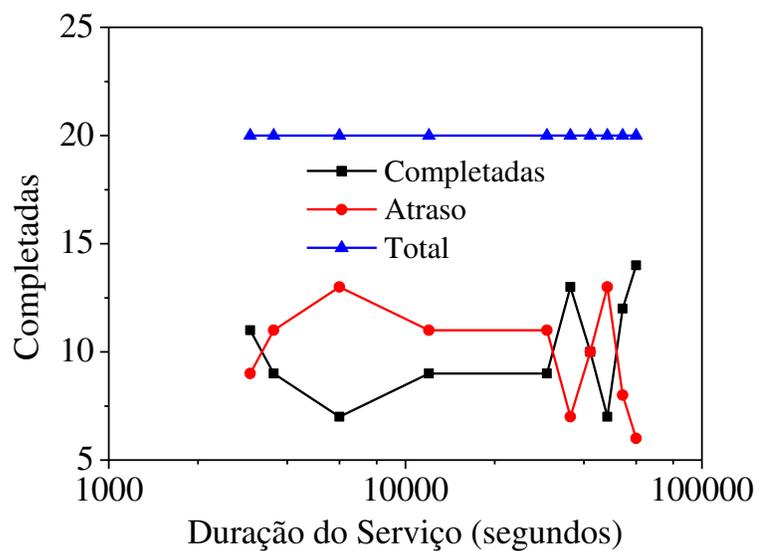


Figura 26 - Quantidade de serviços de dados em simulações com intervalo de requisição fixa considerando a tarifação por nº de chamadas.

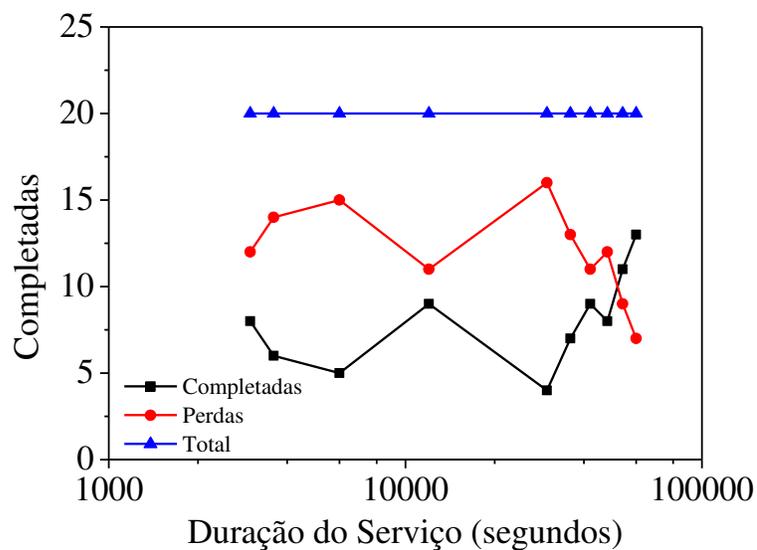


Figura 27 - Quantidade de videoconferências em simulações com intervalo de requisição fixa considerando a tarifação por nº de chamadas.

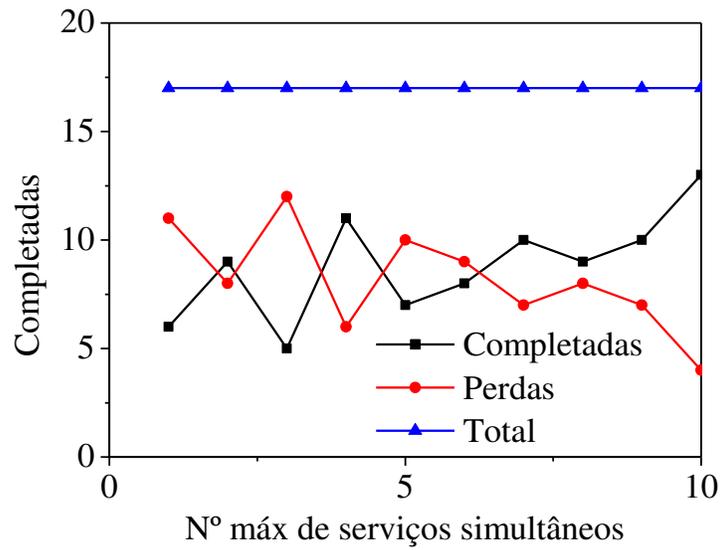


Figura 28 - Quantidade de serviços de voz em simulações com intervalo de requisição e duração média fixa considerando a tarifação por nº de chamadas.

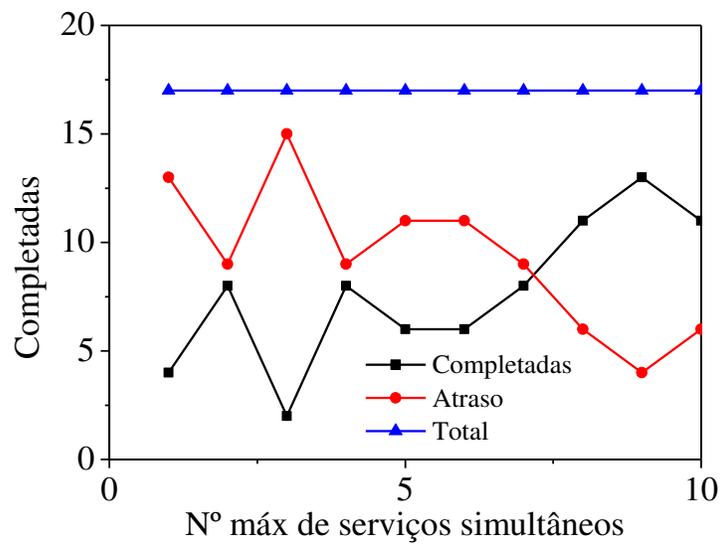


Figura 29 - Quantidade de serviços de dados em simulações com intervalo de requisição e duração média fixa considerando a tarifação por nº de chamadas.

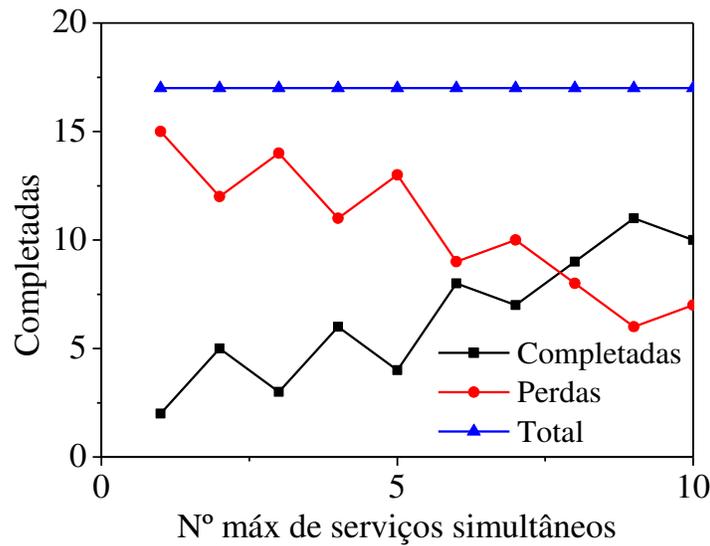


Figura 30 - Quantidade de videoconferências em simulações com intervalo de requisição e duração média fixa considerando a tarifação por nº de chamadas.

As simulações representadas nas Figuras 28-30 mostram os dados considerando uma duração média fixa de 30.000 segundos e um intervalo médio entre as requisições fixo de 3600 segundos. Nota-se que, no segundo caso, as durações aumentam substancialmente e os intervalos entre as requisições diminuem em relação ao primeiro.

A seguir são mostrados os resultados obtidos ao considerar a tarifação fixa, ou seja, o valor cobrado é fixo, independente da duração e do número de conexões. As Figuras 31-33 mostram os as quantidades de serviços em função do intervalo médio entre as requisições de serviço para uma duração média fixa de 900 s (15 min) e o número máximo de serviços simultâneos igual a 10, baseados em uma tarifação fixa. Nesse caso também observamos um comportamento de lei de potência, como nos outros cenários.

Nas Figuras 34-36 as simulações foram realizadas considerando um intervalo médio entre as requisições de serviço fixo de 3600 segundos. Vemos que com a tarifação fixa as perdas ficam da mesma ordem dos serviços completados para voz e dados. No caso do serviço de vídeo, fica ainda mais evidente o aumento das perdas. Esse resultado era esperado, já que com este tipo de

tarifação o tráfego aumenta muito e os recursos da rede não são suficientes para atender um volume tão grande de solicitações de serviço.

Os resultados apresentados nas Figuras 37-39 mostram a quantidade de serviços em função do número máximo de serviços simultâneos, considerando uma duração média fixa de 30.000 segundos e um intervalo médio entre as requisições fixo de 3600 segundos.

Nota-se que os intervalos entre as requisições diminuem em relação ao segundo caso e diminuem ainda mais em relação ao primeiro, diminuindo também em relação ao segundo.

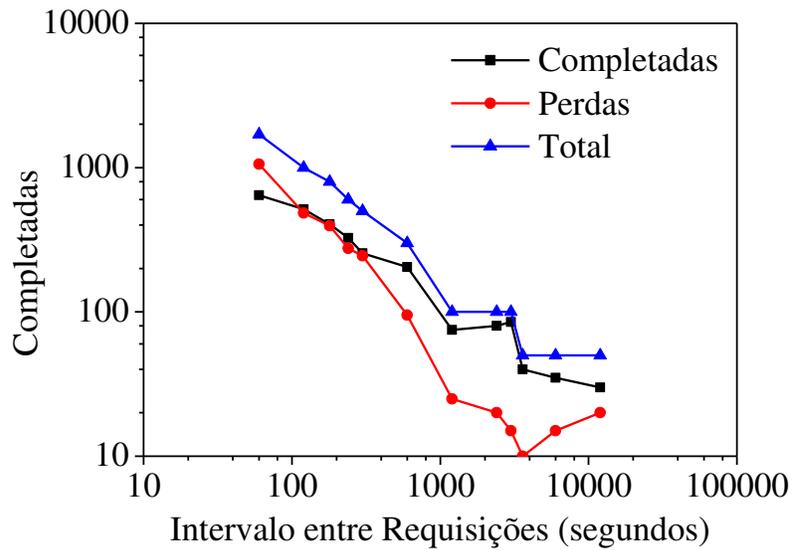


Figura 31 - Quantidade de serviços de voz em simulações considerando a tarifação fixa

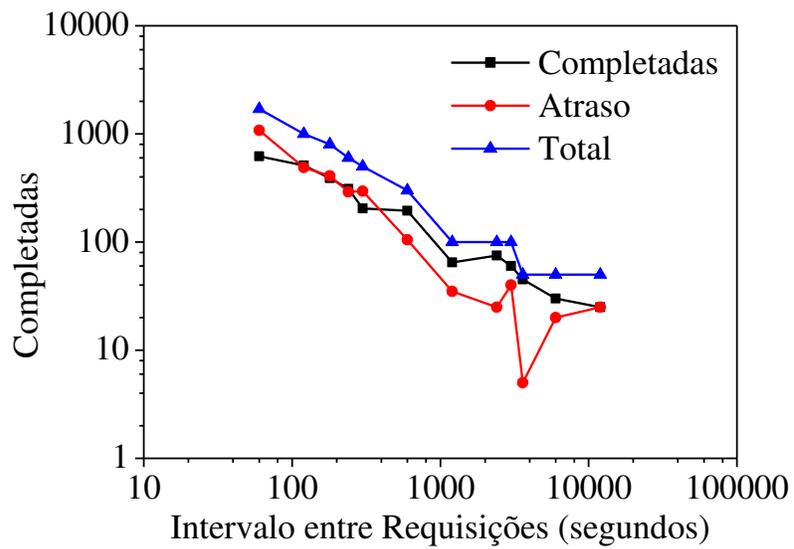


Figura 32 - Quantidade de serviços de dados em simulações considerando a tarifação fixa.

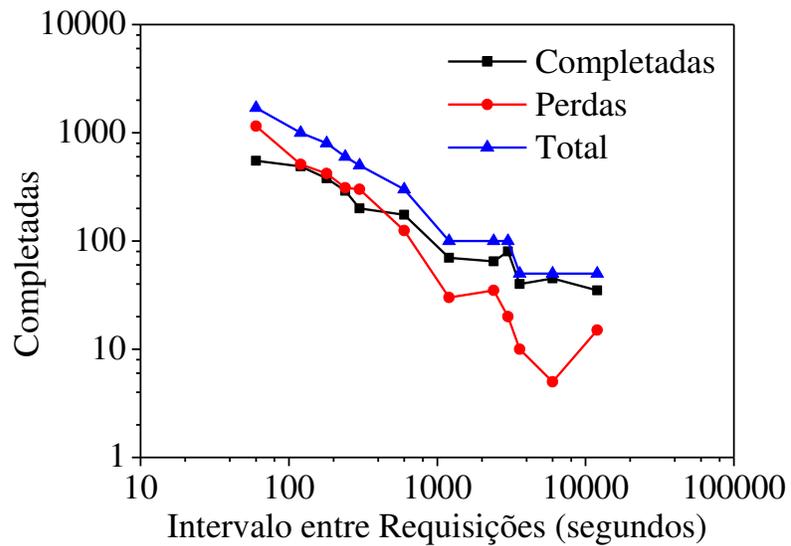


Figura 33 - Quantidade de videoconferências em simulações considerando a tarifação fixa.

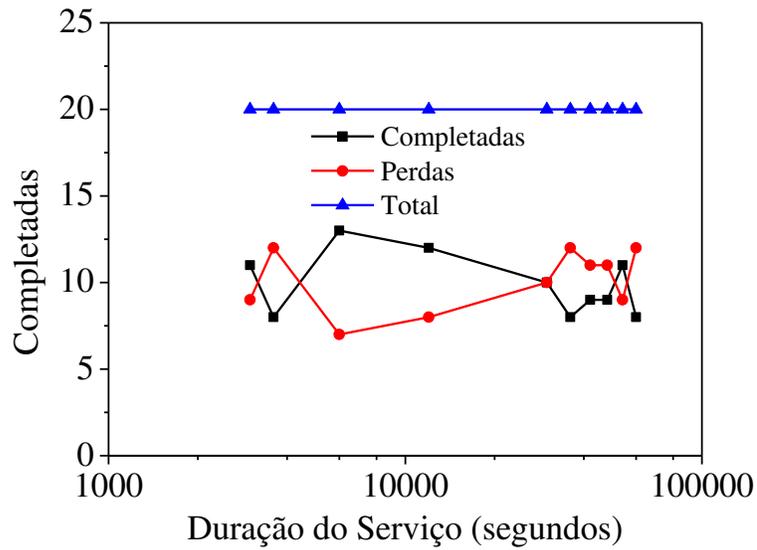


Figura 34 - Quantidade de serviços de voz em simulações com intervalo de requisição fixa considerando a tarifação fixa.

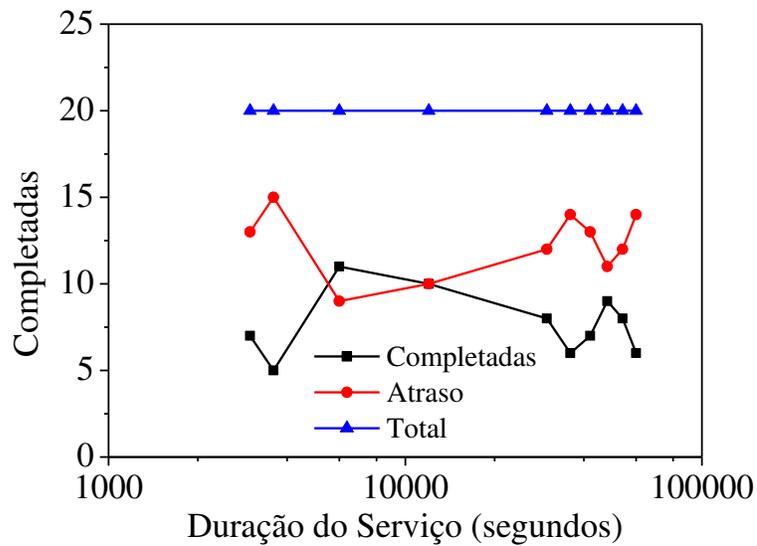


Figura 35 - Quantidade de serviços de dados em simulações com intervalo de requisição fixa considerando a tarifação fixa.

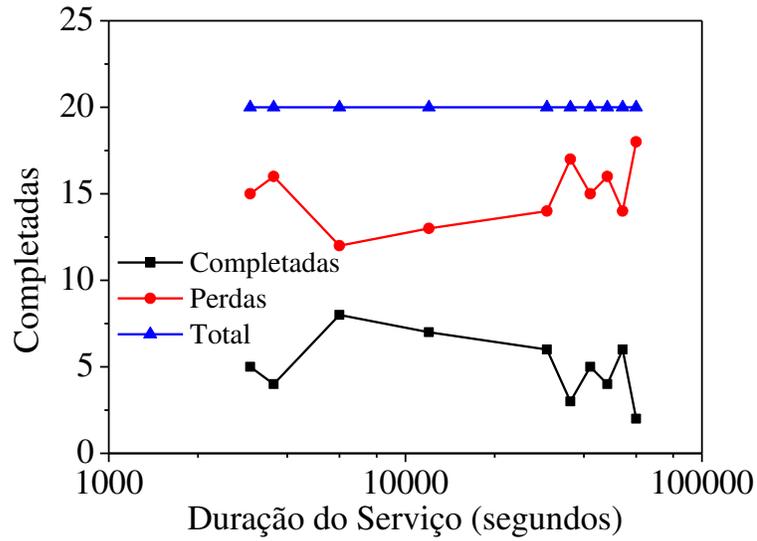


Figura 36 - Quantidade de videoconferências em simulações com intervalo de requisição fixa considerando a tarifação fixa.

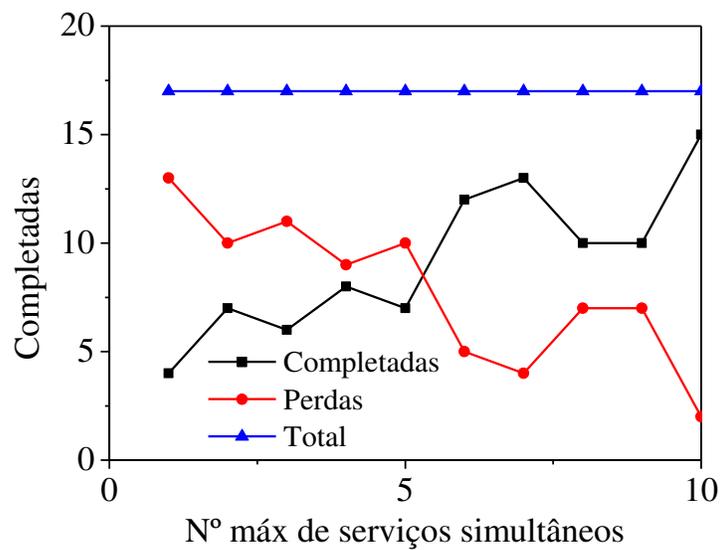


Figura 37 - Quantidade de serviços de voz em simulações com intervalo de requisição e duração média fixa considerando a tarifação fixa.

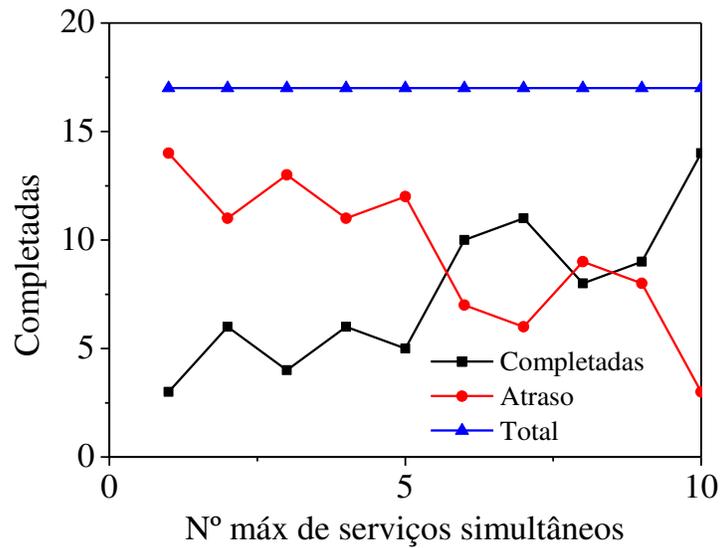


Figura 38 - Quantidade de serviços de dados em simulações com intervalo de requisição e duração média fixa considerando a tarifação fixa.

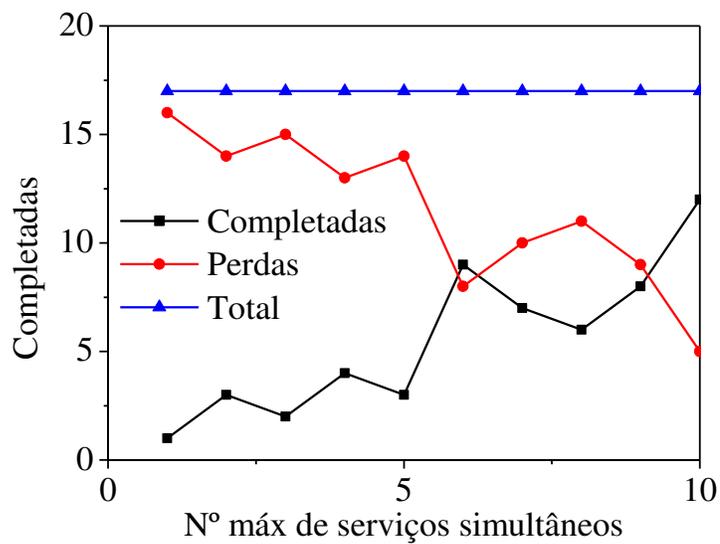


Figura 39 - Quantidade de videoconferências em simulações com intervalo de requisição e duração média fixa considerando a tarifação fixa.

O Modelo 2 ilustra a análise dos três tipos de tarifação acima considerando os atributos de *jitter* e latência, podem causar mais perdas.

As aplicações não interativas não são muito afetadas pelo *jitter*, porém os usuários de Voz sobre IP (VoIP) e videoconferência sofrem bastante interferência. Esse atraso ocorre devido à utilização dos *codecs*. Os pacotes gerados na origem, que inicialmente possuem taxas constantes, sofrem retardos diferentes e não chegam ao destino sincronizados, após atravessarem a rede.

Ao serem utilizados valores de *jitter* e latência nos casos de serviços de voz a perda é insignificante em relação aos valores analisados nos gráficos acima. Para os serviços de videoconferência ocorrem perdas de serviço. Os valores de *jitter* e latência que minimizam a perda dos serviços de videoconferência são apresentados na Tabela 6.

Considerando a tarifação pelo número de chamadas, as Figuras 43-45, mostram os dados das quantidades de serviços em função do intervalo médio entre as requisições de serviço para uma duração média fixa de 900 s (15 min) e o número máximo de serviços simultâneos iguais a 10. Depois é considerado um intervalo médio entre as requisições de serviço fixo de 3600 segundos e por fim uma duração média fixa de 30.000 segundos e um intervalo médio entre as requisições fixo de 3600 segundos. Estes dados são considerados, também, para a tarifação fixa, representados nas Figuras 46-48.

Tabela 6 - Melhores taxas de tempo para 1 única perda de serviço.

Fluxo de Vídeo (Mbit/s)	Latência	Jitter	Perda de serviços	Período de Coleta 30 minutos
3.0	<200ms	<50ms	1	1 evento de erro a cada 30 minutos
3.75	<200ms	<50ms	1	1 evento de erro a cada 30 minutos
5.0	<200ms	<50ms	1	1 evento de erro a cada 30 minutos

Os gráficos a seguir mostram os dados da análise da tarifação pelo tempo de utilização do canal de rádio considerando *jitter* e latência como mostrados na Tabela 6 acima.

A Figura 40 mostra as quantidades de serviços em função do intervalo médio entre as requisições de serviço para uma duração média fixa de 900 s (15 min) e o número máximo de serviços simultâneos iguais a 10. Como no modelo anterior, temos também uma lei de potência. Na Figura 41 podemos observar o colapso da rede quando a duração média do serviço aumenta para um intervalo médio entre as requisições de serviço fixo de 3600 segundos. Vemos que o número de serviços completados vai a zero e as perdas se aproximam do total de solicitações. A Figura 42 mostra a quantidade de serviços em função do número máximo de serviços simultâneos, considerando uma duração média fixa de 30.000 segundos e um intervalo médio entre as requisições fixo de 3600 segundos.

Considerando a tarifação pelo número de chamadas, as Figuras 43-45, mostram os dados das quantidades de serviços em função do intervalo médio entre as requisições de serviço para uma duração média fixa de 900 s (15 min) e o número máximo de serviços simultâneos iguais a 10. Depois é considerado um intervalo médio entre as requisições de serviço fixo de 3600 segundos e por fim uma duração média fixa de 30.000 segundos e um intervalo médio entre as requisições fixo de 3600 segundos. Estes dados são considerados, também, para a tarifação fixa, representados nas Figuras 46-48.

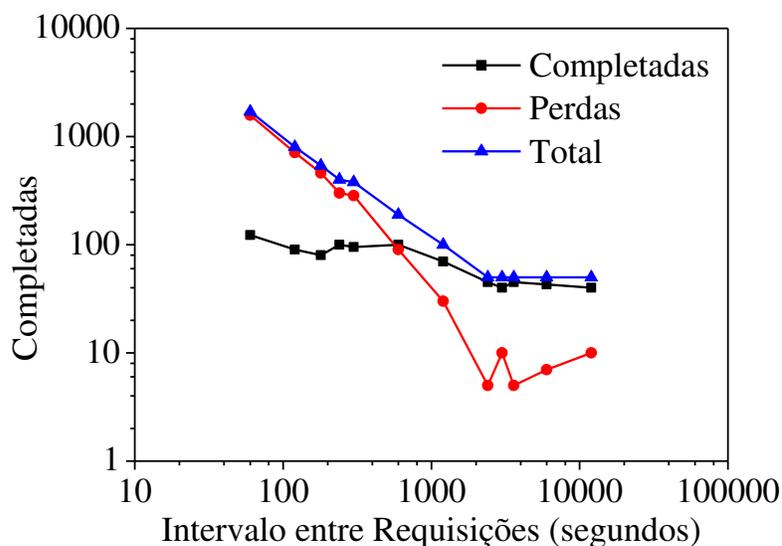


Figura 40 - Quantidade de videoconferências em simulações com intervalo de duração variável, jitter e latência.

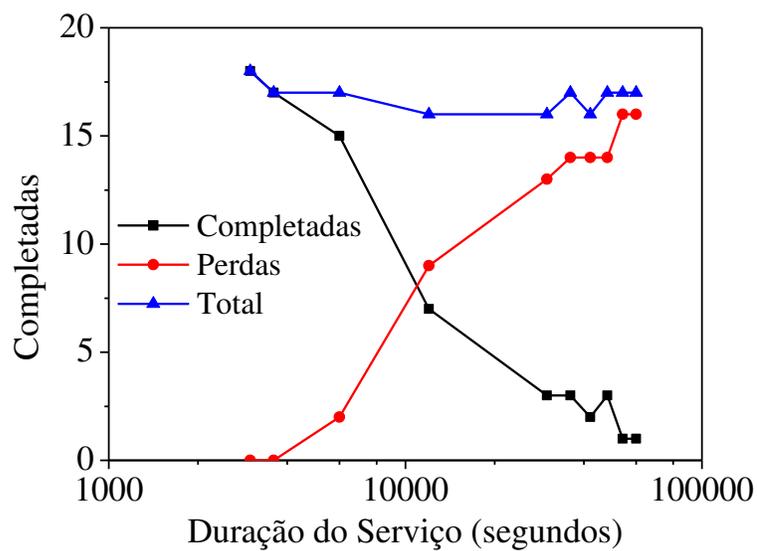


Figura 41 - Quantidade de videoconferências em simulações com duração de serviço variável, *jitter* e latência.

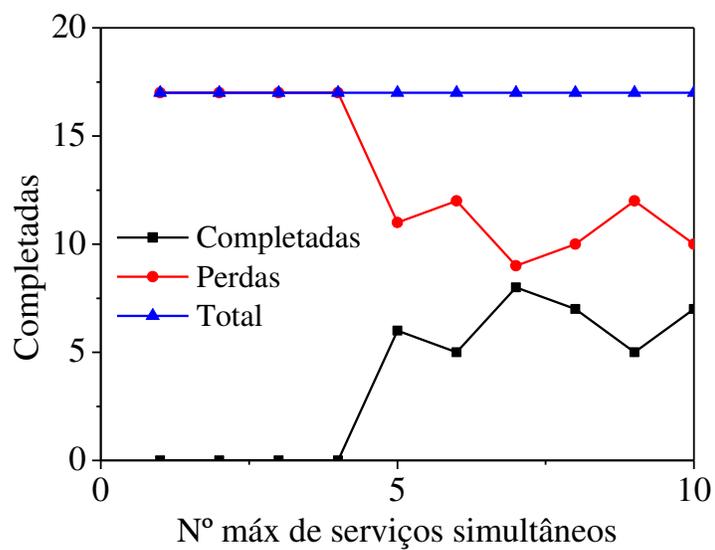


Figura 42 - Quantidade de videoconferências em simulações variação do número máximo de serviços, *jitter* e latência.

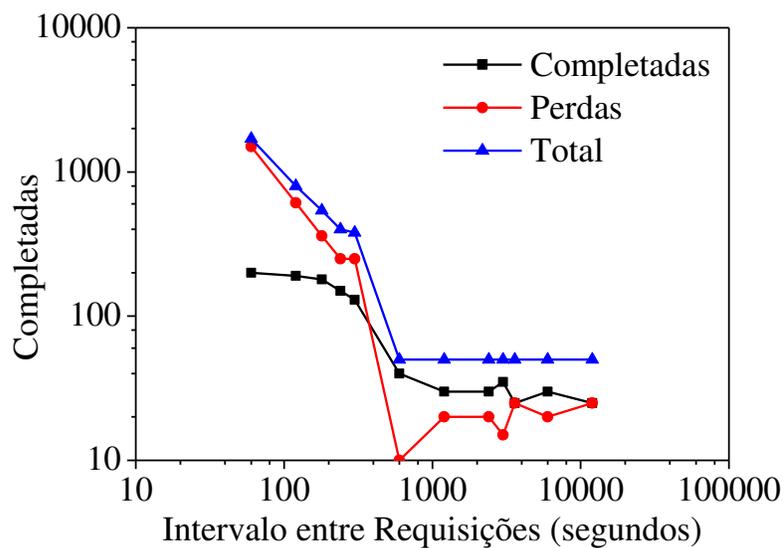


Figura 43 - Quantidade de videoconferências em simulações considerando a tarifação por nº de chamadas, *jitter* e latência.

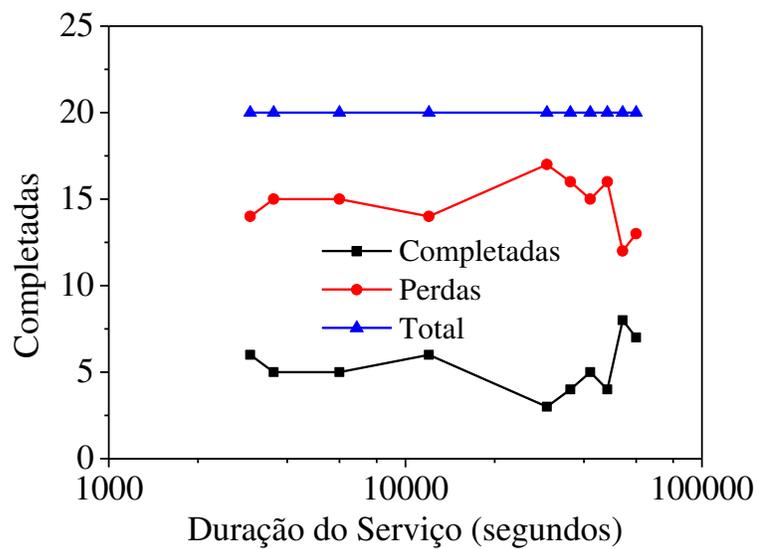


Figura 44 - Quantidade de videoconferências em simulações com intervalo de requisição fixa considerando a tarifação por nº de chamadas, *jitter* e latência.

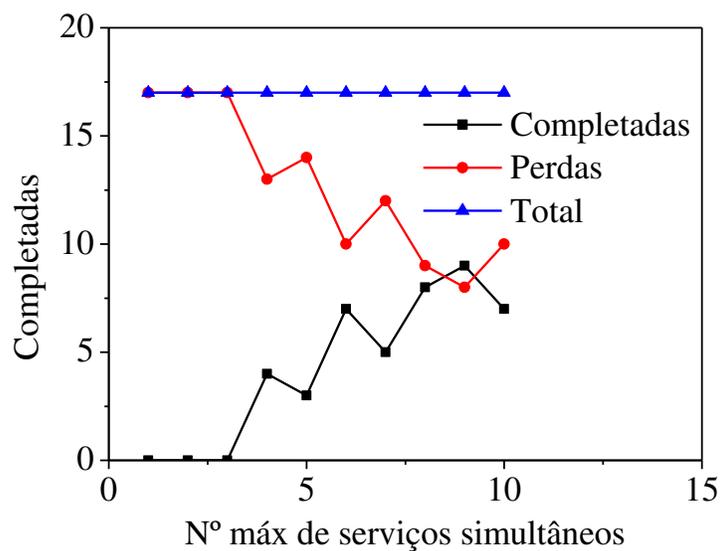


Figura 45 - Quantidade de videoconferências em simulações com intervalo de requisição e duração média fixa considerando a tarifação por nº de chamadas, *jitter* e latência.

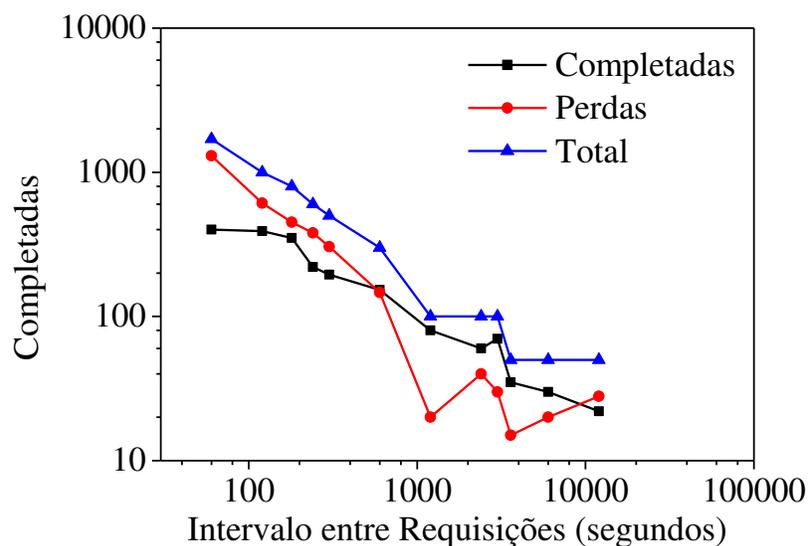


Figura 46 - Quantidade de videoconferências em simulações considerando a tarifação fixa, *jitter* e latência.

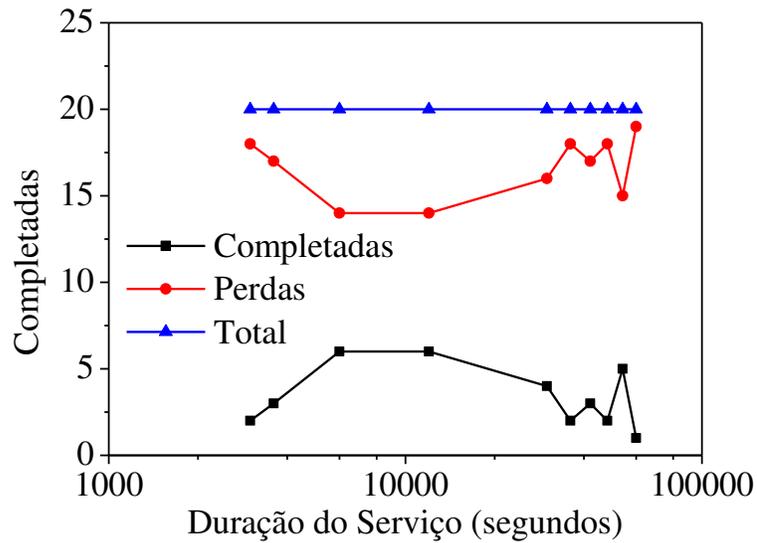


Figura 47 - Quantidade de videoconferências em simulações com intervalo de requisição fixa considerando a tarifação fixa, *jitter* e latência.

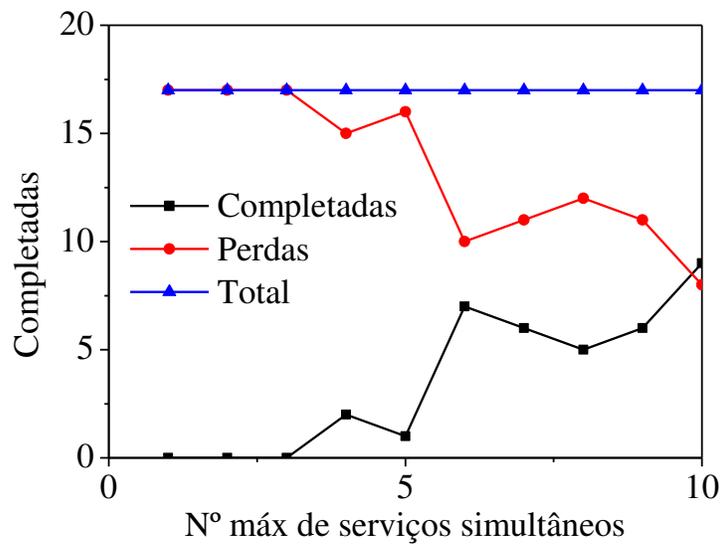


Figura 48 - Quantidade de videoconferências em simulações com intervalo de requisição e duração média fixa considerando a tarifação fixa, *jitter* e latência.

5 CONCLUSÕES E PERSPECTIVAS

Neste trabalho foram analisados dois modelos de simulação considerando diferentes tipos de tarifação para redes 3G. As simulações foram realizadas utilizando o *software* ARENA, que apesar de não ser um *software* específico para simulação de tráfegos de redes, atendeu todas as necessidades das simulações. Foi utilizada uma versão gratuita para estudantes do simulador na versão 12.0. A principal diferença entre essa versão e a versão completa é que ela possui um limite de entidades simultâneas. Devido a essa restrição algumas simulações que gostaríamos de ter executado não foi possível, porque ultrapassavam as 150 entidades que a versão gratuita permite.

O tempo de simulação variou de 1 hora a 10 horas para cada simulação, dependendo dos parâmetros utilizados. No total foram executadas 54 simulações para os dois modelos propostos, cada qual com três diferentes tipos de tarifação. Em cada modelo foram utilizados diferentes valores de replicações. Quanto maior o número de replicações, mais precisos foram os resultados. Por outro lado, o seu aumento resulta em tempos maiores para o término da simulação, inviabilizando simular um número grande de diferentes cenários e parâmetros.

Foram propostos dois diferentes modelos de simulação que simulam um tráfego multimídia, onde serviços *stream* e elástico são requisitados ao mesmo tempo pelos usuários. A função de um modelo é descrever o funcionamento da realidade por meio de um número de parâmetros, tão pequeno quanto possível, que permita a sua compreensão. O primeiro modelo possui três blocos *Arrive*, que são os responsáveis pela geração dos serviços *stream*: *videoconference*, *video-clip*, e *video on demand*. Cada bloco possui valores que representam seus serviços e funções de distribuição. Para cada serviço há duas saídas, serviços rejeitados e serviços completados. O modelo 2 possui seis serviços: *videoconference*, *video-clips*, *video on demand*, VoIP, câmera IP, e *data file* (serviço de Dados), sendo este o único serviço do tipo elástico.

No modelo 1 foram realizadas simulações propondo um estudo de demanda dos usuários, onde três cenários foram estudados. Esses tipos de situações são amplamente estudados em casos

de mudança do comportamento na tarifação das empresas de telefonia móvel e fixa, onde a cobrança não é mais feita por minutos utilizados, mas sim por chamadas feitas aumentando significativamente a duração das chamadas. Os estudos de demanda são cruciais por parte das empresas para não acarretar em um baixo *QoS* (infringindo as regras da ANATEL) e bloqueios de chamadas elevado. Dos resultados obtidos podemos notar que foi possível descrever qualitativamente o comportamento da rede modelada, ainda que quantitativamente os modelos não tenham precisão.

No modelo 2 os parâmetros de *jitter* e da latência mudaram significativamente o comportamento visto no modelo 1. Em todos os cenários simulados neste modelo houve um aumento significativo na perda de serviços multimídia completados. Além disso, mesmo aumentando o volume do tráfego *stream*, o tempo do tráfego elástico em muitos casos não aumentou. Isso se deve à influência do *jitter* e da latência.

As simulações mostraram o colapso no desempenho da rede ao aumentar substancialmente o tráfego para o mesmo dimensionamento, e como a rede deve ser redimensionada para que o desempenho seja mantido mesmo com o tráfego alterado (adição de mais canais).

O trabalho apresentou uma análise qualitativa, ou seja, os modelos não são precisos quantitativamente e os resultados mostram a tendência esperada. Isso mostra que a nossa abordagem é adequada para modelar o tráfego multimídia em redes de comutação por pacotes. Para o uso em problemas de dimensionamento, contudo, é necessária uma calibração dos modelos.

Em trabalhos futuros sugere-se que os modelos de simulações utilizados nesse trabalho sejam estendidos para os descrever as redes de acesso 4G, assim como a utilização da versão completa do *software* ARENA, pois proporcionará estudos com volumes de tráfegos bem maiores sem restrição nos parâmetros dos modelos. Muitos casos interessantes não foram simulados devido às restrições que a versão de gratuita do ARENA tem quanto ao tamanho dos modelos.

REFERÊNCIAS

[1] TRAN-GIA, P.; STAEHLE, D.; LEIBNITZ, K. Source traffic modeling of wireless applications. **International Journal of Electronics and Communications**, v. 55, n. 1, p. 27-36, jan. 2001.

[2] GOLLAUP, A.; AGHVAMI, H. A multimedia traffic-modeling framework for simulation-based performance evaluation studies. **Computer Networks**, v. 50, n. 12, p. 2071-2087, aug. 2006.

[3] ALQAHTANI, S. A. A Simulation-Based Comparison of Multimedia Traffic Prioritization Schemes for High-Performance Input-Queued Packet Switches. **Journal of Computer Science**, v. 2, n. 4, p. 347-354, 2006.

[4] ZHANG, D.; FANG, X.; ZHU, L. Novel Multimedia Traffic Modeling based CAC Scheme for CDMA Communication Systems. **Journal of Electronics (China)**, v. 24, n. 1, p. 39-45, jan. 2007.

[5] LAZARIS, A.; KOUTSAKIS, P.; PATERAKIS, M. A new model for video traffic originating from multiplexed MPEG-4 videoconference streams. **Performance Evaluation: an international journal**, v. 65, p. 51-70, 2008.

[6] ETSI. **Universal Mobile Telecommunications System (UMTS) Radio Resource Control (RRC) Protocol specification (3GPP TS 25.331. Version 9.1.0, Release 9)**. France: European Telecommunications Standards Institute, 2010.

[7] KELTON, W.D.; SADOWSKI, R.P., SADOWSKI, D.A. **Simulation with ARENA**. Boston: McGraw-Hill Higher Education, 2001.

[8] PINOTTI, F. L. [et al.]. An IP-based multimedia traffic generator. In: INTERNATIONAL WORKSHOP ON TELECOMMUNICATIONS, 2011. **Proceedings...** Brasil: Instituto Nacional de Telecomunicações, 2011; PINOTTI, F. L., Simulação e emulação de tráfego multimídia em redes IP, Tese de Mestrado, UNICAMP (2011).

[9] OLIVEIRA, T. R. B.; URSINI, E. L.; TIMÓTEO, V. S. Simulation based assessment of energy consumption optimization in 3G mobiles. In: INTERNATIONAL CONFERENCE ON COMPUTER AND MANAGEMENT, 2011. **Proceedings...** China: IEEE Computer Society, 2011.

[10] OLIVEIRA, T. R. B.; URSINI, E. L.; TIMÓTEO, V. S. Simulation inspired model for energy consumption in 3G always-on mobiles. In: NATIONAL CONFERENCE ON TELECOMMUNICATIONS, 2. 2011. **Proceedings...** Arequipa, Peru: Universidad Católica San Pablo, IEEE, 2011.

[11] RAPPAPORT, T. S. **Comunicações sem fio: princípios e práticas**. São Paulo: Pearson Prentice Hall, 2009.

[12] ANATEL [página da Internet]. Disponível em: <<http://www.anatel.gov.br>>. Acesso em: 10 nov. 2012.

[13] EDGE, HSPA and LTE: The Mobile Broadband Advantage. USA: 3G Americas, 2007.

[14] DEFINING 4G: understanding ITU process for the next generation of wireless technology. USA: 3G Americas, 2008.

[15] HOLMA, H.; TOSAKALA, A. (Ed.). **WCDMA for UMTS**: radio access for third generation mobile communications. New Jersey: Wiley, 2004.

[16] _____. **HSDPA/HSUPA for UMTS**: high speed radio access for mobile communications. New Jersey: Wiley, 2006.

[17] UMTS Evolution: from 3GPP release 7 to release 8. USA: 3G Americas, 2008.

[18] OJANPERA, T. Overview of multiuser detection/interference cancellation for DS-CDMA. In: IEEE INTERNATIONAL PERSONAL WIRELESS COMMUNICATIONS, 1997. **Proceedings...** Mumbai, India: IEEE, 1997.

[19] LOPES, E. M. TP-100: sistemas de comunicações móveis e TV digital. [Apresentação INATEL].

[20] PRADO, D. **Usando o ARENA em simulação**: v. 3. Belo Horizonte, MG: IDNG, 1999.

[21] _____. **Planejamento e Controle de Projetos**. Belo Horizonte, MG: IDNG, 1999.

[22] TELECO [página da Internet]. Disponível em: <<http://www.teleco.com.br>>. Acesso em: 25 out. 2012.