

Universidade Federal do Pará  
Instituto de Tecnologia  
Programa de Pós-Graduação em Engenharia Elétrica

Processo Markoviano de Decisão para Alocação  
Dinâmica de Recursos e Controle de Admissão de  
Conexão em Redes IEEE 802.16

Cynthia Feitosa Leal

DM\_02/2010

UFPA/ITEC/PPGEE  
Campus Universitário do Guamá  
Belém-Pará-Brasil

2010

Processo Markoviano de Decisão para Alocação  
Dinâmica de Recursos e Controle de Admissão de  
Conexão em Redes IEEE 802.16

Cynthia Feitosa Leal

Orientador:

Prof. Dr. João Crisóstomo Weyl A. Costa

Coorientador:

Prof. Dr. Solon Venâncio de Carvalho

*Dissertação de Mestrado submetida à Banca Examinadora do Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Pará como requisito de obtenção do título de “Mestre em Engenharia Elétrica”.*

UFPA/ITEC/PPGEE

Campus Universitário do Guamá

Belém-Pará-Brasil

2010

---

L422p

Leal, Cynthia Feitosa

Processo markoviano de decisão para alocação dinâmica de recursos e controle de admissão de conexão em redes IEEE 802.16 / Cynthia Feitosa Leal; orientador, João Crisóstomo Weyl Albuquerque Costa.-2010.

Dissertação (Mestrado) – Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-graduação em Engenharia Elétrica, Belém, 2010.

1.Redes locais sem fio – normas. 2. Sistemas de comunicação em banda larga. 3. Markov, processos de. 3. Redes metropolitanas de computação I. Orientador. II. Título.

CDD 22. ed. 004.68

---

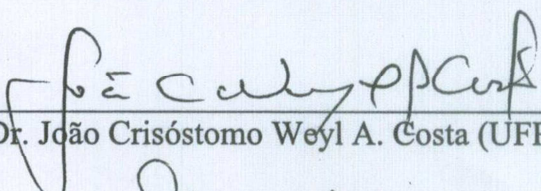
Cynthia Feitosa Leal

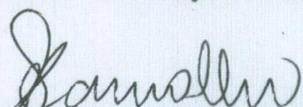
PROCESSO MARKOVIANO DE DECISÃO PARA ALOCAÇÃO  
DINÂMICA DE RECURSOS E CONTROLE DE ADMISSÃO DE  
CONEXÃO EM REDES IEEE 802.16

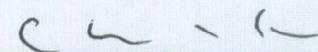
Esta Dissertação foi julgada adequada à obtenção do título de Mestre em Engenharia Elétrica e aprovada em sua forma final pelo Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Pará.

Belém, 08 de Fevereiro de 2010.

BANCA EXAMINADORA:

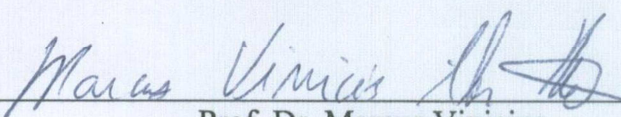
  
\_\_\_\_\_  
Prof. Dr. João Crisóstomo Weyl A. Costa (UFPA) – Orientador

  
\_\_\_\_\_  
Prof. Dr. Solon Venâncio de Carvalho (INPE) – Coorientador

  
\_\_\_\_\_  
Prof. Dr. Carlos Renato Lisboa Francês (UFPA) – Membro

  
\_\_\_\_\_  
Prof. Dr. Antônio Jorge Gomes Abelém (UFPA) – Membro externo

Visto:

  
\_\_\_\_\_  
Prof. Dr. Marcus Vinicius  
Coordenador do PPGEE/ITEC/UFPA

Dedico este trabalho à minha família, amigos, professores e todas as outras pessoas que apoiaram e incentivaram mesmo nas dificuldades e que, de uma forma ou outra, me fizeram crescer

# Agradecimentos

Algumas pessoas marcam a nossa vida para sempre, umas porque nos vão ajudando na construção, outras porque nos apresentam projetos de sonhos e outras ainda porque nos desafiam a construí-los. Quando damos conta, já é tarde para lhes agradecer.

Dedico este trabalho a Todas as pessoas que, sem saberem (?), muito para ele contribuíram, em especial.

À minha Mãe, pelo estímulo e apoio incondicional desde o primeiro momento; pela paciência e grande amizade com que sempre me ouviu, e sensatez com que sempre me aconselhou.

Aos professores Solon Venâncio de Carvalho e João Crisóstomo Weyl A. Costa pela competência com que me orientaram nesta dissertação e pelo tempo que generosamente me dedicaram, transmitindo-me os melhores e mais úteis ensinamentos, com paciência, lucidez e confiança. Pelo acesso que me facilitaram a uma pesquisa mais alargada e enriquecedora e pela suas críticas sempre tão atempada, como construtiva.

A todas as pessoas, que me fizeram companhia durante essa etapa de minha vida. A Lilian e Cleide que mesmo distantes estiveram e estarão sempre presentes em meu coração. Em especial a Adriana e Eduardo, que compartilharam e me ajudaram de perto em toda esta etapa de minha vida, sempre me apoiando.

Agradeço, sobretudo, a Deus por ter colocado essas pessoas em meu caminho que em muito contribuíram para o que hoje sou.

A TODOS, o meu sincero muito obrigada!

# Resumo

Este trabalho apresenta uma solução para o problema de controle admissão de conexão e alocação dinâmica de recursos em redes IEEE 802.16 através da modelagem de um Processo Markoviano de Decisão (PMD) utilizando o conceito de degradação de largura de banda, o qual é baseado nos requisitos diferenciados de largura de banda das classes de serviço do IEEE 802.16. Para o critério de desempenho do PMD é feita a atribuição de diferentes retornos a cada classe de serviço, fazendo assim o tratamento diferenciado de cada fluxo. Nesse sentido, é possível avaliar a política ótima, obtida através de um algoritmo de iteração de valores, considerando aspectos como o nível de degradação médio das classes de serviço, utilização dos recursos e probabilidades de bloqueios de cada classe de serviço em relação à carga do sistema. Resultados obtidos mostram que o método de controle markoviano proposto é capaz de priorizar as classes de serviço consideradas mais relevantes para o sistema.

# Abstract

This work presents a solution to the problem of connection admission control and dynamic resource allocation in IEEE 802.16 networks by modeling a Markov Decision Process (MDP) using the concept of bandwidth degradation, which is based on different bandwidth requirements of IEEE 802.16 service classes. In order to test the performance of the MDP, different returns for each class of service are allocated, thus making the differential treatment of each service class. Therefore, it is possible to evaluate the optimal policy, obtained through a value iteration algorithm, considering aspects such as the service class average adjustment, resource utilization and blocking probability in relation to system load. Results obtained show that the Markov control method proposed is able to prioritize service classes considered most relevant to the system.



# Sumário

## Lista de Figuras

## Lista de Tabelas

## Glossário

<b>Introdução.....</b>	<b>1</b>
1.1 Motivação .....	1
1.2 Objetivo .....	2
1.3 Contribuições .....	2
1.4 Organização da Dissertação.....	3
<b>Capítulo 2 - Tecnologia e Trabalhos Correlatos.....</b>	<b>4</b>
2.1 Introdução .....	4
2.2 O Padrão IEEE 802.16.....	4
2.2.1 Camada Física – PHY .....	6
2.2.2 Camada MAC.....	9
2.3 Trabalhos Correlatos.....	15
2.4 Conclusão.....	16
<b>Capítulo 3 - Controle de Admissão de Conexão.....</b>	<b>17</b>
3.1 Introdução .....	17
3.2 Controle de Admissão de Conexão.....	17
3.3 Esquema de Degradação de Largura de Banda.....	19
3.4 Função Utilidade.....	20
3.4.1 Tráfego de tempo real não adaptativo .....	20
3.4.2 Tráfego de tempo real adaptativo .....	21
3.4.3 Tráfego de tempo não real.....	23
3.4.4 Quantização das funções utilidade .....	25
3.5 Conclusão.....	25
<b>Capítulo 4 - Formulação do CAC como um PMD.....</b>	<b>26</b>
4.1 Introdução .....	26
4.2 Processos Semi-Markovianos de Decisão e Processos Markovianos de	

Decisão a Tempo Contínuo.....	26
4.3 O Algoritmo de Iteração de Valores .....	30
4.4 Formulação do CAC como um PMDTC .....	32
4.4.1 Espaço de Estados .....	33
4.4.2 Dinâmica dos Estados e Ações.....	33
4.4.3 Critério de Otimização .....	35
4.5 Conclusão.....	36
<b>Capítulo 5 - Resultados Numéricos .....</b>	<b>38</b>
5.1 Introdução .....	38
5.2 Método e Parâmetros .....	38
5.3 Resultados e Discussões .....	40
5.3.1 PMD para CAC com Degradação .....	41
5.3.2 Análise da Degradação de Largura de Banda.....	45
5.4 Conclusões .....	48
<b>Capítulo 6 - Conclusões e Trabalhos Futuros .....</b>	<b>49</b>
6.1 Conclusões .....	49
6.2 Trabalhos Futuros .....	50
<b>Referências Bibliográficas .....</b>	<b>51</b>

# Lista de Figuras

Figura 2.1 – Pilha de protocolo do IEEE 802.16. ....	6
Figura 2.2 – Esquema adaptativo em função das condições do canal [10].....	8
Figura 2.3 – Arquitetura de QoS do IEEE 802.16. ....	13
Figura 3.1 – Rede IEEE 802.16 PMP .....	19
Figura 3.2 – Função utilidade para tráfego da classe 1.....	21
Figura 3.3 – Função utilidade para tráfego da classe 2.....	22
Figura 3.4 – Função utilidade para tráfego da classe 3.....	24
Figura 3.5 – Quantização das funções utilidade com intervalos de largura de banda iguais. ....	25
Figura 4.1 – Algoritmo de Iteração de Valores.....	31
Figura 5.1 – Funções Utilidade das Classe de Serviço 1, 2 e 3.....	40
Figura 5.2 – Largura de Banda Utilizada em Percentagem .....	42
Figura 5.3 – Nível Médio de Degradação por Classe de serviço.....	43
Figura 5.4 – Probabilidade de Bloqueio por Classe de Serviço.....	43
Figura 5.5 – Largura de Banda Utilizada em Percentagem .....	44
Figura 5.6 – Largura de Banda Média por Conexão.....	45
Figura 5.7 – Utilização de Recursos por Esquema .....	46
Figura 5.8 – Probabilidade de Bloqueio por Esquema.....	47
Figura 5.9 – Largura de Banda Média por Esquema .....	48

# Lista de Tabelas

Tabela 4.1 – Estado Reagido.....	35
Tabela 4.2 – Próximo Estado e Transição Associada .....	35
Tabela 5.1 – Parâmetros do Sistema .....	39
Tabela 5.2 – Função Utilidade das Classes de Serviço.....	40
Tabela 5.3 – Caracterização da Política Ótima. ....	41

# Glossário

QoS	Quality Of Service
CAC	Controle De Admissão De Conexão
PMD	Processo Markoviano de Decisão
WiMAX	World Interoperability for Microwave Access
BS	Base Station
SS	Subscriber Station
WLAN	Wireless Local Area Network
WAN	Wide Area Network
LMDS	Local Multipoint Distribution System
xDSL	Digital Subscriber Line
MIMO	Multiple-Input Multiple-Output
WMAN	Wireless Metropolitan Area Network
LOS	Line of Sight
NLOS	Non Line of Sight
ISM	Industrial Scientific And Medical
MIB	Management Information Base
4G	Fourth Generation
MAC	Medium Access Control
CPS	Common Part Sublayer
CS	Convergence Sublayer
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
TDD	Time Division Duplexing
FDD	Frequency Division Duplexing
BPSK	Binary Phase Shift Keying
QPSK	Quadrature Phase Shift Keying
QAM	Quadrature Amplitude Modulation
AMC	Adaptative Coding And Modulation
FFT	Fast Fourier Transform

TDMA	Time Division Multiple Access
TDM	Time-Division Multiplexing
UNII	Unlicensed National Information Infrastructure
CS-SAP	Cs - Service Access Point
ATM	Asynchronous Transfer Mode
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
VLAN	Ethernet E Virtual Local Area Network
SDUs	Service Data Units
CID	Conection Identificator
GPSS	Grant Per Subscriber Station
GPC	Grant Per Connection
CBR	Constant Bit Rate
UGS	Unsolicited Grant Service
ertPS	Extended Real-Time Polling Service
rtPS	Real-Time Polling Service
nrtPS	Non Real-Time Polling Service
BE	Best Effort
CS	Complete Sharing
CP	Complete Partitioning
PMP	Ponto Multiponto

# Introdução

---

## 1.1 Motivação

No contexto amazônico, a ampliação real dos serviços de cidadania é dificultada pela falta de infra-estrutura básica adequada em regiões que, por questões históricas e/ou geográficas, acabaram sendo relegadas ao isolamento e atraso tecnológico. Para que esse quadro seja revertido é necessário que se fomente o desenvolvimento e inclusão digital dessas regiões. Nesse sentido, é de suma importância buscar soluções viáveis para o acesso de última milha, o qual permitirá que essas regiões, outrora isoladas ou de acesso precário, se conectem aos *backbones* centrais das linhas de comunicação. Entretanto, vários fatores contribuem para dificultar o acesso de última milha na região amazônica, como: as grandes distâncias entre o ponto de acesso e os *backbones* centrais, condições geográfico-climáticas inerentes à região e finalmente o desinteresse comercial das operadoras concessionárias envolvidas.

O padrão IEEE 802.16 surge como uma solução promissora para o problema de acesso de última milha no contexto apresentado, pois disponibiliza base tecnológica que dá suporte às grandes demandas das aplicações banda de larga com cobertura de rádio frequência relativamente estendida, possibilitando a cobertura de distâncias maiores sem a necessidade de investimento em infra-estrutura de alto custo como ocorre em redes de acesso banda larga cabeada, reduzindo desta forma, o custo de implantação e do tempo necessário para se conectar residências, escolas e escritórios aos *backbones* das linhas de comunicação.

O padrão IEEE 802.16 permite ainda tratamento diferenciado dos fluxos de tráfego possibilitando o provimento garantido de qualidade de serviço (QoS), em termos de atraso, *jitter* e taxa de transferência. Esta característica é de grande importância para aplicações como as de tele-educação, consideradas estratégicas para a região amazônica no sentido de promover conhecimento e capacitação através de núcleos de educação à distância. No contexto do provimento de QoS, o controle de admissão de conexão (CAC) é estratégico, pois possibilita o controle da utilização dos recursos da rede, entretanto não foi especificado pelo padrão IEEE 802.16. A principal função do CAC é decidir adequadamente se canal de

comunicação deve ou não aceitar uma nova conexão. Caso a política do controle aceite um número excessivo de conexões o sistema não terá com garantir a QoS das conexões existentes, entretanto caso a política de controle admita um número muito pequeno de conexões, ou seja, rejeite muitas conexões, pode ocorrer um desperdício dos recursos da rede. Objetivando a maximização de utilização de recursos da rede pode-se utilizar o conceito de degradação de largura de banda, o qual é baseado nos requisitos diferenciados de largura de banda das classes de serviço do IEEE 802.16. Assim, combinação do CAC com esquema de degradação de largura de banda torna a alocação de recursos mais eficiente, porém, os projetos desses esquemas ficam cada vez mais complexos. Nesta dissertação, esses assuntos são abordados através de modelagem markoviana, dada sua capacidade de produzir soluções rápidas e precisas de sistemas com ênfase em otimização.

## 1.2 Objetivo

O objetivo deste trabalho, no contexto de acesso de ultima milha, é estudar o problema de controle de admissão de conexões em redes IEEE 802.16, buscando propor uma solução otimizada que combine um esquema degradação de largura de banda e que atente para a qualidade de serviço percebida pelo usuário final. A solução proposta deve ainda ser adequada para análise de métricas relevantes no projeto da rede, e ter como objetivo atender o maior número de usuários com a melhor QoS possível.

## 1.3 Contribuições

As contribuições técnicas deste trabalho são:

Estudo sobre modelagem e análise de desempenho de redes IEEE 802.16 destacando:

- Controle de admissão de conexão (CAC) em redes ponto multiponto (PMP) combinado com esquema de degradação de largura de banda, levando em consideração a QoS percebida pelo usuário final;
- Proposta de modelo analítico para a representação e avaliação de desempenho desses procedimentos no controle de admissão de conexão;
- Otimização modelo proposto para o esquema de alocação de recursos e CAC através de um processo markoviano de decisão.



## 1.4 Organização da Dissertação

O texto desta dissertação está organizado da seguinte maneira:

Capítulo 2: Abordagem teórica sobre IEEE 802.16 e trabalhos correlatos;

Capítulo 3: Trata do problema de controle de admissão e apresenta conceitos utilizados na modelagem do sistema;

Capítulo 4: Apresenta uma proposta de modelagem do problema de controle de admissão como processo markoviano de decisão;

Capítulo 5: Apresenta os resultados obtidos através da simulação numérica do processo markoviano de decisão para várias cargas do sistema;

Capítulo 6: Traz as conclusões do trabalho e sugestões para melhorias e trabalhos futuros.

## Capítulo 2 - Tecnologia e Trabalhos Correlatos

---

### 2.1 Introdução

Este capítulo apresenta uma descrição sucinta do padrão IEEE 802.16 destacando os conceitos-chaves desta tecnologia de acesso no que se refere ao controle de admissão de conexão, e ao final apresenta um levantamento bibliográfico de como o problema de controle de admissão de conexão vem sendo tratado na literatura e limites das soluções atuais.

### 2.2 O Padrão IEEE 802.16

O padrão IEEE 802.16, popularmente conhecido por WiMAX, acrônimo para *World Interoperability for Microwave Access*, ostenta um alcance teórico de 50Km em áreas rurais, de 10Kms em áreas suburbanas e cerca de 5Km em densas áreas urbanas com taxas médias de transferência de até 100Mbps [1-5]. A versão inicial do IEEE 802.16, publicada em abril de 2002, operava nas frequências de 10 a 66 GHz e operava com linha de visada direta (LOS – *Line Of Sight*). A extensão 802.16a, aprovada em janeiro de 2003, adicionou características que possibilitaram transmissões sem linha de visada direta (NLOS – *Non Line Of Sight*) e permitiram o uso de frequências mais baixas (2 a 11 GHz), muitas das quais não são licenciadas, tornando o padrão mais atrativo comercialmente. As emendas subsequentes ao padrão, como a IEEE 802.16e, publicada em dezembro de 2005, possibilitaram que uma única estação base (BS) oferecesse cobertura tanto para estações assinantes (SS) fixas quanto móveis. Essas correções ajudaram a preencher a lacuna entre as altas taxas de dados das redes locais sem fio (WLAN) e a alta mobilidade celular das redes metropolitanas (WAN). A seguir, são descritos de forma sucinta a família de padrões que compõem o IEEE 802.16[1, 6-8]:

**IEEE 802.16:** Corresponde a especificação original, projetado para padronizar implementações LMDS (*Local Multipoint Distribution System*). Desenvolvido para trabalhar

em frequências de 10 –66 GHz

**IEEE 802.16a:** Atende frequências mais baixas (2 - 11 GHz). Especificado para competir com as tecnologias que oferecem acesso à última milha, como xDSL e cable modems. Pode obter taxas de transmissão de até 75 Mbps com um alcance máximo de 50Km.

**IEEE 802.16b:** Aborda aspectos relativos à qualidade de serviço para operação na faixa de frequência ISM (*Industrial Scientific and Medical*) de 5 GHz.

**IEEE 802.16c:** Trata de interoperabilidade através dos perfis de sistemas na range de 10 GHz a 66 GHz, protocolos e especificação de testes de conformação.

**IEEE 802.16-REVd:** Atualização do padrão 802.16, a qual consolidou as revisões dos padrões 802.16a e 802.16c em um único padrão, substituindo o 802.16a como o padrão base. Destaca-se pela provisão de suporte para antenas MIMO (*Multiple-InputMultiple-Output*), possibilitando maior confiabilidade do alcance com multipercurso. Facilitando instalações de antenas internas (indoor) por não necessitar de linha de visada direta entre antenas.

**IEEE 802.16e:** Acrescenta especificações de mobilidade em WMANs com LOS ou NLOS, em frequências de 10-66 GHz e 2-11 GHz, respectivamente. Além dessas frequências, o IEEE 802.16e também oferece suporte à mobilidade em frequências entre 2-6 GHz.

**IEEE 802.16f:** Apresenta uma introdução do conceito de redes em malha (*mesh networks*).

**IEEE 802.16g:** Descreve melhoras para o suporte à mobilidade.

**IEEE 802.16h:** Possibilita o suporte à contenção de acesso ao meio que permite a operação em bandas ISM na faixa de 2,4 GHz e 5,8 GHz.

**IEEE 802.16i:** Introduce o conceito de base de informações de gerência (MIB) especificando quais variáveis são mantidas pelos elementos da rede.

**IEEE 802.16j:** Especifica a operação em saltos múltiplos com retransmissão (*Multihop Relay Specification*) e interoperabilidade entre estações retransmissoras (*Relay Stations*) e estações base.

**IEEE 802.16k:** Propõe a união do padrão IEEE 802.1D (padrão de *bridge*

transparente) no reconhecimento do padrão IEEE 802.16.

**IEEE 802.16m:** Descreve uma interface aérea avançada que permite a compatibilidade entre o IEEE 802.16REVd (nômade) e futuras redes de quarta geração (4G). Prever cinco especificações de velocidades de transmissão. Uma pra transferências a 16 Kbps. Outra pra transferências de dados e multimídia a 144 Kbps e três categorias de tráfego em multimídia com 2 Mbps, 30 Mbps e uma que pode alcançar até 1 Gbps.

A provisão de acesso banda larga sem fio pelo IEEE 802.16 é realizada através da especificação de uma pilha de protocolos, a qual foi definida de forma semelhante à adotada na maioria das famílias de padrões do IEEE 802, porém, contendo um número maior de subdivisões. Assim, a pilha de protocolo do IEEE 802.16 é composta por duas camadas principais: a camada PHY (Física) e a camada MAC (*Medium Access Control*), conforme é ilustrado na Figura 2.1, sendo a que camada de enlace de dados é dividida em três subcamadas: subcamada de segurança (*Security Sublayer*), subcamada de convergência comum (CPS - *Common Part Sublayer*) e subcamada de convergência de serviços específicos (CS - *Convergence Sublayer*). A seguir, descrevem-se cada uma das camadas do padrão 802.16.

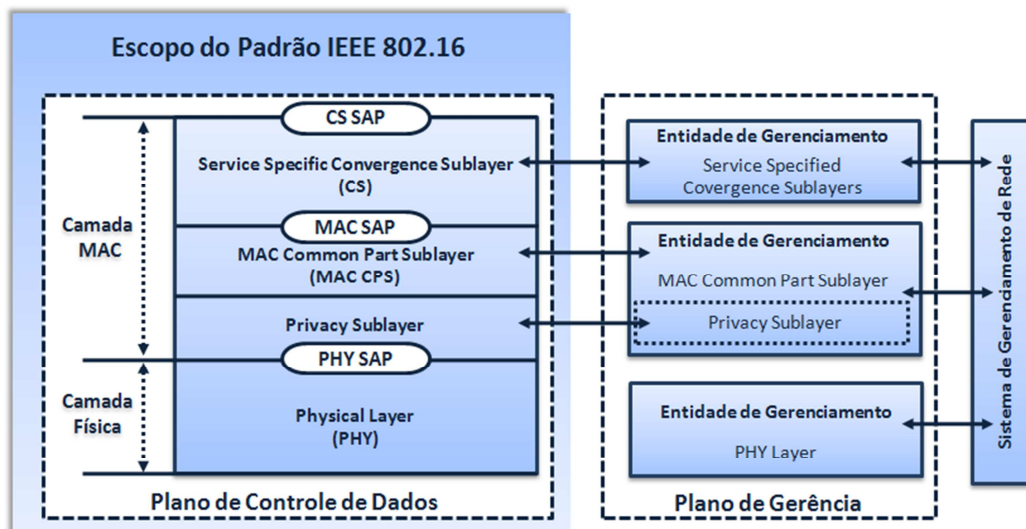


Figura 2.1 – Pilha de protocolo do IEEE 802.16.

### 2.2.1 Camada Física – PHY

O IEEE 802.16 faz uso de várias tecnologias já conhecidos para atingir seu objetivo de fornecer máximo *throughput* a máximas distâncias com alta confiabilidade. Dentre estas

tecnologias pode-se citar: *Orthogonal Frequency Division Multiplexing* (OFDM), *Time Division Duplexing* (TDD), *Frequency Division Duplexing* (FDD), *Quadrature Phase Shift Keying* (QPSK) e *Quadrature Amplitude Modulation* (QAM). Todas são tecnologias de camada física, conhecidas e utilizadas em outras soluções que foram combinadas tornando o IEEE 802.16 um grande avanço na área de redes metropolitanas sem fio. A camada física pode ser descrita com maiores detalhes, analisando as especificações diferentes interfaces aéreas suportadas. Assim, a seguir descreve-se com mais detalhes cada uma destas especificações.

### **WirelessMAN-SC PHY**

Esta especificação da camada física é baseada no esquema de modulação *single carrier*, foi projetada para trabalhar na faixa de frequência de 10-66GHz, em transmissões do tipo ponto-a-ponto, com linha de visada direta (LOS), suporta dois tipos de métodos de duplexização, o FDD e o TDD. Podem-se utilizar vários perfis de transmissão adaptativa ou modulação e codificação adaptativa (AMC - *Adaptive Coding and Modulation*), onde os parâmetros de transmissão podem ser ajustados de acordo com cada estação [9]. Este esquema de AMC possui três tipos de modulações diferentes: o *Quadrature Amplitude Modulation* (QAM-64), *Quadrature Amplitude Modulation* (QAM-16) e *Quadrature Phase Shift Keying* (QPSK). A escolha de uma modulação está condicionada à qualidade do enlace. Nos casos onde se deseja qualidade do enlace elevada, o perfil escolhido é o QAM-64. Onde os enlaces requerem estabilidade e qualidade da conexão, o perfil ideal é o QPSK e para enlaces onde há possibilidade de atenuação de sinal é utilizado o QAM-16. Esse sistema de perfil também leva em conta a distância da SS até a BS. SS's que estiverem mais longe da BS, utilizarão o esquema de modulação QPSK, SS's que estiverem a uma distância mediana da BS, utilizará o esquema QMA-16 e para as SSs que estiverem a uma distância mais curta, o esquema a ser aplicado é o QAM-64, conforme Figura 2.2 [9, 10].

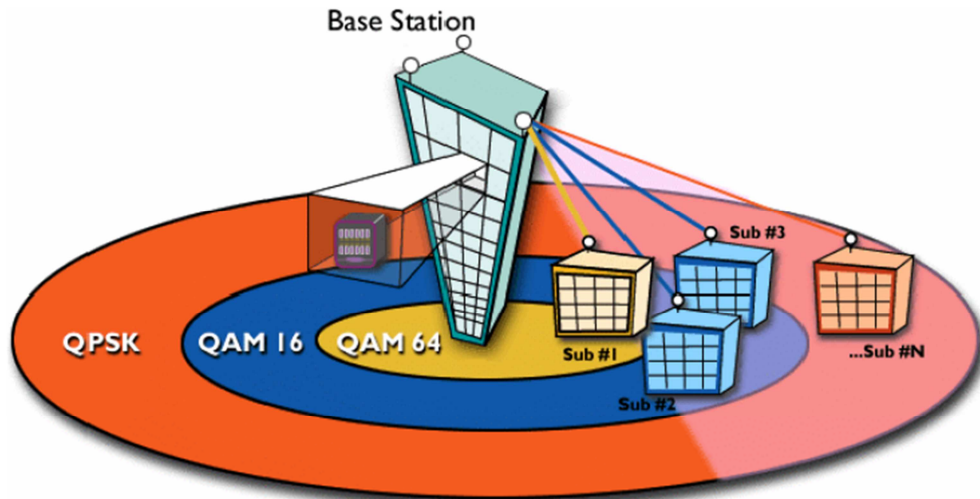


Figura 2.2 – Esquema adaptativo em função das condições do canal [10]

### WirelessMAN-SCa PHY

O WirelessMAN-SCa PHY também é baseada na modulação *single carrier* e foi projetado para trabalhar com transmissões do tipo ponto-a-ponto, sem linha de visada direta (NLOS) e com frequências menores que 11 GHz [8]. Assim como a especificação anterior, esta suporta duplexização FDD ou TDD e o mecanismo AMC. Utiliza a tecnologia TDMA para *uplink*, ou seja, o envio de informações da estação assinante para a estação base é feito através de acesso múltiplo por divisão do tempo. Para o *downlink* existe suporte para TDMA ou TDM, ou seja, o envio de dados da estação base para a base do assinante pode utilizar tanto acesso múltiplo por divisão do tempo como multiplexação por divisão do tempo.

### WirelessMAN-OFDM PHY

A WirelessMAN-OFDM é baseada no esquema de modulação *Orthogonal Frequency Division Multiplexing* (OFDM), que transmite em múltiplos canais espaçados ortogonalmente e ao mesmo tempo, evitando interferências [11]. A transmissão OFDM possui 256 subportadoras trabalhando em faixas de frequência abaixo de 11 GHz, também opera sem linha de visada. Esta variante é utilizada para pontos de acessos fixos, onde as estações dos assinantes ficam normalmente em residências ou prédios comerciais. Este PHY suporta subcanalização no *uplink*, suporta também operações com TDD e FDD em modo *half* e *fullduplex*. Esta especificação suporta múltiplos níveis de modulação como o BPSK, QPSK, 16-QAM e 64-QAM.

## WirelessMAN-OFDMA PHY

A WirelessMAN-OFDMA trabalha em faixas de frequência abaixo de 11 GHz, utiliza o modelo de modulação *Orthogonal Frequency Division Multiple Access* (OFDMA) com um número elevado de subportadoras: 2048 [8]. Este grande número de subportadoras acarreta no aumento de requisições de sincronização, tornando assim, mais lento o *Fast Fourier Transform* (FFT). Por estes e alguns outros motivos esta interface não tem despertado grande interesse na sua utilização, sendo mais utilizada a com 256 subportadoras.

## WirelessHUMAN

O WirelessHUMAN, que significa *Wireless High Speed Unlicensed Metro Area Network*, é similar as outras especificações citadas acima, sendo que esta é direcionada para dispositivos e bandas não licenciadas, que em inglês são chamados de *Unlicensed National Information Infrastructure* (UNII) *devices*. Para duplexação, apenas o TDD é suportado [12].

### 2.2.2 Camada MAC

A camada MAC do padrão IEEE 802.16 foi projetada para tratar diferentes demandas de fluxos de tráfegos (dados, voz e vídeo), ser eficiente o suficiente para atender um elevado número de usuários, e robusta para suportar tanto tráfego contínuo quanto em rajadas, garantindo a qualidade destes serviços (QoS). Para atender essas necessidades a camada de MAC foi subdividida em três subcamadas: Subcamada de Convergência de Serviços Específicos, Subcamada Comum e Subcamada de Segurança, (Figura 2.1) as quais serão detalhadas nas próximas subseções.

#### Subcamada de Convergência de Serviços Específicos

A subcamada de convergência de serviços específicos realiza a interface entre as camadas de mais alto nível com a subcamada comum de acesso ao meio, utilizando o ponto de serviço de acesso CS-SAP (CS - *Service Access Point*). Nesta camada é feito o mapeamento de serviços que não têm por origem e fim as conexões MAC 802.16 como: ATM, IPv4 (*Internet Protocol Version 4* – Protocolo de Internet Versão 4), IPv6 (*Internet Protocol Version 6* – Protocolo de Internet Versão 6), *Ethernet e Virtual Local Area Network* (VLAN). A função principal desta subcamada é classificar as unidades de serviços de dados SDUs (*Service Data Units*) para a apropriada conexão MAC, preservando ou habilitando QoS

e alocação de largura de banda [13].

### **Subcamada de Convergência Comum**

A camada de convergência comum da camada MAC tem como funcionalidades o estabelecimento e o gerenciamento das conexões realizadas entre as BS e as SSs, suporte a qualidade de serviço (QoS) e gerenciamento de largura de banda. Com todas essas funções, ela recebe e envia dados de outras camadas de convergência e faz todos os ajustes que forem necessários para adequação do tipo de conexão à camada MAC [8]. A subcamada de convergência comum é orientada a conexão, com o propósito de mapeamento de serviço nas SSs e associação de níveis de QoS de uma conexão. Uma vez estabelecida uma conexão, atribui-se um código único de identificação, *Connection Identifier (CID)*, sendo necessária uma manutenção contínua desta, dependendo do tipo de serviço conectado [14].

No decorrer de uma conexão em uma rede IEEE 802.6 deverão ser realizadas concessões, pela BS, e requisições, pela SS, de largura de banda pelas SS. A concessão pode ser realizada de duas formas: *Grant Per Subscriber Station (GPSS)* ou *Grant Per Connection (GPC)*. No mecanismo de alocação de largura de banda por estação assinante (GPSS) a estação base concede o recurso para a SS e esta pode redistribuir a largura de banda entre as suas conexões mantendo a QoS de acordo com o nível de serviço negociado. Esse mecanismo se adequa a cenários onde existem muitas conexões por terminal, o que possibilita ajustes mais sofisticadas de acordo com as necessidades de QoS das aplicações. Nesse contexto, a sobrecarga na BS é reduzida, pois a implantação de GPSS requer SSs mais “inteligentes”. A utilização de GPSS é obrigatória na especificação da camada física do padrão IEEE 802.16 que utiliza a faixa de frequência entre 10 – 66 GHz (padrão base). No mecanismo de alocação de largura de banda por conexão (GPC) a estação base concede o recurso para a estação assinante por conexão. Esse mecanismo é recomendado para cenários onde existem poucos usuários por SS. Sua implantação provoca uma sobrecarga de processamento na BS, mas permite que a implementação das SSs seja mais simples.

A requisição de largura de banda, processo pelo qual uma SS indica para uma BS que necessita de uma alocação de largura de banda, é feita de duas formas: incremental e agregada. A requisição incremental é aquela onde será adicionada a banda requerida à banda já existente da conexão, já a requisição agregada substitui a largura de banda da conexão existente pela nova requisição de banda. Para que haja este dois tipos de requisição de banda,



as SSs devem enviar periodicamente requisições de banda e essas atualizações irão depender da função da classe de serviço utilizada e da qualidade do enlace. A SSs podem requisitar largura de banda de quatro formas:

- *Piggyback Request*: os pedidos de largura de banda são realizados junto com os quadros de dados e é sempre incremental.
- *Implicit Request* ou *BW-Request*: pedido de largura de banda mais tradicionalmente utilizado. A solicitação de banda usando *slots* específicos do quadro de transmissão. A SS pode pedir banda implicitamente, sendo a largura de banda negociada no momento do estabelecimento da conexão.
- *Unicast Polling (Poll-Me Bit)*: Para contornar o processo de *polling* normal, uma estação assinante com conexão CBR (*Constant Bit Rate*) pode enviar um bit *poll-me* para a BS, que sorteia quais serão agraciadas com largura de banda.
- *Contention Basead Polling*: as SSs enviam mensagens *BW-Request* durante um intervalo de reserva. A contenção é resolvida usando o algoritmo *back-off*. Todas as mensagens de pedido de largura de banda são recebidas na BS pelo módulo *UL-BW Grant Generator*, que de acordo com a gerência do escalonador geral das BS concede tempos de transmissão.

Além dos esquemas de concessões e requisições de largura de banda pelas SS é necessário definir a prioridade de transmissão para cada conexão existente, com base no seu tipo de tráfego. Isto é realizado através de um escalonador que classifica cada conexão em um tipo pré-definido de classe de escalonamento, observado na Figura 2.3. Com base nos tipos de fluxo de serviços o padrão IEEE 802.16 define cinco categorias de serviço as quais devem ser tratadas de forma diferenciada pelo mecanismo de escalonamento da camada MAC: *Unsolicited Grant Service (UGS)*, *extended Real-Time Polling Service (ertPS)*, *Real-Time Polling Service (rtPS)*, *Non Real-Time Polling Service (nrtPS)* e *Best Effort (BE)*. Nas próximas subseções elas são brevemente detalhadas [11,15].

#### ***Unsolicited Grant Service – UGS***

A categoria UGS é projetada para oferecer suporte aos fluxos de serviço de tempo real que geram pacotes de dados de tamanho fixo em intervalos periódicos, ou seja, tráfego CBR.

Essa categoria de serviço pode ser representada pelo tráfego gerado por emulação T1/E1 e por aplicações de voz sobre IP sem supressão de silêncio. O serviço oferece periodicamente concessões não solicitadas para transmissão de dados. Isso elimina a sobrecarga e a latência das requisições das estações dos assinantes (SS – *Subscriber Stations*) para enviar pedidos de transmissão. No UGS, a SS é proibida de usar qualquer requisição de contenção e a estação base (BS – *Base Station*) não oferece qualquer oportunidade de requisição *unicast* para a SS. Requisições de *piggyback* também são proibidas no UGS. Os principais parâmetros para essa categoria de serviço são: *Maximum Sustained Rate*, *Maximum Latency Tolerance* e *Jitter tolerance* [11].

### ***Extended Real-Time Polling Service – ertPS***

A categoria ertPS foi adicionada na ratificação do 802.16e [8] oferece uma solução de compromisso entre a classe UGS e rtPS, oferecendo suporte aos fluxos de serviço de tempo real que geram pacotes de dados de tamanho variável em intervalos periódicos e com restrições de atraso. Um exemplo de aplicações deste tipo é ou voz sobre IP com supressão de silêncio. O serviço oferece periodicamente oportunidades de requisição *unicast*, as quais vão ao encontro das necessidades do fluxo de tempo real (largura de banda e retardo) e permitem que a SS especifique o tamanho da concessão desejada. A SS não tem permissão de utilizar qualquer esquema de requisição de contenção ou de *piggyback*. Os principais parâmetros para essa categoria de serviço são: *Jitter Tolerance*, *Traffic Priority*, *Maximum Latency Tolerance*, *Maximum Sustained Rate* e *Minimum Reserved Traffic Rate* [11].

### ***Real-Time Polling Service – rtPS***

A categoria rtPS é projetada para oferecer suporte aos fluxos de serviço de tempo real que geram pacotes de dados de tamanho variável em intervalos periódicos, como, por exemplo, o tráfego gerado por transmissão de vídeo no formato MPEG. Os mecanismos de requisição de largura de banda para esta categoria de serviço são os mesmos apresentados para serviços ertPS. Os principais parâmetros para essa categoria de serviço são: *Traffic Priority*, *Maximum Latency Tolerance*, *Maximum sustained rate* e *Minimum Reserved Traffic Rate* [11].

### ***Non Real-Time Polling Service – nrtPS***

A categoria nrtPS é projetada para oferecer suporte aos fluxos de serviço não tempo

real que geram pacotes de tamanho variável em intervalos periódicos, como, por exemplo, o tráfego gerado por aplicações FTP, e-mail, SMS, *multicast/broadcast*, MMS, entre outros. O serviço oferece periodicamente oportunidades de requisição *unicast (polls)*, utilizando intervalos de tempo mais espaçados do que na categoria rtPS. Isso assegura que o fluxo receba oportunidades de requisição mesmo durante períodos em que a rede esteja congestionada. Além disso, a SS pode utilizar oportunidades de requisição de contenção e de *piggyback*. Os principais parâmetros para essa categoria são: *Maximum Sustained Rate*, *Minimum Reserved Traffic Rate* e *Traffic Priority* [11].

### **Best Effort – BE**

Na categoria BE a SS pode utilizar oportunidades de requisição de contenção e de *piggyback*, mas não pode usar *polls* periódicos, além disso, não é permitido o envio de concessões periódicas para a transmissão de dados pela BS. O serviço de melhor esforço é tipicamente oferecido pela Internet para o tráfego gerado por navegação na Web. Os principais parâmetros para essa categoria de serviço são: *Maximum Sustained Rate*, e *Traffic Priority* [11]. É importante mencionar que para as categorias nrtPS e BE, o padrão especifica que a BS deve usar o parâmetro prioridade de tráfego para determinar a precedência na requisição do serviço e na geração da concessão para a transmissão de dados. Além disso, a BS deve oferecer preferencialmente oportunidades de requisição de contenção baseadas em prioridade.

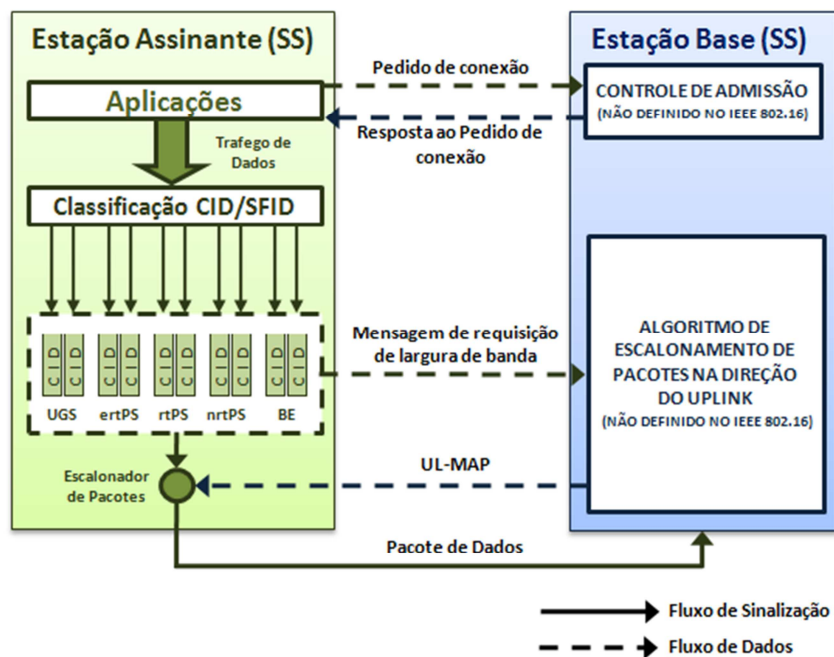


Figura 2.3 – Arquitetura de QoS do IEEE 802.16.

O escalonamento de fluxos de tráfego desempenha um papel crucial em sistemas que fornecem garantias de qualidade de serviço [16]. A classificação permite a distinção entre diferentes fluxos de tráfego e a possibilidade de tratamento diferenciado. Entretanto, simplesmente classificar os fluxos não garante que eles recebam um serviço com a QoS desejada. A classificação é apenas um mecanismo para distingui-los. A diferenciação no tratamento destes fluxos é uma decisão de um escalonador de fluxos de tráfego. Porém, o padrão IEEE 802.16 não especifica qual algoritmo de escalonamento deve ser utilizado (Figura 2.3). Essa implementação está a cargo dos fabricantes dos equipamentos, que com isso, irão pesquisar e desenvolver soluções com diferentes objetivos e desempenhos. Outra técnica que não é especificada pela norma é o controle de admissão de conexões, que possui como objetivo principal a preservação da QoS dos fluxos já existentes na rede, quando da chegada de uma nova solicitação. Um novo fluxo só é admitido se, por algum critério de decisão previamente estabelecido, não comprometer as garantias de qualidade já oferecidas aos demais fluxos. O controle de admissão tem ainda como objetivo secundário a maximização dos níveis de utilização da rede. Também é desejável que tenha baixo custo computacional, já que deve oferecer uma resposta em tempo real às solicitações de entrada dos novos fluxos [17]. Quando o controle de admissão aceita uma nova conexão, a qualidade de serviço será garantida se a fonte obedecer aos descritores de tráfego que são especificados durante o estabelecimento desta conexão. Entretanto, se o fluxo de tráfego violar o “contrato” inicial ou mudanças ocorrerem no enlace, como modulação e controle de erros, a rede poderá não atingir um desempenho aceitável. Assim, para impedir a violação dos contratos estabelecidos, deve existir algum mecanismo de policiamento do tráfego na rede [18].

### **Subcamada de Segurança**

O padrão IEEE 802.16 [11] buscou definir um ambiente seguro e robusto. A maior parte dos serviços de segurança é fornecida por primitivas criptográficas, na subcamada de acesso ao meio (MAC). Assim, a preocupação do padrão IEEE 802.16 foi definir uma suíte criptográfica – conjunto de funcionalidades desejáveis para atender aos serviços de segurança necessários. O conceito aqui envolvido é o de associações criptográficas, que contém informações sobre quais algoritmos devem ser aplicados, quais as chaves usar, entre outros. O padrão define ainda um processador dedicado à segurança, na estação base (BS). Há ainda alguns requisitos de funcionalidades criptográficas para o tráfego, bem como para a autenticação ponto-a-ponto. O tráfego em uma rede IEEE 802.16 deve ser cifrado

empregando o *Counter Mode* com *Cipher Block Chaining Message Authentication Code Protocol* (CCMP) - o qual emprega o AES (*Advanced Encryption Standard*), tanto para a segurança da transmissão, quanto para o serviço de autenticação da integridade dos dados. Excetuando-se a primeira conexão, todas as conexões são mapeadas para uma associação de segurança, seja em tempo de configuração, seja em operação. Para autenticação ponto-a-ponto, emprega-se a metodologia do protocolo de autenticação extensível (ou PKM-EAP - *Extensible Authentication Protocol*), que se baseia no padrão TLS (*Transport Layer Security*) de criptografia de chave pública. Atualmente, o protocolo PKM emprega o certificado X.509 com criptografia de chave pública RSA para autenticação dos SS, e troca de chaves de autorização. O próprio protocolo de mensagens PKM é autenticado empregando o protocolo *Hashed Message Authentication Code* (HMAC). Já a autenticação de mensagens nas funções da camada de acesso ao meio é provida pelo PKM [11].

## 2.3 Trabalhos Correlatos

Os esquemas de controle de admissão de conexão propostos em diversos outros trabalhos na literatura [20-22] usam apenas o requisito de taxa mínima das conexões no processo de decisão. Essa abordagem é válida mesmo para as conexões de tempo real, pois, de acordo com o padrão IEEE 802.16, a latência máxima precisa ser garantida apenas para o tráfego que não ultrapassa a taxa mínima requisitada pela conexão e essa garantia pode ser implementada pelo escalonador de pacotes. Algoritmos de controle de admissão que incluem garantias de latência para o tráfego de tempo real tendem a ser mais complexos, como, por exemplo, o algoritmo apresentado em [23]. Em [24] é apresentada uma variante do requisito de taxa mínima das conexões fazendo uma abordagem utilizando um modelo de degradação de largura de banda, o qual é aplicado apenas para conexões da classe nrtPS. [24] não avaliando o impacto que o modelo de degradação proposto tem sobre a QoS percebida pelo usuário final. Uma solução para mensurar este impacto pode ser encontrada em [25], que propõe uma forma para quantificá-lo através de funções utilidade, as quais, por sua vez, são definidas com base no tipo de tráfego associado.

Este trabalho diferencia-se dos demais encontrados na literatura por apresentar uma solução para o problema de controle de admissão através da modelagem de um processo markoviano de decisão utilizando o conceito de degradação proposto por [23], estendendo-o para as conexões das classes ertPS e rtPS, objetivando maior utilização da rede. Assim, o

processo de decisão do modelo proposto será guiado pelo impacto que esquema de degradação tem sobre a QoS percebida pelo usuário final, o qual será quantificado segundo as funções utilidade propostas em [25].

## 2.4 Conclusão

Este capítulo apresentou uma visão do padrão IEEE 802.16, descrevendo a família de padrões e seu modelo de referência. Alguns aspectos da camada Física do Padrão IEEE 802.16 formaram a composição deste capítulo, como os esquemas de modulação utilizados no Padrão IEEE 802.16. Detalhes do funcionamento da camada MAC no Padrão IEEE 802.16 também foram apresentados com a descrição das subcamadas de Convergência de Serviços Específicos, Comum e de Segurança. A camada MAC é de fundamental importância nas redes IEEE 802.16, pois é nesta camada que se implementa o suporte à provisão de QoS. Encerrando o capítulo, foram apresentados, de forma breve, os trabalhos correlatos mais relevantes para este trabalho.

## Capítulo 3 - Controle de Admissão de Conexão

---

### 3.1 Introdução

Este capítulo apresenta o problema do controle de admissão em uma rede IEEE 802.16, bem como alguns conceitos utilizados na sua abordagem, como o esquema de degradação de largura de banda como abordado em [24] e o conceito de funções utilidade introduzido em [25]. Os conceitos desenvolvidos neste capítulo serão utilizados posteriormente na formulação problema de controle de admissão de conexão como um processo markoviano de decisão.

### 3.2 Controle de Admissão de Conexão

O controle de admissão de conexão (CAC) é de suma importância em qualquer rede que permita provisão de QoS para suas conexões, pois possibilita o controle da utilização dos recursos da rede. A principal função do CAC é decidir adequadamente se canal de comunicação deve ou não aceitar uma nova conexão. Caso a política do controle aceite um número excessivo de conexões, o sistema não terá com garantir a QoS das conexões existentes, entretanto caso a política de controle admita um número muito pequeno de conexões, ou seja, rejeite muitas conexões, pode ocorrer um desperdício dos recursos da rede. Dentre as estratégias utilizadas para solucionar o problema de admissão destacam-se a *Complete Sharing* (CS) e a *Complete Partitioning* (CP), [26]. Em linhas gerais, a política CS permite que as conexões tenham um acesso igualitário dos recursos disponíveis na rede. Esta estratégia resulta em máxima utilização de largura de banda, especialmente em redes com muito tráfego. Entretanto, o tratamento igualitário de conexões com prioridade diferentes pode trazer conseqüências severas quando conexões de uma classe necessitam de muito menos recurso que outras. Nesta situação poderá ser desejável rejeitar conexões com baixo requisito de largura de banda para aumentar a probabilidade de aceitar uma futura conexão maior necessidade de largura de banda. A política CP atribui a cada classe uma faixa da largura de banda total da rede, que não pode ser utilizada por conexões de outras classes.

Sendo assim, ela suporta serviços diferenciados e controle da probabilidade de bloqueio das classes. Entretanto, a política CP não maximiza a utilização de todos os recursos disponíveis.

O controle de admissão de conexão (CAC) de uma rede IEEE 802.16 deve permitir tanto o tratamento diferenciado das classes de serviço como objetivar a maximização da utilização dos recursos da rede. Neste contexto, este trabalho propõe uma solução para este problema através da modelagem do problema de CAC como um processo markoviano de decisão (PMD) [27] cujo objetivo seja obter uma política ótima de admissão que maximize a utilização de recursos da rede, considerando a qualidade de serviço percebida pelo usuário.

A maximização de utilização de recursos da rede será feita utilizando-se o conceito de degradação de largura de banda proposto por [24], o qual é baseado nos requisitos diferenciados de largura de banda das classes de serviço do IEEE 802.16, por exemplo, conexões das classes ertPS, rtPS e nrtPS possuem uma taxa máxima e mínima de largura de banda. O padrão IEEE 802.16 especifica que deve ser garantido pelo menos uma largura de banda mínima para que este parâmetro QoS seja satisfeito, isto permite que o excedente de banda da conexão em questão possa ser “emprestado” para ser utilizado em novas conexões. É interessante notar que não pode haver “empréstimo” de largura de banda de conexões da classe UGS, pois estas necessitam de largura de banda constante, conforme estabelecido no padrão IEEE 802.16. Vale ressaltar ainda, que o conceito de degradação apresentado em [24] considera que somente conexões da classe nrtPS podem ser degradadas, porém neste trabalho esse conceito será estendido para as conexões da classes ertPS e rtPS, objetivando obter maior utilização da rede.

Assim, para a modelagem do problema considera-se uma rede IEEE 802.16 com topologia Ponto Multiponto (PMP), conforme mostrada na Fig. 3.1. O tráfego na estação base (BS) é dividido em três classes de serviços. A classe 1, corresponde às conexões da classe de serviço UGS, a classe 2, corresponde às conexões da classes de serviço ertPS e rtPS, agrupou-se essas duas classes devido serem equivalentes em termos requisitos de largura de banda. Finalmente, a classe 3 do modelo está associada à classe nrtPS. Não se considera a classe de serviço BE na modelagem, pois se assume que a mesma é sempre admitida pela BS, porém não há reserva de recurso no momento da admissão, isto é possível devido ao fato de o padrão IEEE 802.16 não definir largura de banda mínima para esta classe de serviço.



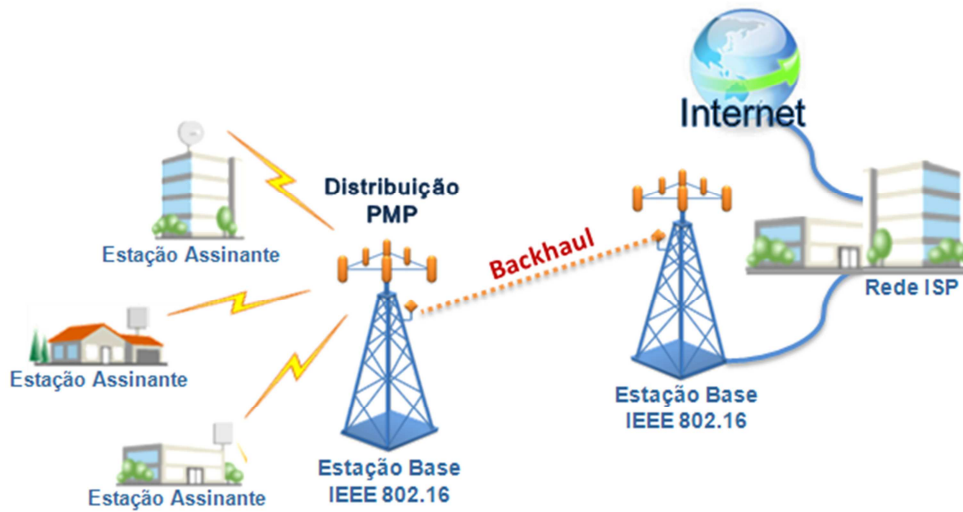


Figura 3.1 – Rede IEEE 802.16 PMP

### 3.3 Esquema de Degradação de Largura de Banda

Assume-se que a largura de banda requerida pelas conexões da classe 1 é  $b_1^{cte}$ , devendo ser satisfeita para que garantir os requisitos sua QoS. As taxas de transmissão das conexões das classes 2 e 3 podem variar na faixa  $[b_j^{min}, b_j^{max}]$ , onde  $b_j^{min}$  e  $b_j^{max}$  são a banda máxima e mínima para conexões da classe  $j$ , onde  $j \in \{2, 3\}$ . Quando existe banda suficiente disponível no sistema (baixo número de conexões ativas), as conexões podem transmitir a altas taxas, porém conforme o número de conexões no sistema aumenta, as conexões das classes 2 e 3 existentes podem dispensar certa quantidade de largura de banda com o objetivo permitir que um maior número de novas conexões seja aceite, maximizando desta forma a utilização da rede. Considera-se que uma conexão somente terá sua largura de banda diminuída para que uma conexão de maior ou igual prioridade seja admitida. Assim, a classe 2 só deverá ter sua largura de banda degradada com o intuito de aceitar novas conexões desta mesma classe ou da classe 1, enquanto que a classe 3 poderá ter sua largura de banda degradada para que seja possível aceitar um nova conexão de qualquer uma das classes definida no modelo.

A degradação é feita em passos onde  $\delta$  é a quantia de largura de banda “emprestada” para cada passo de degradação das conexões da classe 2 e 3. Considera-se que todas as conexões de uma classe devem manter o mesmo nível de degradação, sendo  $l_j^n$  o nível de degradação da classe  $j$ . Assim, a largura de banda reservada para cada conexão é  $b_j^{max} - l_j^n \delta$  que satisfaz  $b_j^{max} - l_j^n b_j^{min} \leq b_j^{min}$ . O passo máximo de degradação é definido por  $l_j^{max} =$

$\frac{b_j^{max}-b_j^{min}}{\delta}$ , onde  $j \in \{2, 3\}$ . O processo de degradação de largura de banda das classes 2 e 3 pode, potencialmente, terminar por servir as conexões destas classe somente a banda mínima definida para cada uma. Para evitar isso se deve fazer uma redistribuição da largura de banda sempre que houver disponibilidade para tal, ou seja, sempre que uma conexão sair do sistema e conseqüentemente liberar o recurso (largura de banda) que detinha, este deverá ser redistribuindo entre as conexões da classe 2 e 3, obedecendo a prioridade de cada classe.

### 3.4 Função Utilidade

O esquema de degradação de largura de banda tem impacto sobre QoS percebida pelo usuário final cuja conexão sofreu degradação. [25] propõe uma forma para quantificar este impacto através de funções utilidade, as quais são definidas com base no tipo de tráfego associado. Assim, [25] definiu três funções utilidade, cada uma associada a um tipo de tráfego: tempo real não adaptativo, tempo real adaptativo e tempo não real. Em [28] o conceito de utilidade é definido em detalhes e justificativas teóricas são apresentadas.

#### 3.4.1 Tráfego de tempo real não adaptativo

Conexões da classe de serviço com tráfego de tempo real não adaptativo possuem requisito de largura de banda crítico, de forma que, se a largura de banda reservada para conexões desta classe for menor que um dado limiar  $b_1^{cte}$ , esta terá utilidade nula do ponto de vista do usuário final. Este tipo de tráfego pode ser associado à classe de serviço UGS do IEEE 802.16 e conseqüentemente à classe 1 do modelo definido neste trabalho. Assim, a função utilidade de uma conexão da classe 1 em relação à largura de banda é dada por:

$$u_1(b_1) = \begin{cases} 1, & \text{se } b_1 \geq b_1^{cte} \\ 0, & \text{se } b_1 < b_1^{cte} \end{cases} \quad (3.1)$$

O formato exato desta função utilidade, figura 3.2, depende apenas de  $b_1^{cte}$ , o qual pode ser definido pelo operador da rede.

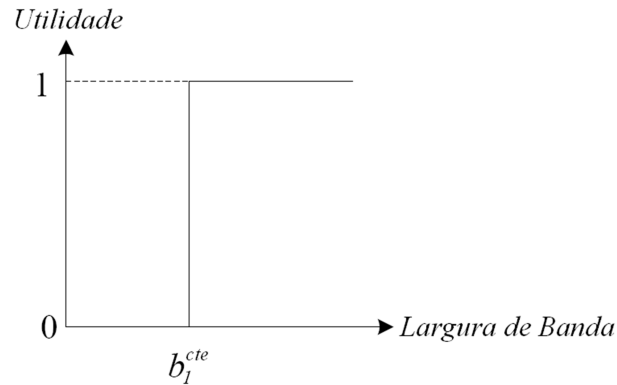


Figura 3.2 – Função utilidade para tráfego da classe 1.

### 3.4.2 Tráfego de tempo real adaptativo

Conexões com tráfego de tempo real adaptativo pode ter sua taxa transmissão reduzida em casos de elevada carga na rede, desde que lhes seja garantido uma banda mínima, abaixo da qual a utilidade de conexões desta classe de serviço torna-se nula. Devido à restrição de tempo real característica deste tipo de tráfego, o usuário é mais sensível às alterações na taxa de transmissão, pois estas aplicações possuem uma exigência intrínseca de largura de banda ( $b_2^{intr}$ ) dado que a taxa de geração de dados é independente do congestionamento da rede. Desta forma, a qualidade começa a cair drasticamente assim que a largura de banda é reduzida abaixo do  $b_2^{intr}$  e se torna inaceitável quando a largura de banda é reduzida abaixo de  $b_2^{min}$ . Este tipo de tráfego pode ser associado às classes de serviço rtPS e ertPS do IEEE 802.16 e conseqüentemente à classe 2 do modelo definido neste trabalho. Desta forma, a função utilidade de uma conexão da classe 2 em relação à largura de banda é dada por:

$$u_2(b_2) = \begin{cases} 1 - e^{-\frac{k_1 b_2^2}{k_2 + b_2}}, & \text{se } b_2 \geq b_2^{min} \\ 0, & \text{se } b_2 < b_2^{min} \end{cases} \quad (3.2)$$

Onde,  $k_1$  e  $k_2$  são parâmetros variáveis que determinam a forma da função utilidade e garantem que quando  $b_2 = b_2^{max}$ ,  $u_2(b_2) \approx 1$ , conforme a figura 3.3.

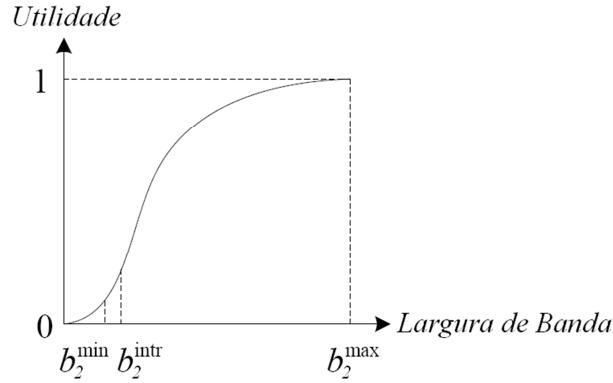


Figura 3.3 – Função utilidade para tráfego da classe 2.

Para se determinar a forma exata desta função é necessário calcular os parâmetros  $k_1$  e  $k_2$ . Observa-se que quando a largura de banda alocada  $b_2 = b_2^{max}$ , a utilidade correspondente  $u(b_2)$  é igual a  $u_2^{max}$ , conforme a equação a seguir.

$$1 - e^{\frac{k_1 b_2^2}{k_2 + b_2}} = u_2^{max} \quad (3.3)$$

A partir da equação 3.3 a relação entre  $k_1$  e  $k_2$  pode ser derivada como segue:

$$k_1 = \frac{\ln(1 - u_2^{max}) (k_2 + b_2^{max})}{-(b_2^{max})^2} \quad (3.4)$$

É necessário defini mais uma equação para calcular  $k_1$  e  $k_2$ . Assim, observando que a largura de banda intrínseca requerida  $b_2^{intr}$ , é definida como a largura de banda antes da qual a função utilidade é convexa e após a qual se torna côncava, fato este que ocorre no ponto onde a derivada de segunda ordem da função utilidade é igual a zero, temos:

$$\left(1 - e^{\frac{k_1 b_2^2}{k_2 + b_2}}\right)'' \quad (3.5)$$

Após calcular a derivada de segunda ordem, tem-se a equação a seguir:

$$2k_1 \frac{e^{\frac{k_1 (b_2^{intr})^2}{k_2 + b_2^{intr}}}}{(k_2 + b_2^{intr})} \left( k_2^3 + (b_2^{intr} - 2(b_2^{intr})^2 k_1) k_2^2 - 2(b_2^{intr})^3 k_1 k_2 - \frac{(b_2^{intr})^4 k_1}{2} \right) = 0 \quad (3.6)$$

Como  $k_1$  e  $k_2$  são números positivos, a equação 3.6 é igual a zero apenas quando equação cúbica entre parênteses é igual a zero.

$$k_2^3 + \left(b_2^{intr} - 2(b_2^{intr})^2 k_1\right) k_2^2 - 2(b_2^{intr})^3 k_1 k_2 - \frac{(b_2^{intr})^4 k_1}{2} \quad (3.7)$$

Após a substituição de  $k_1$  por  $\frac{\ln(1-u_2^{max})(k_2+b_2^{max})}{-(b_2^{max})^2}$ , a equação 3.7 fica como segue:

$$\begin{aligned} & \left(1 + 2 \ln(1 - u_2^{max}) \left(\frac{b_2^{intr}}{b_2^{max}}\right)^2\right) k_2^3 + \\ & \left(b_2^{intr} + 2 \ln(1 - u_2^{max}) \frac{(b_2^{intr})^2}{b_2^{max}} + 2 \ln(1 - u_2^{max}) \frac{(b_2^{intr})^3}{(b_2^{max})^2}\right) k_2^2 + \\ & \left(2 \ln(1 - u_2^{max}) \frac{(b_2^{intr})^3}{b_2^{max}} + \ln(1 - u_2^{max}) \frac{(b_2^{intr})^4}{2(b_2^{max})^2}\right) k_2 + \\ & \left(\ln(1 - u_2^{max}) \frac{(b_2^{intr})^4}{2b_2^{max}}\right) = 0 \end{aligned} \quad (3.8)$$

Dado que  $b_2^{intr}$ ,  $b_2^{max}$  e  $u_2^{max}$  são constantes pré-definidas no projeto da rede, a equação 3.8 é uma equação polinomial cúbica em  $k_2$  com todos os coeficientes constantes. Esta equação pode ser resolvida e a sua raiz cúbica positiva é o valor de  $k_2$ . Depois de encontrado  $k_2$ ,  $k_1$  pode ser calculado usando a equação 3.4. Neste trabalho a equação cúbica foi resolvida através do método de Tartaglia, também conhecido como método de Cardano cuja definição pode ser encontrada em [29].

### 3.4.3 Tráfego de tempo não real

Conexões com tráfego de tempo não real apresentam comportamento semelhante às conexões com tráfego de tempo real adaptativo no que se refere à possibilidade de degradação de largura de banda, ou seja, da mesma forma que nestas últimas, conexões com tráfego de tempo não real podem ter sua taxa transmissão reduzida em casos de elevada carga na rede, desde que lhes seja garantido uma banda mínima, abaixo da qual a utilidade de conexões

desta classe de serviço se torna nula. Porém, conexões de tempo não real apresentam menor sensibilidade à redução de largura de banda, do ponto de vista da utilidade percebida pelo usuário final. Este tráfego pode ser associado às classes de serviço nrtPS do IEEE 802.16 e conseqüentemente à classe 3 do modelo definido neste trabalho. Desta forma, a função utilidade de uma conexão da classe 3 em relação à largura de banda é dada por:

$$u_3(b_3) = \begin{cases} 1 - e^{-\frac{kb_3}{b_3^{max}}}, & \text{se } b_3 \geq b_3^{min} \\ 0, & \text{se } b_3 < b_3^{min} \end{cases} \quad (3.9)$$

Onde,  $k$  é um parâmetro variável que determina a forma da função utilidade e garante que quando  $b_3 = b_3^{max}$ ,  $u_3(b_3) \approx 1$ . A figura 3.7 mostra o formato desta função utilidade

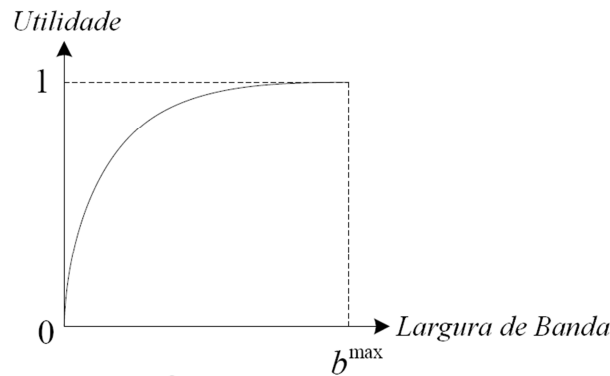


Figura 3.4 – Função utilidade para tráfego da classe 3.

Para determinar o exato formato da função utilidade para tráfego sem requisitos de tempo real, o parâmetro  $k$  precisa ser calculado. Dado que para uma função utilidade, quando a largura de banda alocada  $b_3 = b_3^{max}$ , a utilidade correspondente  $u$  é igual a  $u_3^{max}$ , conforme a seguinte equação.

$$1 - e^{-\frac{kb_3}{b_3^{max}}} = u_3^{max} \quad (3.10)$$

O parâmetro  $k$  é derivado da equação 3.10 como segue:

$$k = -\ln(1 - u_3^{max}) \quad (3.11)$$

Como  $u_3^{max}$  é uma constante que pode ser pré-definida pelo operador da rede,  $k$  pode ser facilmente calculado utilizando a equação 3.11.

### 3.4.4 Quantização das funções utilidade

Observando que no esquema de degradação proposto para este trabalho, a degradação é feita em passos  $\delta$ , as funções utilidade para cada uma das classes definidas em projeto devem ser quantizadas de acordo com este parâmetro. Assim, após a quantização uma função de utilidade  $u(b^l)$ , onde  $b_j^l$  é a largura de banda dado o nível de degradação da classe  $j$ ,  $j \in \{1,2,3\}$ , seria representada por uma lista  $\langle \text{largura de banda}, \text{utilidade} \rangle$  de pontos na ordem crescente de largura de banda dada por:

$$u(b^l) = \{ \langle b_1^0, u^0 \rangle, \langle b_1^1, u^1 \rangle \dots \langle b_1^n, u^n \rangle \} \quad (3.12)$$

Onde  $b^0$ , representa a largura de banda com menor nível de degradação e  $b^n$  a largura de banda para o nível máximo de degradação. A figura 3.8 demonstra a quantização para as classes 1, 2 e 3, conforme definidas neste trabalho. Observa-se que devido a não degradabilidade da largura de banda da classe 1, suas funções de utilidade são iguais antes e depois da quantização, contendo apenas um ponto, ou seja,  $u(b^l) = \{ \langle b_1^0, u^0 \rangle \}$ .

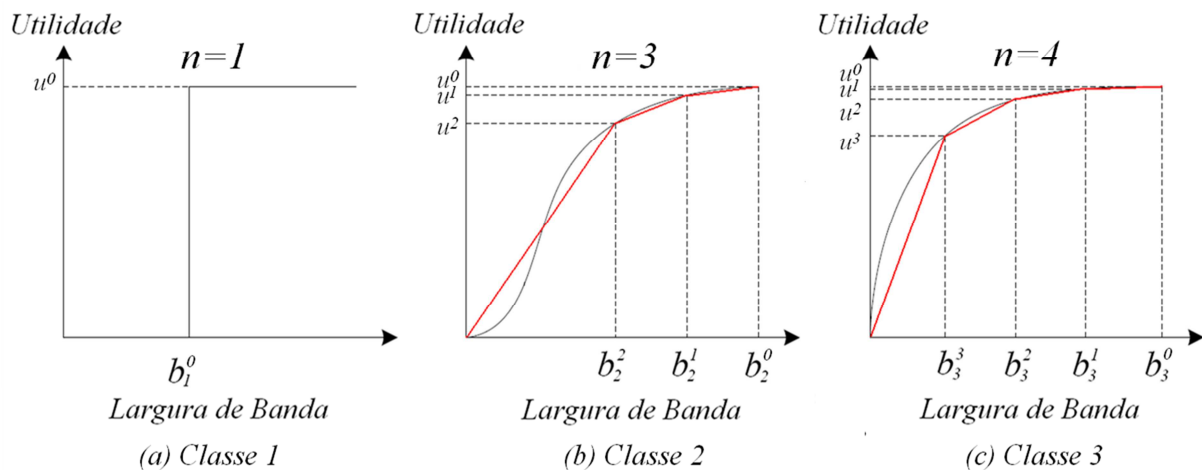


Figura 3.5 – Quantização das funções utilidade com intervalos de largura de banda iguais.

## 3.5 Conclusão

Neste capítulo, foram apresentados o modelo de degradação e as funções utilidade correspondentes para o problema de controle de admissão de conexão (CAC) em uma rede IEEE 802.16. Sendo assim, foram definidos os conceitos-chave que serão utilizados na formulação do problema de controle de admissão de conexão como um processo markoviano de decisão (PMD), a qual será apresentada no próximo capítulo.

## Capítulo 4 - Formulação do CAC como um PMD

---

### 4.1 Introdução

O objetivo deste capítulo é apresentar alguns conceitos básicos sobre processos semi-markovianos e processos de decisão markovianos de decisão a tempo contínuo, bem como a utilização destes conceitos na modelagem do problema apresentado no capítulo anterior, o qual será otimizado sob o critério do custo médio esperado a longo prazo.

### 4.2 Processos Semi-Markovianos de Decisão e Processos Markovianos de Decisão a Tempo Contínuo

Um Processo Semi-Markoviano de Decisão (PSMD) descreve um processo que é observado em pontos no tempo chamados instantes de decisão  $t$ . Em cada instante de decisão, o sistema é classificado em um estado  $i$ . Denota-se  $E$  o conjunto dos estados possíveis, também chamado espaço de estados. Se, em um instante de decisão, o sistema é observado no estado  $i \in E$ , deve-se escolher uma ação  $a \in A(i)$  para o estado  $i$ , onde  $A(i)$  é o conjunto de ações possíveis para o estado  $i$ . Assim, espaço de ações do sistema  $A$  é dado pela união de todos os conjunto  $A(i)$   $i \in E$ ,  $A = \bigcup_{i \in E} A(i)$ . O caráter markoviano do processo deve-se a hipótese do comportamento futuro do processo ser independente dos estados e ações passadas dado que o estado corrente e a ação escolhida são conhecidos em cada instante de decisão.

Um PSMD deve ser controlado por uma política que forneça um critério para a escolha de ações ao longo do processo decisório e uma regra de decisão específica a ação a ser escolhida num particular instante de decisão. Assim, uma política é uma seqüência de regras de decisão que pode ser descrita como  $R = (f^0, f^1, f^2, \dots)$ , onde  $f^t$ ,  $t = 0, 1, 2, \dots$ , é a regra de decisão a ser aplicada no instante de decisão  $t$ . Consideram-se políticas markovianas estacionárias determinísticas, ou seja, políticas que atribuem em cada instante de decisão uma ação que depende apenas do último estado observado, sem considerar o comportamento do processo no passado. Uma política desse tipo é construída a partir de uma regra de decisão



$f: E \rightarrow A$ , onde  $A = \cup_{i \in E} A(i)$ , tal que, se o estado observado é  $i \in E$ , então uma única ação  $f(i) \in A(i)$  é escolhida, independente do instante de decisão corrente  $t$ . Uma política markoviana estacionária determinística  $R$ , caracterizada pela regra de decisão  $f$ , atribui a ação  $f(i)$  sempre que o estado  $i$  for observado. Denota-se tal política  $R = (f, f, f, \dots)$  ou  $R = (f)^\infty$ .

Para um PSMD, dado que em um instante de decisão o sistema está no estado  $i \in E$  e a ação  $a \in A(i)$  foi escolhida, definem-se:

$\tau_i(a)$  = tempo esperado até o próximo instante de decisão;

$p_{ij}(a)$  = probabilidade de que no próximo instante de decisão o sistema estará no estado  $j$ ;

$C_i(a)$  = retorno ou custo esperado incorrido até o próximo instante de decisão.

O problema da otimização de um PSMD consiste em obter uma política de controle que otimiza o processo. No presente trabalho, considera-se como critério para a otimalidade a maximização do retorno esperado médio a longo prazo do processo. Para escrever a expressão deste retorno, define-se  $Z(t)$  como o retorno total incorrido até o instante  $t \geq 0$ . Para uma política de controle  $R$  e para um estado inicial  $i$ , define-se o retorno esperado médio a longo prazo do processo sob esta política como:

$$g_i(R) = \lim_{t \rightarrow \infty} \frac{1}{t} E_{i,R}[Z(t)] \quad (4.1)$$

onde  $E_{i,R}$  corresponde ao operador valor esperado quando o estado inicial é  $i$  e a política  $R$  é usada. [30] mostrou que o limite acima existe e que no caso *unichain*, ou seja, quando sob a política  $R$  o processo possui um único conjunto fechado de estados, o custo  $g_i(R)$  é independente do estado inicial  $i$  e pode ser denotado por  $g(R)$ .

Para os PSMD, é provado [30] que existe uma política markoviana estacionária determinística  $R^*$  ótima, ou seja, que maximiza o retorno médio do sistema a longo prazo dentro do espaço de políticas de controle possíveis.

Quando o sistema é controlado por uma política markoviana estacionária determinística

$R$  fixada, pode-se considerar a cadeia de Markov imersa no processo  $\{X_0, X_1, X_2, \dots\}$ . Esta cadeia representa a seqüência de estados visitados pelo processo, ou seja,  $X_0$  é o estado inicial do processo e, para  $n = 1, 2, \dots, X_n$  é o estado atingido pelo processo logo após a  $n$ -ésima transição. Esta cadeia de Markov tem o mesmo espaço de estados do processo original e probabilidades de transição dadas por  $p_{ij}(R_i)$ , onde  $R_i$  é a ação atribuída pela política  $R$  para o estado  $i$ . Se, sob a política  $R$ , a cadeia de Markov imersa no processo é *unichain*, então esta possui probabilidades limite denotadas por  $\{x_i^\infty(R), i \in E\}$ . [30] mostra que o retorno  $g(R)$  pode ser escrito como:

$$g(R) = \frac{\sum_{i \in E} C_i(R_i) x_i^\infty(R)}{\sum_{i \in E} \tau_i(R_i) x_i^\infty(R)} \quad (4.2)$$

A expressão do retorno de uma política  $R$ , em função das probabilidades limites da cadeia de Markov imersa no processo, não é utilizada no Algoritmo de Iteração de Valores considerado neste trabalho para a obtenção de uma política  $R^*$  de retorno máximo mínimo. Apesar disto, esta é importante porque pode facilmente ser estendida para a obtenção de medidas de desempenho do sistema sob esta política. De forma geral estas medidas representam o valor esperado médio de “funções de estado” definidas de forma análoga aos custos.

O tratamento computacional do modelo baseado num PSMD pode ser resumido em quator partes:

- Construção do modelo computacionalmente.
- Obtenção de uma política de controle markoviana estacionária determinística ótima pelo Algoritmo de Iteração de Valores apresentado em [29];
- Cálculo das probabilidades limite da cadeia de Markov imersa no processo quando a política de controle ótima é utilizada;
- Cálculo de medidas de desempenho do processo sob a política de controle ótima baseadas nas probabilidades limite calculadas.

Mais especificamente, o modelo a ser apresentado neste trabalho baseia-se num processo markoviano de decisão a tempo contínuo (PMDTC). Os PMDTCs são casos

particulares dos PSMD onde o tempo entre decisões sucessivas são exponencialmente distribuídos com parâmetro dependente do estado corrente do processo (último estado observado). Para estes processos, define-se  $\Lambda_{ij}(a)$  como a taxa de transição do estado  $i$  ao estado  $j$  ( $i, j \in E$ ), quando a última ação escolhida foi  $a \in A(i)$ . A partir das taxas de transição, obtém-se facilmente a taxa total de saída de cada estado dada por  $\Lambda_i(a) = \sum_{i \neq j} \Lambda_{ij}(a)$ . Note que  $\Lambda_i(a)$  é o parâmetro da distribuição exponencial negativa que descreve o tempo de permanência no estado  $i$  quando a ação  $a$  é escolhida. A partir das taxas de transição, as probabilidades de transição são dadas por  $p_{ij}(a) = \Lambda_{ij}(a)/\Lambda_i(a)$  e o tempo esperado entre transições é dado por  $\tau_i(a) = 1/\Lambda_i(a)$ . Assim, um PMDTC pode ser caracterizado pelo espaço de estados  $E$ , pelos conjuntos de ações  $A(i)$ , pelas taxas de transição  $\Lambda_{ij}(a)$  e pelos custos ou retorno  $C_i(a)$ .

O tratamento matemático e computacional dos PMDTCs é idêntico àquele dos PSMDs. Na prática, a diferença entre esses dois processos se encontra na forma que são construídos. Geralmente, para um PSMD, a obtenção das probabilidades de transição  $p_{ij}(a)$  e dos retornos  $C_i(a)$  envolve a teoria das renovações [31] e dificilmente os cálculos efetuados para um modelo podem ser diretamente reutilizado num outro modelo. Por outro lado, as probabilidades de transição e os custos necessários à construção de um PMDTC possuem expressões simples. A construção de um PMDTC permite o uso direto de uma modelagem por eventos, onde cada taxa de transição  $\Lambda_{ij}(a)$  está associada a um evento, que causa a transição de estado do processo.

Neste trabalho pretende-se modelar controle de admissão de uma rede PMP IEEE 802.16 como um processo markoviano de decisão a tempo contínuo. Estes processos são completamente definidos quando se descreve o espaço de estado, o espaço de ações, as probabilidades, custos (ou recompensas) e tempos esperados até o próximo instante de decisão.

O problema de decisão seqüencial é obter uma política que minimize (maximize) uma função desta seqüência de custos (recompensas). Neste trabalho esta função é o custo esperado médio a longo prazo, cuja demonstração pode ser encontrada em [32] e a política ótima é obtida utilizando técnicas de Programação Dinâmica, como o Algoritmo de Iteração de Valores que será brevemente apresentado na próxima seção.

### 4.3 O Algoritmo de Iteração de Valores

Esta seção apresenta uma descrição sucinta do algoritmo de iteração de valores, sendo que os resultados apresentados assim como prova da convergência podem ser encontrados em [30].

O Algoritmo de Iteração de Valores calcula recursivamente, para  $n = 1, 2, \dots$ , o valor da função  $V_i^n$ , definida por:

$$V_i^n = \max_{a \in A(i)} \left\{ C_i(a) + \sum_{j \neq i'} p_{ij}(a) V_j^{n-1} \right\}, \quad i \in E \quad (4.3)$$

Inicializado com uma função  $V_i^0$ ,  $i \in E$  arbitrariamente escolhida.

A quantidade  $V_i^n$  pode ser interpretada como o retorno esperado total máximo com um horizonte de  $n$  períodos quando o estado atual é  $i$ , e um custo terminal  $V_i^0$  é acarretado ao sistema quando este pára no estado  $j$ .

Esta interpretação sugere que, para  $n$  suficientemente grande, a diferença em um passo  $V_i^n - V_i^{n-1}$  estará muito próxima do retorno médio máximo por unidade de tempo para qualquer  $i \in E$ , quando  $n \rightarrow \infty$ , os limites:

$$m_n = \min_{i \in E} \{V_i^n - V_i^{n-1}\} \text{ e } M_n = \max_{i \in E} \{V_i^n - V_i^{n-1}\} \quad (4.4)$$

se aproximam da taxa de retorno máximo.

Escolhendo  $V_i^0$  tal que  $0 \leq V_i^0 \leq \max_a C_i(a)$  para todo  $i \in E$ , tem-se que  $V_i^1 \geq V_i^0$  para todo  $i$ , o que implica que cada termo da seqüência não decrescente  $\{m_n, n \geq 1\}$  é não negativo. Então, tem-se que:

$$\frac{M_n - m_n}{m_n} \leq \varepsilon \Rightarrow 0 \leq \frac{g(f^n) - g^*}{g^*} \leq \varepsilon \quad (4.5)$$

Isto é, o custo  $g(f^n)$  da política  $f^n$  obtido na  $n$ -ésima equação não pode diferir mais que  $\varepsilon$  do suposto custo médio mínimo  $g^*$  quando  $(M_n - m_n)/m_n \leq \varepsilon$ , onde  $\varepsilon$  é a precisão desejada. Em aplicações práticas, isto é usualmente satisfeito com uma política cujo custo médio está próximo do retorno médio máximo.

Então, tem-se que o Algoritmo de Iteração de Valores é dado por:

Para cada estado  $i \in E$  e para cada ação  $a \in A(i)$  escolher  $0 \leq V_i^0 \leq \max_a C_i(a)$ ;  
 Inicializar os parâmetros  $m_n, M_n$  e  $\varepsilon \in \mathcal{R}$ ;  
 Para  $n = 0, 1, 2, \dots$  fazer {  
   Para cada  $i \in E$  fazer {  
     
$$V_i^n = \max_{a \in A(i)} \left\{ C_i(a) + \sum_{j \neq i} p_{ij}(a) V_j^{n-1} \right\}, i \in E$$
  
     Determinar  $f^n$  como a política estacionária cujas ações maximizam o lado direito da equação acima.  
   }  
   Calcular os limites:  
     
$$m_n = \min_{i \in E} \{V_i^n - V_i^{n-1}\} \text{ e } M_n = \max_{i \in E} \{V_i^n - V_i^{n-1}\}$$
  
   O algoritmo para com a política  $f^n$  se  $(M_n - m_n)/m_n \leq \varepsilon$ .  
 }

Figura 4.1 – Algoritmo de Iteração de Valores

Para utilizar o Algoritmo de Iteração de Valores no caso contínuo, aplica-se o método de uniformização que pode ser encontrado em [30].

Através da uniformização, converte-se o modelo markoviano de decisão a tempo contínuo num modelo markoviano de decisão a tempo discreto tal que para cada política estacionária os custos (retornos) médios por unidade de tempo são os mesmos em ambos os modelos.

A seguir introduz-se a seguir os passos para a uniformização:

- Escolher um número  $\tau$  tal que  $0 < \tau \leq \min_{i,a} \tau_i(a)$ .
- Considerar o modelo markoviano de decisão a tempo contínuo cujo espaço de

estados, espaço de ações, probabilidade, custo e tempo entre transições são dados respectivamente por:  $E, A(i), p_{ij}(a), C_i(a)$  e  $\tau_i(a)$ .

- Obter o modelo markoviano de decisão a tempo discreto cujo espaço de estados, espaço de ações, probabilidades e custo entre transições são dados por:

$$\begin{aligned}\bar{E} &= E, \\ \bar{A}(i) &= A(i), \quad i \in \bar{E}, \\ \bar{C}_i(a) &= \frac{C_i(a)}{\tau_i(a)} \quad i \in \bar{E} \text{ e } a \in \bar{A}(i),\end{aligned}\tag{4.6}$$

$$\bar{p}_{ij}(a) = \begin{cases} \frac{\tau}{\tau_i(a)} p_{ij}(a), & j \neq i, i \in \bar{E} \text{ e } a \in \bar{A}(i), \\ \frac{\tau}{\tau_i(a)} p_{ij}(a) + \left[1 - \frac{\tau}{\tau_i(a)}\right], & j \neq i, i \in \bar{E} \text{ e } a \in \bar{A}(i). \end{cases}$$

Deve-se utilizar uma perturbação  $\tau$  estritamente menor que  $\min_{i,a} \tau_i(a)$ , para garantir que toda política ótima induza a uma cadeia de Markov aperiódica, condição necessária para garantir que  $\lim_{n \rightarrow \infty} m_n = \lim_{n \rightarrow \infty} M_n = g^*$ , onde  $g^*$  é o custo médio mínimo (retorno médio máximo) por unidade de tempo. Com a introdução de  $\tau$ , tem-se que, se as probabilidades de transição  $p_{ij}(a)$  da cadeia de Markov são perturbadas por alguma constante  $\tau$  com  $0 < \tau \leq \min_{i,a} \tau_i(a)$  a cadeia perturbada é aperiódica e tem as mesmas propriedades limites que a cadeia de Markov original. Em geral, recomenda-se escolher  $\tau = \min_{i,a} \tau_i(a)$  quando para esta escolha o requisito de a periodicidade é satisfeito; caso contrário  $\tau = \frac{1}{2} \min_{i,a} \tau_i(a)$  é uma escolha razoável.

#### 4.4 Formulação do CAC como um PMDTC

Tendo como base os conceitos apresentados nas seções anteriores, esta seção será dedicada à formulação do problema de controle de admissão de uma rede IEEE 802.16 como processo markoviano de decisão a tempo contínuo. Assim, definiremos primeiramente o espaço de estados, a seguir as dinâmicas dos estados e ações, e por fim o critério de otimização do PMD.

#### 4.4.1 Espaço de Estados

Para o modelo de controle de admissão assume-se que, o processo de chegada de novas conexões das classes 1, 2, 3 são processos de Poisson com taxas  $\lambda_1, \lambda_2$  e  $\lambda_3$ , respectivamente. Os tempos de serviço das conexões das classes 1, 2 e 3, são exponencialmente distribuídos com médias  $1/\mu_1, 1/\mu_2$  e  $1/\mu_3$ , respectivamente. Cada BS é modelada como uma cadeia de Markov a tempo contínuo cujo estado é representado por  $i = (n_1, n_2, n_3, l_2^n, l_3^n, ev)$ , onde  $n_1, n_2$  e  $n_3$  são o número de conexões ativa das classes 1, 2 e 3, respectivamente;  $l_j^n$  é o nível de degradação das conexões da classe  $j$  do sistema, para  $j \in \{2,3\}$ ; e  $ev$  representa a ocorrência de evento que pode ser:

- $a_1, a_2, a_3$ : chegada de uma nova conexão da classe 1, 2 e 3, respectivamente.
- $s_1, s_2, s_3$ : término de conexão da classe 1, 2 e 3, respectivamente.

Desta forma, o espaço de estados  $S$  do modelo proposto é:

$$\begin{aligned}
 E = \{ & (n_1, n_2, n_3, l_2^n, l_3^n, ev) | n_1 \geq 0, n_2 \geq 0, n_3 \geq 0, \\
 & ev \in \{a_1, a_2, a_3, s_1, s_2, s_3\}, \\
 & n_1 b_1^{cte} + n_2 (b_2^{max} - l_3^n \delta) + n_3 (b_3^{max} - l_3^n \delta) \leq B, \\
 & l_3^n \leq l_3^{max}, \\
 & l_2^n \leq l_2^{max}, \\
 & \text{se } n_1 = 0 \text{ então } ev \neq s_1, \\
 & \text{se } n_2 = 0 \text{ então } ev \neq s_2, \\
 & \text{se } n_3 = 0 \text{ então } ev \neq s_3, \\
 & \text{se } ev = a_3 \text{ então } l_2^n = l_2^{m'}.
 \end{aligned} \tag{4.7}$$

#### 4.4.2 Dinâmica dos Estados e Ações

Por hipótese, o sistema é observado continuamente no tempo. A cada chegada de uma nova chamada, deve-se decidir sobre sua aceitação ou rejeição (A ou R). Uma nova chamada só poderá ser aceita quando houver largura de banda suficiente para alocá-la. A cada término de conexão, esta sai do sistema e o recurso alocado antes alocado para ele é liberado, ou seja, deve-se adotar a ação L (libera), a qual redistribui o recurso liberado entre as classe 2 e 3 obedecendo a ordem de prioridade.

Assim para cada estado  $i = (n_1, n_2, n_3, l_2^n, l_3^n, ev) \in E$ , as ações possíveis  $A(i)$ :

$$A(i) = \begin{cases} \{A, R\} \text{ se } \begin{cases} (ev = a_1) \wedge ((n_1 + 1)b_1^{cte} + n_2b_2^{min} + n_3b_3^{min} \leq B) \vee \\ (ev = a_2) \wedge (n_1b_1^{cte} + (n_2 + 1)b_2^{min} + n_3b_3^{min} \leq B) \vee \\ (ev = a_3) \wedge (n_1b_1^{cte} + n_2(b_2^{max} - l_2^n\delta) + (n_3 + 1)b_3^{min} \leq B) \end{cases} \\ \{R\} \text{ se } \begin{cases} (ev = a_1) \wedge ((n_1 + 1)b_1^{cte} + n_2b_2^{min} + n_3b_3^{min} > B) \vee \\ (ev = a_2) \wedge (n_1b_1^{cte} + (n_2 + 1)b_2^{min} + n_3b_3^{min} > B) \vee \\ (ev = a_3) \wedge (n_1b_1^{cte} + n_2(b_2^{max} - l_2^n\delta) + (n_3 + 1)b_3^{min} > B) \end{cases} \\ \{L\} \text{ se } ev = s_1 \vee ev = s_2 \vee ev = s_3 \end{cases} \quad (4.8)$$

O comportamento dinâmico do sistema é definido pela observação contínua do estado corrente do sistema  $i = (n_1, n_2, n_3, l_2^n, l_3^n, ev) \in E$ , e pela ação adotada  $a \in A(i)$  cada vez que este estado se altera, ou seja, quando ocorre um novo evento. Admite-se que a decisão é tomada imediatamente após a observação da ocorrência do evento.

Se o estado  $i = (n_1, n_2, n_3, l_2^n, l_3^n, ev) \in E$  é observado e uma ação  $a \in A(i)$  é escolhida, o sistema passa imediatamente a um “estado reagido”  $s_r$  apresentado na Tabela 4.1, e aguardar a ocorrência do próximo evento, quando ocorrerá a próxima transição. Deve-se notar que o conceito de “estado reagido” não faz parte das definições teóricas de PMD, entretanto foi utilizado neste trabalho para facilitar a obtenção das probabilidades de transição de estado e dos retornos do processo. Nesse sentido temo que, se o último evento em  $i \in E$  é a chegada de uma chamada e é tomada a decisão de aceitá-la, o sistema imediatamente muda para um “estado reagido”  $j$  em que a nova chamada é incorporada ao sistema. Se o último evento é o término de uma chamada, a única ação possível é a sua retirada imediata do sistema, devendo-se ainda fazer a redistribuição de largura de banda utilizada antes por esta conexão por todas as demais conexões em andamento no sistema, objetivando manter sempre o nível de degradação  $l_k^{m'}$  mínimo para o novo “estado reagido”  $j$ , onde  $k \in \{2,3\}$ . As condições do estado  $i \in E$ , a ação tomada  $a \in A(i)$  e o “estado reagido”  $j$ , são apresentados na Tabela 4.1.

Na Tabela 4.2 apresentam-se o próximo estado e a taxa de cada transição, em função do “estado reagido” do sistema e do próximo evento a ocorrer. Nota-se que as transições



possíveis são combinações das linhas das Tabelas 4.1 e 4.2.

Condição	Ação	Estado Reagido
$(ev = a_1) \wedge ((n_1 + 1)b_1^{cte} + n_2(b_2^{max} - l_2' \delta) + n_3(b_3^{max} - l_3' \delta) \leq B)$	A	$(n_1 + 1, n_2, n_3, l_2', l_3')$
$(ev = a_2) \wedge (n_1 b_1^{cte} + (n_2 + 1)(b_2^{max} - l_2' \delta) + n_3(b_3^{max} - l_3' \delta) \leq B)$	A	$(n_1, n_2 + 1, n_3, l_2', l_3')$
$(ev = a_3) \wedge (n_1 b_1^{cte} + n_2(b_2^{max} - l_2' \delta) + (n_3 + 1)(b_3^{max} - l_3' \delta) \leq B)$	A	$(n_1, n_2, n_3 + 1, l_2', l_3')$
$(ev = s_1) \wedge (n_3 > 0)$	L	$(n_1 - 1, n_2, n_3, l_2', l_3')$
$(ev = s_2) \wedge (n_2 > 0)$	L	$(n_1, n_2 - 1, n_3, l_2', l_3')$
$(ev = s_3) \wedge (n_3 > 0)$	L	$(n_1, n_2, n_3 - 1, l_2', l_3')$
Caso contrário	R	$(n_1, n_2, n_3, l_2', l_3')$

Tabela 4.1 – Estado Reagido

Próximo evento	Próximo estado	Taxa de transição
$a_1$	$(n_1', n_2', n_3', l_2', l_3', a_1)$	$\lambda_1$
$a_2$	$(n_1', n_2', n_3', l_2', l_3', a_2)$	$\lambda_2$
$a_3$	$(n_1', n_2', n_3', l_2', l_3', a_3)$	$\lambda_3$
$s_1$	$(n_1', n_2', n_3', l_2', l_3', s_1)$	$t_1 \mu_1$
$s_2$	$(n_1', n_2', n_3', l_2', l_3', s_2)$	$t_2 \mu_2$
$s_3$	$(n_1', n_2', n_3', l_2', l_3', s_3)$	$t_3 \mu_3$

Tabela 4.2 – Próximo Estado e Transição Associada

#### 4.4.3 Critério de Otimização

A escolha de uma ação produz dois resultados: incorre-se em um retorno  $C_i(a)$

imediatamente, e o sistema muda para um novo estado num ponto subsequente no tempo de acordo com uma distribuição de probabilidade determinada pela ação escolhida. Assim, definiu-se a função  $g_i(a)$ , a qual corresponde à recompensa total esperada do sistema até o próximo instante de observação, dado que sistema é observado no estado  $i = (n_1, n_2, n_3, l_2^n, l_3^n, ev) \in E$ , e a ação tomada  $a \in A(i)$  é escolhida.

No modelo apresentado, este retorno corresponde ao número de chamadas sendo transmitidas no estado corrente, ponderados por um peso  $r_i$  atribuído às conexões de cada classe de serviço  $i$ , onde  $i \in \{1, 2, 3\}$ . Como imediatamente após a escolha da ação, o sistema passa ao “estado reagido”  $j = (n'_1, n'_2, n'_3, l_2^m, l_3^m, ev_r)$  mostrado na Tabela 4.1, é em função do “estado reagido” que se deve obter o retorno  $g_i(a)$ , o qual se calculado da seguinte forma:

$$g_i(a) = (r_1 n'_1 + r_2 n'_2 + r_3 n'_3) \tau_i(a) \quad (4.9)$$

Onde  $\tau_i(a)$  é o tempo esperado até o próximo instante de decisão, que é dado por  $\tau_i(a) = \sum_{i \neq j} \Lambda_{ij}(a) \Lambda_j(a)$ . Sendo  $\Lambda_{ij}(a)$  a taxa de transição do estado  $i \in E$  para o estado  $j \in E$ , quando a ação  $a \in A(s)$  é escolhida.

Os valores de  $r_1, r_2$  e  $r_3$  são definidos por funções de utilidade descritas na seção 3.3, obtidas através do mapeamento da largura de banda atribuída para cada conexão de classe de serviço e o desempenho percebido pelo usuário final.

Assim, os valores  $r_1, r_2$  e  $r_3$  da equação 4.9 são dados por:

$$\begin{aligned} r_1 &= r_1^{max} u_1(b_1) \\ r_2 &= r_2^{max} u_2(b_2) \\ r_3 &= r_3^{max} u_3(b_3) \end{aligned} \quad (4.10)$$

Onde,  $b_j$  e  $r_j^{max}$  são a largura de banda e retorno máximo atribuídos a uma conexão da classe de serviço  $j$ , respectivamente, onde  $j \in \{1, 2, 3\}$ . Os valores  $u_1(b_1)$ ,  $u_2(b_2)$  e  $u_3(b_3)$  são obtidos pelas funções utilidades definidas nas equações 3.1, 3.2 e 3.9, respectivamente.

## 4.5 Conclusão

Neste capítulo foi apresentada uma breve base teórica sobre processos semi-

markovianos e processos de decisão markovianos de decisão a tempo contínuo, bem como aplicação destes na definição e modelagem do problema de controle de admissão como um processo markoviano de decisão. Também se levou em consideração na modelagem do PMD os conceitos e restrições apresentado no capítulo 3. No próximo capítulo serão apresentados os resultados numéricos obtidos através do da resolução do PMD proposto neste capítulo.

# Capítulo 5 - Resultados Numéricos

---

## 5.1 Introdução

Neste capítulo apresentam-se os resultados obtidos por meio de resolução analítica do PMD para controle de admissão proposto no capítulo anterior. Na Seção 5.2 descreve-se o cenário e parâmetros utilizados. Na Seção 5.3 os resultados numéricos obtidos são apresentados e discutidos. Finalmente, a Seção 5.5 contempla as conclusões deste capítulo.

## 5.2 Método e Parâmetros

A modelagem do processo markoviano de decisão proposto para o CAC teve com base uma topologia PMP conforme a apresenta na figura 2.4. Objetivando avaliar a eficiência do PMD, este foi construído e solucionado utilizando a biblioteca MODESTO, desenvolvida em C++ pelo coorientador desta dissertação. Os resultados obtidos foram por sua vez visualizados utilizando MATLAB.

Especificaram-se duas métricas de avaliação para o esquema de CAC proposto. A primeira métrica refere-se ao retorno médio da rede, o qual é otimizado considerando a priorização de diferentes classes de serviço e a maximização da função definida na equação 4.9. A diferenciação das classes de serviço é feita através dos parâmetros de retorno. A segunda métrica mensura a probabilidade de bloqueio de cada classe de serviço, obtida na solução do PMD.

Por simplicidade e para poupar tempo e recursos computacionais, porém sem perda de generalidade, simula-se uma rede com topologia PMP conforme apresentado na figura 2.4, com capacidade de transmissão de 2Mbps. A diferença do tratamento de redes com capacidade de transmissão maior seria o aumento no tamanho do espaço de estados e, conseqüentemente, a exigência de maiores recursos computacionais ou mesmo técnicas analíticas mais avançadas.

Considerou-se que as taxas de processamento das chamadas são  $\mu_1 = \mu_2 = \mu_3 = 1$ ; e os retornos atribuídos às classes,  $r_1^{max} = 16$ ;  $r_2^{max} = 12$ ;  $r_3^{max} = 4$ . Nota-se que a atribuição de diferentes retornos a cada classe de serviço é uma forma de fazer o tratamento diferenciado de cada fluxo, assim atribuí-se um retorno maior à classe com que se deseja priorizar. Considera-se ainda que a largura de banda total  $BW$  da rede é 2Mbps e que o esquema de degradação é feito em passos  $\delta$  de 32Kbps. Os demais dados considerado são apresentados nas Tabelas 5.1 e 5.2, onde as funções utilidade das classes 1, 2 e 3 apresentadas nesta Tabela 5.2, são obtidos através dos conjuntos de equações apresentados na seção 3.3 calculados com base nos parâmetros da tabela 5.1. Pode-se observar o formato das funções utilidade para os parâmetros considerados na Figura 5.1. Deve-se ressaltar que os valores definidos na Tabela 5.1 foram baseados nos dados da referência [24].

Classe de serviço do IEEE 802.16	Parâmetro	Valor
–	BW	2Mbps
UGS	$b_1^{cte}$	64Kbps
ertPS/rtPS	$b_2^{min}$	64Kbps
	$b_2^{max}$	128Kbps
nrtPS	$b_3^{min}$	32Kbps
	$b_3^{max}$	128Kbps
–		32Kbps

Tabela 5.1 – Parâmetros do Sistema

Classe de Serviço IEEE 802.16	Classe de serviço correspondente na modelagem do CAC	Função Utilidade ( $b_j$ em Kbps)
UGS	1	$1, \quad se \ b_1 \geq b_1^{cte}$ $0, \quad se \ b_1 < b_1^{cte}$
ertPS/rtPS	2	$1 - e^{\frac{0,077b_2^2}{148,67+b_2}}, \quad se \ b_2 \geq b_2^{min}$ $0, \quad se \ b_2 < b_2^{min}$

nrtPS	3	$1 - e^{-\frac{4,6b_3}{128}}, \quad \text{se } b_3 \geq b_3^{\min}$	$0, \quad \text{se } b_3 < b_3^{\min}$
-------	---	---	--

Tabela 5.2 – Função Utilidade das Classes de Serviço

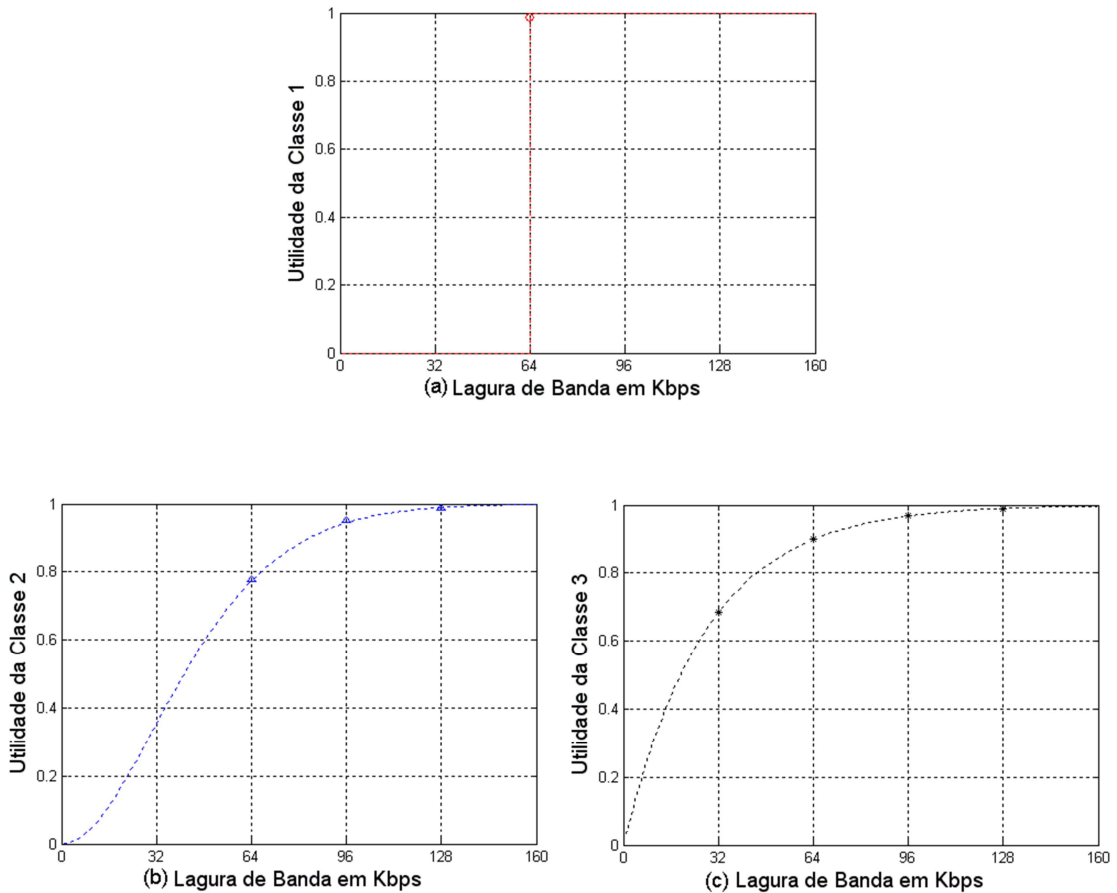


Figura 5.1 – Funções Utilidade das Classe de Serviço 1, 2 e 3.

Para obter uma avaliação do modelo com relação à carga do sistema (número de chamadas) variou-se a taxa de chegadas de conexões nas classes, fazendo,  $\lambda_1 = \lambda_2 = \lambda_3 = \{2, 4, 6, 8, 10, 12, 14, 16, 18 \text{ e } 20\}$ .

### 5.3 Resultados e Discussões

Nesta seção serão discutidos os resultados obtidos através da resolução analítica do PMD de decisão proposto para o problema de CAC. Mas antes convém comentar um pouco sobre a caracterização da política ótima obtida, que neste trabalho se trata de uma tabela indicando para cada estado do modelo a ação a ser tomada. Dado que o modelo, para os

parâmetros considerados, possui mais de 70.000 estados, a caracterização da política ótima graficamente é demasiadamente complexa. Porém, como se trata de uma tabela, do ponto de vista computacional é simples a sua consulta. Uma parte da política é mostrada na Tabela 5.3. Deve-se notar que uma ação “rej” é equivalente a ação R (rejeita) para o sistema, porém colocou-se essa diferenciação para facilitar a análise. Assim, “rej” significa que a nova conexão deve ser rejeitada, devido à impossibilidade de aceitar-la por falta de recursos ou restrição definida no modelo, e R significa que a nova conexão deve ser rejeitada por decisão da política, mesmo tendo recursos para aceitar-la. Analisando a linha 5 da Tabela 5.1, observa-se que a conexão da classe 2 do modelo (A2) foi rejeitada, pois a política optou por reservar recurso para a classe 1. Já na linha 6 a conexão da classe 3 do modelo (A3) foi rejeitada porque seria necessário degradar a classe 2 do modelo para liberar recurso e aceitá-la. Entretanto, a degradação de uma conexão pertencente a uma classe de maior prioridade para aceitar uma conexão de uma classe de menor prioridade é proibida, conforme definido no modelo do PMD, dessa forma a política não tem alternativa senão rejeitar a nova conexão.

<b>Linha</b>	<b>Estado</b> $(n_1, n_2, n_3, l_2^n, l_3^m, ev)$	<b>Ação</b>
1	(10, 8, 20, 1, 3, S1)	L
2	(10, 8, 20, 1, 3, S2)	L
3	(10, 8, 20, 1, 3, S3)	L
4	(10, 8, 20, 1, 3, A1)	A
5	(10, 8, 20, 1, 3, A2)	R
6	(10, 8, 20, 1, 3, A3)	rej
7	(10, 8, 21, 2, 3, S1)	L
8	(10, 8, 21, 2, 3, S2)	L
9	(10, 8, 21, 2, 3, S3)	L
10	(10, 8, 21, 2, 3, A1)	A
11	(10, 8, 21, 2, 3, A2)	R
12	(10, 8, 21, 2, 3, A3)	R

Tabela 5.3 – Caracterização da Política Ótima.

### 5.3.1 PMD para CAC com Degradação

Dado sequência à análise de resultados, nesta seção serão apresentados e discutidos os resultados obtidos para controle de admissão de conexão com degradação ponderado por funções utilidades modelado como um PMD. Assim, pode-se visualizar nas figuras subsequentes, modelo proposto é eficiente no que se propõe, ou seja, possibilita a maximização da utilização de recursos da rede e o tratamento diferenciado de classes de serviço. Pode-se constatar através dos resultados apresentados na Figura 5.2, que quando há baixa carga no sistema (taxa de chegada de conexões pequena) o *throughput* da rede é quase 100% para todas, ou seja, as conexões que chegam não são bloqueadas, pois a rede dispõe de recursos suficientes para suprir a demanda, sendo possível atribuir a todas as conexões a largura de banda máxima permitida pela classe de serviço a qual estão associadas, de forma que o nível médio de degradação da classe 2 e 3 são próximos de zero para  $\lambda_1 = \lambda_2 = \lambda_3 = 2$  e 4, conforme Figura 5.3. Entretanto, conforme elevamos a carga do sistema observa-se há favorecimento das classes com maior retorno (classe 1, 2), aquelas consideradas mais relevantes para o sistema no projeto da rede, em detrimento da classe 3, com menor retorno. Ou seja, a política do CAC tenta sempre manter o *throughput* da classe UGS (classe 1 do CAC proposto) próximo de 100%, sendo esta a classe mais favorecida, conforme definido no dados de projeto, seguida pela classe ertPS/rtPS (classe 2 do CAC proposto). A classe nrtPS (classe 3 do CAC proposto) acaba por sua vez tendo seu *throughput* reduzido e, conseqüentemente, sua probabilidade de bloqueio elevada conforme a carga do sistema aumenta, Fig. 5.4.

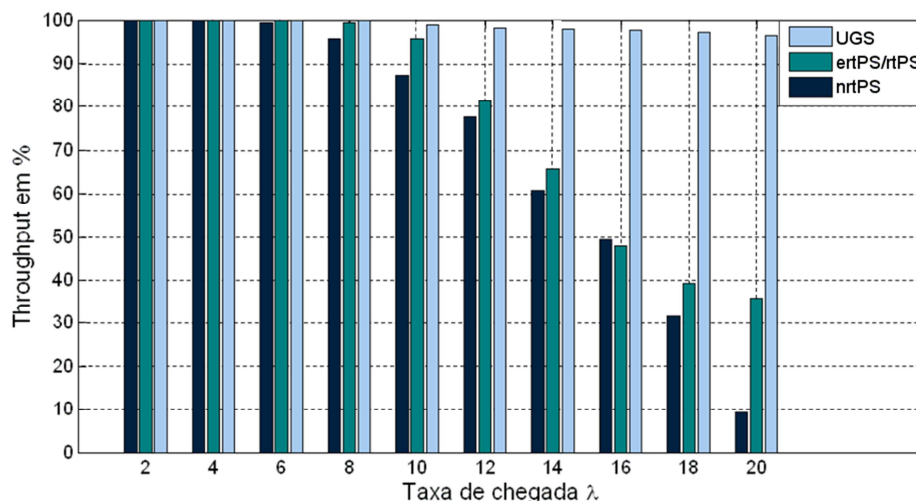


Figura 5.2 – Largura de Banda Utilizada em Percentagem



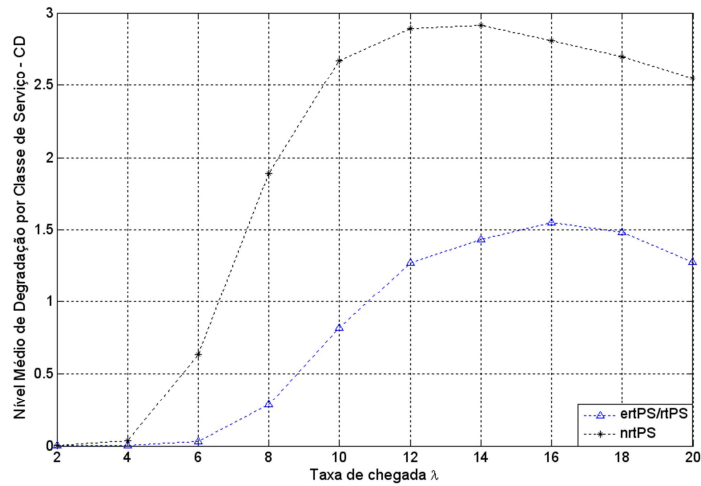


Figura 5.3 – Nível Médio de Degradação por Classe de serviço

É interessante observar a relação entre o nível médio de degradação (Fig. 5.3) e as probabilidades de bloqueio das classes ertPS/rtPS e nrtPS (Fig. 5.4). Conforme aumentamos a carga sistema, a política de controle do CAC, através da elevação do nível de degradação de largura de banda das classes rtPS e ertPS/nrtPS, tenta manter as probabilidades de bloqueio do sistema baixas. Entretanto, para cargas mais elevadas ainda  $\lambda_1 = \lambda_2 = \lambda_3 > 14$ , o sistema precisa bloquear mais conexões, para manter QoS das conexões em andamento. O aumento da probabilidade de bloqueio faz com que mais recursos fiquem disponíveis para as conexões em andamento, permitindo uma redistribuição de largura de banda e conseqüente diminuição no nível de degradação.

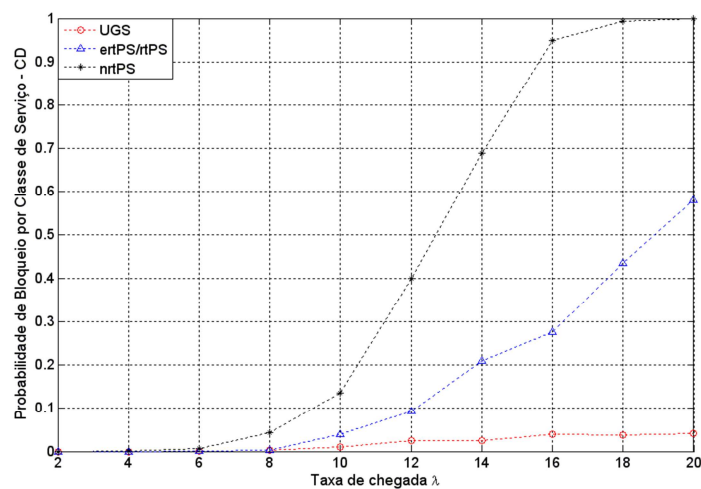


Figura 5.4 – Probabilidade de Bloqueio por Classe de Serviço

Na Figura 5.5 pode-se observar o comportamento do sistema com relação à utilização

de recursos em percentagem (largura de banda). Quando a taxa de chegada  $\lambda_1 = \lambda_2 = \lambda_3 = 4$ , o sistema passa a ter utilização de quase 100%, mantendo probabilidade de bloqueio próximo de zero para todas as classes, Figura 5.4, com todas as classe de serviço sendo atendidas com a largura de banda máxima permita para a classe de serviço a qual está associada, Figura 5.5, ou seja todas as conexões podem ser atendidas sem que seja necessário degradar a largura de banda das classes ertPS/rtPS e nrtPS, Figura 5.3. Ainda através da Fig. 5.5 é possível observar com maior clareza a maximização da utilização de recursos e o favorecimento das classes com maior retorno, pois conforme, a carga do sistema aumenta a utilização sempre se mantém próximo de 100%, entretanto a proporção de conexões existentes de cada uma das classes de serviço vai se alterando até o ponto que para uma taxa de chegada de  $\lambda_1 = \lambda_2 = \lambda_3 \geq 16$  a percentagem de conexões da classe UGS, mais relevante para o sistema, ultrapassa 50% da largura de banda total do sistema, isto é reflexo do aumento da probabilidade de bloqueio das classes ertPS/rtPS e nrtPS.

Na Figura 5.6 observa-se, em temos largura de banda média por conexão de serviço, o efeito causado pela degradação da largura de banda das classes ertPS/rtPS e nrtPS conforme a carga do sistema aumenta. Observa-se ainda que a largura de banda média da classe UGS se mantém constante, pois assim foi definida, não sendo permitindo degradação alguma de sua largura de banda.

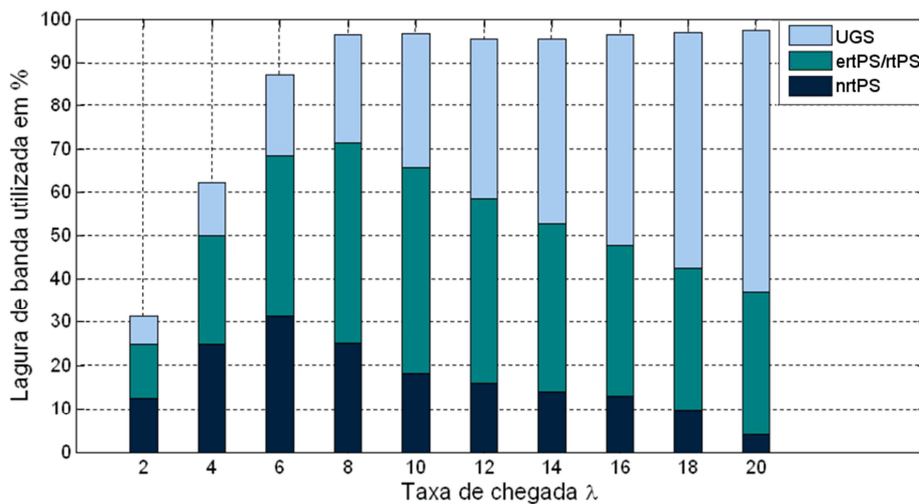


Figura 5.5 – Largura de Banda Utilizada em Percentagem

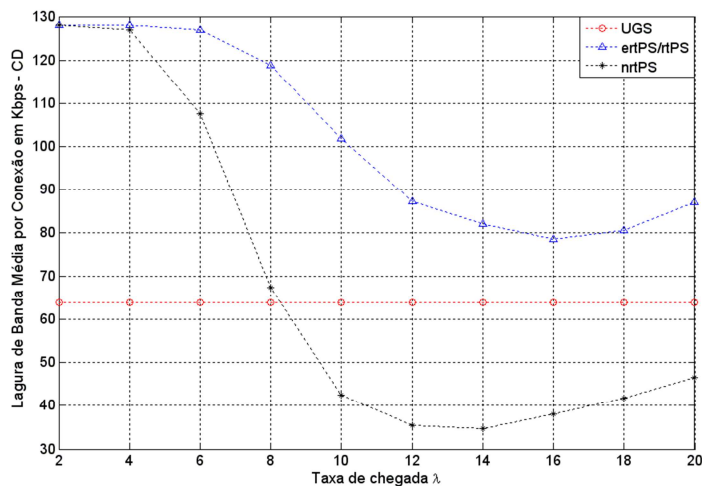


Figura 5.6 – Largura de Banda Média por Conexão

### 5.3.2 Análise da Degradação de Largura de Banda

Nesta seção será feita uma análise dos efeitos da degradação de largura de banda no sistema. Para isso, será mostrado um comparativo o qual entre modelo de CAC tal como definido nesse trabalho que, para efeito didático, será referenciado de agora em diante de Esquema – CD com uma variante sua (Esquema – SD), a qual terá como única diferença a não capacidade de degradação de largura de banda. O Esquema – SD pode ser entendido como um Esquema – CD onde a banda máxima  $b_j^{max}$  é igual à banda mínima  $b_j^{min}$  para conexões da classe  $j$ , onde  $j \in \{2, 3\}$ , é será definido e solucionado como um PMD tal como o Esquema – CD. Ambos os esquemas utilizam em sua definição e solução os mesmos parâmetros discutidos anteriormente, com exceção que no Esquema – SD, que devido sua não degradabilidade de largura de banda, só poderá atribuir à cada conexão a largura de banda máxima estipulada pela sua classe de serviço, conforme a tabela 5.1 e, como consequência disso, as funções utilidade para este esquema serão sempre máximas.

Os resultados obtidos permitem avaliar a importância da degradação de largura de banda no desempenho do modelo proposto. Primeiramente, como se pode observar na Figura 5.7, em termo de utilização temos que o Esquema – CD, que permite degradação, é superior independente a elevação da carga da rede. Este fato é esperado, pois a degradação permite um melhor aproveitamento da largura de banda da rede, ao permitir que as conexões em andamento sejam ajustadas objetivando admitir o maior número de conexões possível.

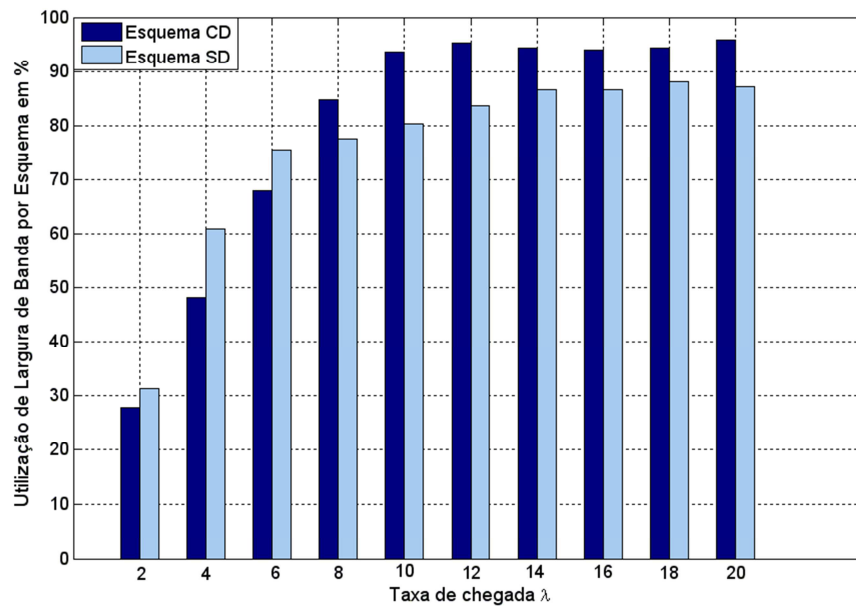


Figura 5.7 – Utilização de Recursos por Esquema

Na Figura 5.8 é apresentada uma comparação entre as probabilidades de bloqueio para cada classe, em cada esquema, onde mais uma vez se constata a superioridade do Esquema – CD sobre Esquema – CD. Este resultado está relacionado diretamente na capacidade de degradação, observando que, ao se permitir o ajuste de largura de banda das conexões em andamento, é possível aceitar um maior número de conexões como a mesma quantidade de recurso, e como consequência disso a probabilidade de rejeitar (bloquear) uma nova conexão será menor.

Prosseguindo o comparativo, temos na Figura 5.9 os resultados referentes à largura de banda média utilizada por classe de serviço em cada um dos esquemas. Pelo menos dois fatos interessantes podem se observados nestes gráficos. Primeiro, que a largura de banda média utilizada pela classe de serviço UGS Figura 5.8-a é muito próxima nos dois esquemas, reflexo da baixa probabilidade de bloqueio dessas classes conforme Figura 5.8. Vale lembrar que essa baixa probabilidade de bloqueio é causada (em ambos os esquemas) pelo favorecimento desta classe através do alto retorno atribuído a ela. O segundo fato interessante observado na Figura 5.9-c refere-se à largura de banda média da classe nrtPS (classe de menor prioridade no modelo), observa-se que para baixas cargas no sistema, ou seja,  $\lambda_1 = \lambda_2 = \lambda_3 < 8$ , temos que a largura de banda média da classe nrtPS é muito maior no Esquema – SD do que no Esquema – CD, caindo drasticamente conforme a carga do sistema aumenta. Este fato pode ser explicado tomando como base a probabilidade de bloqueio desta classe, que por ser a menos favorecida pelo modelo (menor retorno), acaba apresentando maior probabilidade de

bloqueio.

Assim, a queda drástica da largura de banda média da classe nrtPS no Esquema – SD se deve basicamente a baixo retorno desta classe, não degradabilidade do esquema e alta necessidade de recursos desta classe, lembrando que no Esquema – SD deve ser atribuído a largura de banda máxima definida na tabela 4.1. Quando a carga do sistema é baixa, é possível atender todas as conexões com sua largura máxima, inclusive as da classe nrtPS, entretanto conforme a carga do sistema aumenta os recursos se tornam escassos e é necessário utilizá-los para atender as classes de maior prioridade, e uma classe de baixa prioridade e alta necessidade de recurso como a nrtPS precisa ser bloqueada. No Esquema – CD isto ocorre de forma mais suave devido à degradabilidade, assim em casos de elevada carga no sistema pode-se atribuir uma quantidade de recurso menor, diminuindo a necessidade de bloqueio desta. É importante ressaltar que o desfavorecimento da classe nrtPS quando o sistema possui alta carga ocorre em ambos os esquemas, porém a capacidade de degradação permite isto ocorra mas suavemente no Esquema – CD.

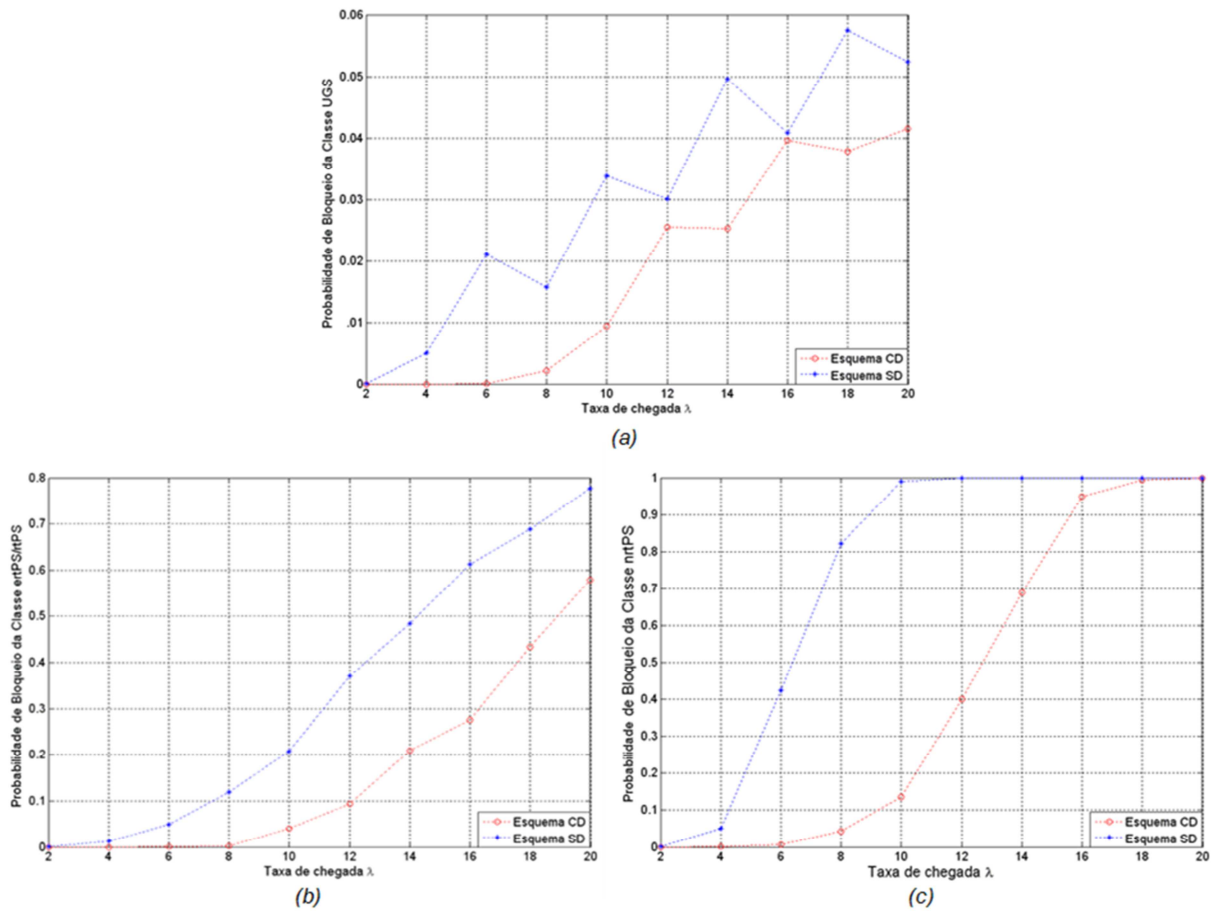


Figura 5.8 – Probabilidade de Bloqueio por Esquema

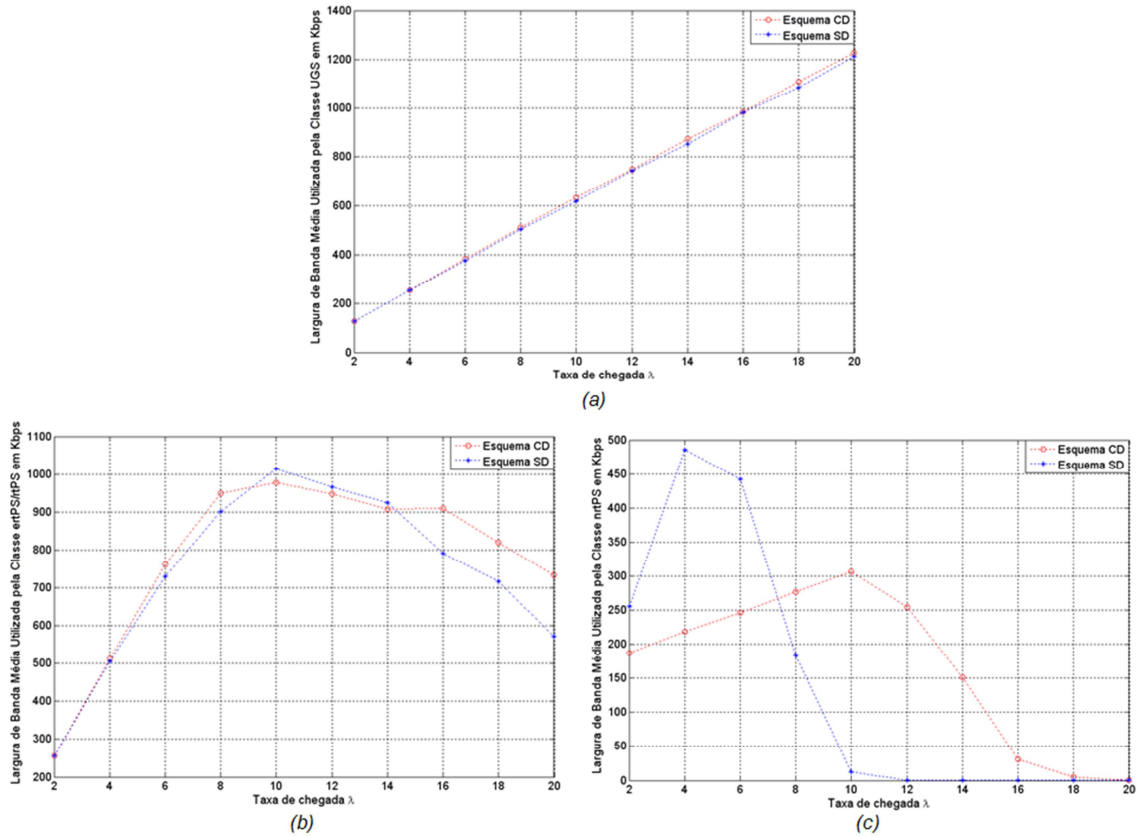


Figura 5.9 – Largura de Banda Média por Esquema

### 5.4 Conclusões

Neste capítulo avaliou-se, através de modelo analítico, o desempenho dos esquemas de controle de admissão propostos no Capítulo 4, vinculado ao cenário descrito na seção 5.2. Assim, os parâmetros de QoS: probabilidade de bloqueio e *throughput* foram avaliados no cenário supracitado. Dessa forma, o mecanismo de controle de admissão proposto conduziu a bons níveis nos parâmetros de QoS mensurados. Foi possível avaliar também a relevância do conceito de degradação de largura de banda através do comparativo de resultados do modelo analítico proposto com e sem a utilização do mesmo. No próximo capítulo será feito um apanhado geral do de todo o trabalho apresentado destacando as suas contribuições bem como sugestões de continuidade.

## Capítulo 6 - Conclusões e Trabalhos Futuros

---

### 6.1 Conclusões

O padrão IEEE 802.16 tem se mostrado uma solução promissora para o problema de acesso de última milha. Suas especificações dão suporte às grandes demandas das aplicações banda de larga. Por ser um padrão sem fio com cobertura de rádio frequência relativamente estendida, possibilita a cobertura de grandes distâncias com custo de investimento em infraestrutura reduzido e tempo de implantação bem menor aos das soluções concorrentes cabeadas.

Um dos grandes atrativos do padrão IEEE 802.16 é o suporte ao tratamento diferenciado dos fluxos de tráfego, o qual possibilita a provisão garantida de qualidade de serviço (QoS). Nesse contexto, o controle de admissão de conexão (CAC), objeto de estudo deste trabalho, é estratégico, pois possibilita o controle da utilização dos recursos da rede. Todavia, o CAC não foi especificado pelo padrão IEEE 802.16, umas das justificativas para que isso ocorresse foi que o CAC deveria ser objeto de pesquisa e serviria para diferenciação de produtos entre fabricantes.

O problema do CAC foi abordado neste trabalho, onde se buscou propor uma solução baseada na modelagem e otimização de processo markoviano de decisão em uma rede IEEE 802.16 PMP. A otimização do modelo foi feita de forma a encontrar uma política de controle ótima de admissão de conexão que maximizasse a utilização da rede utilizando um esquema de adaptação de largura de banda, e ainda considerando a utilidade percebida pelo usuário final.

A abordagem markoviana permitiu um tratamento analítico elegante do problema de CAC, produzindo resultados que mensuram a capacidade da rede dada uma configuração de projeto, através de métricas de desempenho de longo prazo como a probabilidade de bloqueio. A modelagem da admissão de tráfegos com retornos diferenciados permitiu o tratamento diferenciado das classes de serviços, possibilitando que, em caso de grande carga no sistema,

as classes consideradas mais relevantes do ponto de vista do usuário sejam favorecidas.

Assim, baseando-se nos resultados apresentados, conclui-se que a política ótima encontrada para o processo markoviano de decisão proposto tratou as conexões do sistema de forma diferenciada com base em na classe de serviço a qual pertenciam. A política encontrada também fez um balanceamento entre a maximização da utilização de recursos do sistema, através do esquema de degradação de largura de banda e a qualidade de serviço percebida pelo usuário final, quantificada através de funções utilidades. A combinação dos conceitos de maximização de utilização e QoS percebida pelo usuário final, divergentes entre si, em uma solução elegante e otimizada foi a principal contribuição deste trabalho. Entretanto, partido dos conceitos apresentados, muito ainda pode ser desenvolvido. Na próxima seção serão apresentadas algumas sugestões para trabalhos futuros.

## 6.2 Trabalhos Futuros

Para trabalhos futuros, sugere-se estudar uma forma de caracterização mais eficiente da política ótima obtida, que permita uma análise qualitativa e independente dos parâmetros do modelo, fato que não ocorre com a política atual, já esta se dá na forma de tabela contendo a ação a ser tomada para cada estado do modelo, sendo assim a tabela que caracteriza a política está intimamente ligada à quantidade de estados, que por sua vez depende da capacidade da rede.

Pretende-se ainda estender o modelo para fazer com que a adaptação de largura de banda seja feita levando-se em consideração a necessidade de cada conexão individualmente. Para isso sugere-se a modelagem de um escalonador combinado ao modelo de CAC de proposto, o que permitirá avaliação de métricas de desempenho de curto prazo como *jitter* e atraso.

Outra sugestão para trabalhos futuros é estender o modelo de CAC proposto para redes IEEE 802.16 *mesh*. Nesse sentido, a abordagem seria feita primeiramente considerando uma topologia infraestrutura, na qual as estações base são interconectadas formando *backhauls* de transmissão, com gerenciamento feito de forma centralizada. Essas considerações facilitariam a migração do modelo proposto nesse trabalho para o cenário *mesh*, uma vez que o canal alocado para uma conexão conforme considerado neste trabalho seria estendido, no contexto *mesh*, para todo o caminho (*path*) alocado para uma conexão.



## Referências Bibliográficas

---

- [1] IEEE 802.16 Working Group. IEEE Standard for Local and Metropolitan Area Networks – Part. 16: Air Interface for Fixed Broadband Wireless Access Systems. 2004.
- [2] SENGUPTA S. et all, Exploiting MAC flexibility in WiMAX for media streaming. In: Proceedings of sixth IEEE international symposium on a world of wireless mobile and multimedia networks. Giardini Naxos, Italy, pp. 338–343, 2005.
- [3] DRASKOVIC, D. High Efficiency Power Amplifiers for WiMAX Transmitter RF Front Ends. Masters Thesis, University of Westminster, London, 2006.
- [4] ABICHAR, Z. et all, WiMAX: The Emergence of Wireless Broadband, IT Professional, vol. 8, issue 4, pp. 44-48, 2006.
- [5] SIVCHENKO, D. et all, Internet Traffic Performance in IEEE 802.16 Networks, Proceedings in European Wireless, Athens, Greece, 2006.
- [6] BURBANK, J. L. & KASH, W. T. IEEE 802.16 Broadband Wireless Technology and Its Applications to the Military Problem Space. Military Communications Conference, IEEE, 2005.
- [7] KURAN, M. S. & TUGCU, T. A Survey on Emerging Broadband Wireless Access Technologies. IEEE Computer Networks, vol. 51, no. 11, pp. 3013-3046, 2007.
- [8] IEEE 802.16-2005 Working Group. IEEE Standard for Local and Metropolitan Area Networks – Part. 16: Air Interface for Fixed Broadband

Wireless Access Systems for Mobile Users. 2005.

- [9] FIGUEIREDO, Fabrício Lima. Fundamentos da Tecnologia WiMAX. Centro de Pesquisa e Desenvolvimento em Telecomunicações, CpqD. 2006.
- [10] MARKS, R. B. et al, The 802.16 WirelessMAN MAC: It's done, but what is it?. 2002.
- [11] JEFFREY G. Andrews et al, Fundamentals of WiMAX: Understanding Broadband Wireless Networking. Prentice Hall PTR, First Edition, 2007.
- [12] NAVES, Sanzio Guilherme. CHAN, Rodrigo Adolfo. WiMAX – IEEE 802.16: Estudo da Tecnologia e Requisitos para a Modelagem e Simulação, 2006.
- [13] EKLUND, C., MARKES, R., STANWOOD, K., AND WANG, S. IEEE standard 802.16: A technical overview of the WirelessMAN air interface for broadband wireless access. IEEE Communications Magazine, 2002.
- [14] RIVEST R., SHAMIR A., ADLEMAN L. A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. 2 Edição. Ed. Instituto MIT. 1997.
- [15] HAWA, Mohammed et al. Quality of Service Broadband Wireless Access Systems, 2002.
- [16] KUROSE, J. and ROSS, K. Computer Networking – A Top-down Approach Featuring the Internet. Addison-Wesley, Massachusetts, third edition edition, 2004.
- [17] COELHO, J. Mecanismo de controle de qualidade de serviço em redes IEEE 802.11. Universidade Federal do Rio de Janeiro – COOPPE/PESC, 2003.
- [18] MACIEL, P. Modelagem e análise de um protocolo de acesso alternativo para o padrão IEEE 802.16 de redes metropolitanas sem fio. Universidade Federal do Rio de Janeiro – COOPPE/PESC, 2005.

- [19] WiMAX Forum. Disponível em <http://www.wimaxforum.org>.
- [20] CHEN, J., JIAO, W., GUO, Q. An integrated QoS control architecture for IEEE 802.16 broadband wireless access systems. Lucent Technologies), Bell Labs Research China, 2005.
- [21] WANG, L. et al, Admission Control for Non-preprovisioned service Flow in Wireless Metropolitan Area Networks. Fourth European Conference on Universal Multiservice Networks (ECUMN), pp. 243-249, 2007.
- [22] WANG, H., HE, B., AGRAWAL, D. P. Above packet layer level admission control and bandwidth allocation for IEEE 802.16 wireless MAN. Simulation Modeling Practice and Theory, vol. 15, no. 14, pg. 266-382, 2007.
- [23] NIYATO, D. HOSSAIN, E. A Queuing-Theoretic and Optimization-Based Model for Radio Resource Management in IEEE 802.16 Broadband Wireless Networks. IEEE Transactions on Computers, vol. 55, issue 11, pp. 1473-1488, November, 2006.
- [24] HAITANG, W., BING, H., DHARMA, A., Above packet level admission control and bandwidth allocation for IEEE 802.16 wireless MAN. In Proceedings of the 12<sup>th</sup> international Conference on Parallel and Distributed Systems – Volume 1. ICPADS. IEEE Computer Society, Washington, DC, 599-604, 2006.
- [25] LU, N., BIGHAM, J. 2005. Utility-maximization bandwidth adaptation for multi-class traffic QoS provisioning in wireless networks. In Proceedings of the 1<sup>st</sup> ACM international Workshop on Quality of Service & Security in Wireless and Mobile Networks. Q2Swinet '05. ACM, New York, NY, 136-143, 2005.
- [26] FREDERICK S. Lai, JELENA M., SAMUEL T. Chanson., Complete Sharing versus Partitioning: Quality of Service Management for Wireless Multimedia Networks. In Proceedings of the international Conference on

Computer Communications and Networks IC3N. IEEE Computer Society, Washington, DC, 584, 1998.

- [27] PUTERMAN, M. L. Markov Decision Processes. New York:Wiley, 1994.
- [28] LU, N., BIGHAM, J. Utility-based Bandwidth Adaptation for QoS Provisioning in Multimedia Wireless Networks, PhD thesis, 2007
- [29] LIMA L. Elon, A Equação do Terceiro Grau, Revista Matemática Universitária, No.5, 1987
- [30] TIJMS C. Henk, Stochastic Models: An Algorithmic Approach. Wiley, Chichester, 1994
- [31] CINLAR, Erhan, Introduction to stochastic processes. Prentice-Hall. New Jersey. US. 1975.
- [32] HORDIJK, A. & LOEVE, J.A. Undiscounted Markov decision chain with partial information; an algorithm for computing a locally optimal periodic policy. ZOR-Mathematical Methods of Operations Research, 40, 163-181, 1994.