

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E  
INFORMÁTICA INDUSTRIAL

EDEN RICARDO DOSCIATTI

**UMA NOVA ARQUITETURA PARA PROVISÃO DE QoS  
UTILIZANDO ENXAME DE PARTÍCULAS EM REDES WiMAX  
FIXAS**

TESE

CURITIBA

2015

EDEN RICARDO DOSCIATTI

**UMA NOVA ARQUITETURA PARA PROVISÃO DE QoS  
UTILIZANDO ENXAME DE PARTÍCULAS EM REDES WiMAX  
FIXAS**

Tese apresentada ao Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do grau de “Doutor em Ciências” – Área de Concentração: Telemática.

Orientadores: Prof. Dr. Walter Godoy Júnior (*i. m.*)  
Prof. Dr. Emílio Carlos Gomes Wille

Coorientador: Prof. Dr. Augusto Foronda

CURITIBA  
2015

---

#### Dados Internacionais de Catalogação na Publicação

---

D722n Dosciatti, Eden Ricardo  
Uma nova arquitetura para provisão de QoS utilizando enxame de partículas em redes WiMAX fixas/  
Eden Ricardo Dosciatti. – 2015.  
128 f. : il. ; 30 cm

Texto em Português, com resumo em Inglês.

Tese (Doutorado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. Curitiba, 2015.

Bibliografia: f. 122-127.

1. IEEE 802.16 (Normas). 2. Arquitetura de redes de computadores. 3. Qualidade de serviço (Redes de computação). 4. Inteligência artificial. 5. Otimização matemática. 6. Sistemas de comunicação em banda larga. 7. Sistemas de comunicação sem fio. 8. Métodos de simulação. 9. Engenharia elétrica - Teses.

I. Godoy Júnior, Walter (in memoriam), orient. II. Wille, Emílio Carlos Gomes, orient. III. Foronda, Augusto, coorient. IV. Universidade Tecnológica Federal do Paraná - Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial. V. Título.

CDD 22 - - 621.3

---

Biblioteca Central da UTFPR, Câmpus Curitiba

Título da Tese Nº. \_\_\_\_\_

# **Uma Nova Arquitetura Para Provisão De QoS Utilizando Enxame De Partículas Em Redes Wimax Fixas**

por

**Eden Ricardo Dosciatti**

**Orientador: Prof. Dr. Emílio Carlos Gomes Wille (UTFPR)**

**Coorientador: Prof. Dr. Augusto Foronda (UTFPR)**

Esta tese foi apresentada como requisito parcial à obtenção do grau de DOUTOR EM CIÊNCIAS – Área de Concentração: Telecomunicações e Rede, pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial – CPGEI – da Universidade Tecnológica Federal do Paraná – UTFPR, às 14h do dia 29 de abril de 2015. O trabalho foi aprovado pela Banca Examinadora, composta pelos doutores:

---

Prof. Dr. Augusto Foronda  
(Presidente – UTFPR)

---

Prof. Dr. Luiz Carlos Pessoa Albini  
(UFPR)

---

Prof. Dr. Marcelo Eduardo Pellenz  
(PUC)

---

Prof<sup>a</sup> Dr<sup>a</sup> Keiko Verônica Ono Fonseca  
(UTFPR)

---

Prof. Dr. Hermes Irineu Del Monego  
(UTFPR)

Visto da Coordenação:

---

Prof. Dr. Emílio Carlos Gomes Wille  
(Coordenador do CPGEI)

Pai e Mãe! Procurei por uma palavra que tentasse representar o meu sentimento neste momento. Por ter mania de sempre estar fazendo cálculos e pela razão de ter adquirido um raciocínio lógico e matemático, analisei e calculei a melhor palavra e a melhor frase para lhes expressar esse meu sentimento. Deduzi, por indução do coração junto com a alma, que a palavra é AMOR, e a frase é EU AMO VOCÊS. Sinto falta de vocês, do carinho de vocês, do jeito de vocês, das palavras, dos ensinamentos, das piadas e brincadeiras. Queria acordar e estar junto de vocês, mas o problema é que eu não estou dormindo, luto por uma vida, por um futuro, para ser o orgulho de vocês, onde quer que vocês estejam. A única coisa que posso fazer agora é lhes dar este presente, realizando um grande sonho teu meu Pai, que era ter um filho doutor, lógico que seu sonho era para ser doutor-médico, mas aqui está seu filho, um verdadeiro doutor. Por estes motivos dedico este trabalho a vocês, meu pai Protásio e minha mãe Deulinda. Também dedico este trabalho ao exemplar Prof. Dr. Walter Godoy Júnior, que me acolheu no CPGEI, desde o mestrado até sua partida (ainda me orienta com seu largo sorriso) e sempre me falava: “Acredite, é possível fazer pesquisa neste País, então faça”, e fizemos...

## AGRADECIMENTOS

Agradeço, com imensa alegria no coração e na alma, a Deus e a todos aqueles que acreditaram que um dia eu estaria digitando essas páginas.

Dentre eles, em primeiro lugar, a Família, base de nossa estrutura. Ao meu pai Protásio, já falecido a muitos anos. A minha querida e adorável mãe Deulinda que nos deixou neste ano de 2015 e abriu uma profunda lacuna em meu coração. A minha filha amada Lívia, minha verdadeira riqueza, que nasceu durante a realização deste Doutorado, e nos trouxe mais alegria, lógico, mais trabalho também, mas tudo vale a pena quando a alma não é pequena. Agradeço a minha esposa Mariza que também está neste barco chamado Doutorado, quase nos finalmente também, obrigado por chegarmos juntos ao final, porque no final tudo dá certo. A minha irmã Rita que defendeu seu Doutorado, em outra área, Agronomia, neste final de março. Ao cunhado Marcelo Serrão e a sobrinha Julinha que seguraram a barra da minha irmã com a partida prematura de nossa mãe. Ao irmão “Doutor” advogado Carlos, cunhada Gile e sobrinha Laura, que mesmo estando mais longe fisicamente, sempre estavam nos alegrando e dando incentivos de que um dia o final chegaria, e chegou... Ainda ao Tio Valda, que agora é o pai/mãe da Família, e aos primos André e Kellen que sempre estão ao lado nestes momentos que antecederam este momento. Aos meus sogros, Armerindo e Loraide, que foram fundamentais e importantes para este trabalho. Ao cunhado Rodrigo e a cunhada Elisangela, obrigado pelas visitas e pela força sempre para frente.

Aos amigos, segunda parte da nossa estrutura, aos antigos que permanecem em nossas lembranças quando algo nos aperta, e aos novos que estudam, trabalham e se divertem conosco. Ao Omero Bertol (Juca) e Vanessa, das antigas, ele colega de UTFPR Pato Branco e colega de Doutorado, obrigado por tudo e que fase! Ao ex-colega de UTFPR Pato Branco, Edenilson, das antigas, que agora está servindo a UTFPR Curitiba, obrigado. Fernando Schutz, das antigas, da UTFPR Medianeira, colega de Doutorado e parceiro para todas as obras. Ao Dr. Lotário Fank, das antigas, da UTFPR Medianeira, pagador de ovelhas, obrigado. Ao aposentado Paulo Ferrazza, das antigas, que muitas dicas me deu para escrever a tese, claro, em torno de uma mesa na churrasqueira.

Aos colegas do NATEC e arredores, Foronda, Kleber, Lincoln, Danillo, Joilson, Ronska, Guilherme, Rubão, Hermes, Christian, Marcos, Euclides, Talau, ao filósofo estagiário

Vilmar e ao amigo do Tocantins, Brunão, obrigado a todos.

Aos colegas do DAINF de Pato Branco, que trabalharam um pouco a mais para possibilitar meu afastamento integral.

Ao Prof. Dr. Emílio Wille, que com a repentina partida do Prof. Godoy, nos deu o aporte de confiança para terminar este trabalho.

Ao Prof. Dr. Augusto Foronda, incansável coorientador-orientador que desde o mestrado vem me acompanhando e me assumiu integralmente na reta final e foi fundamental para chegar neste momento, sempre sugerindo e orientando.

Ao Prof. Dr. Walter Godoy Júnior (*in memoriam*), por ter me adotado ainda no mestrado e por ter acreditado que eu conseguiria. Foi uma pena a sua partida, em todos os sentidos, amigo, colega, orientador. Sentimos muita falta daquele sorriso largo quando uma piada era contada.

E, finalmente, agradeço a mim! Lutei, cai, perdi; voltei a lutar, levantei, venci. Sempre acreditei nas palavras que o Prof. Dr. Kleber Nabbas proferia: “no final tudo dá certo, se não deu certo ainda, é porque não é o final”. Agora é o final...

*“Maior que a tristeza de não haver vencido é a vergonha de não ter lutado!”* **Rui Barbosa**



## RESUMO

DOSCIATTI, Eden Ricardo. UMA NOVA ARQUITETURA PARA PROVISÃO DE QoS UTILIZANDO ENXAME DE PARTÍCULAS EM REDES WiMAX FIXAS. 128 f. Tese – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2015.

Os avanços tecnológicos ocorridos nos últimos anos provocaram um crescimento na base de usuários que utilizam as redes de comunicação, principalmente com serviços de multimídia, como IPTV, videoconferência e VoIP. Como estes serviços requerem mais recursos e geram uma grande demanda sobre a infraestrutura da rede, cada usuário deve ter as suas aplicações tratadas de acordo com determinadas prioridades, para oferecer um nível aceitável de serviço. Partindo deste pressuposto, as redes de comunicação para acesso sem fio em banda larga, baseadas no padrão IEEE 802.16, também conhecidas como WiMAX, conseguem atender as várias demandas dos usuários finais, como a necessidade de acesso aos dados em todos os momentos e em qualquer lugar e a conectividade de banda larga eficiente, oferecendo uma excelente relação custo-benefício para o utilizador final, pois possibilitam uma elevada capacidade de transmissão de dados a um custo relativamente baixo de implantação. A provisão de qualidade de serviços é um fator de grande importância para o desempenho das redes de comunicação, para isso, os mecanismos de escalonamento, de controle de admissão de conexões e de policiamento de tráfego, devem estar presentes. Porém, o padrão IEEE 802.16 apenas especifica as políticas de como estes mecanismos devem ser implementados, mas não define como implementá-los. Esta pesquisa tem como objetivo apresentar uma nova arquitetura para o tráfego *uplink*, nas estações bases das redes WiMAX, com provisão de qualidade de serviços, utilizando a metaheurística de otimização por enxame de partículas para o cálculo da duração do tempo do quadro, quando houver a necessidade, possibilitando encontrar o valor ideal para esta quantidade, proporcionando uma melhor alocação de usuários na rede.

**Palavras-chave:** IEEE 802.16, WiMAX, QoS, PSO, Tempo do Quadro, Escalonamento

## ABSTRACT

DOSCIATTI, Eden Ricardo. A NEW ARCHITECTURE TO GUARANTEE QoS WITH THE USE OF PARTICLE SWARM OPTIMIZATION IN FIXED WiMAX NETWORKS. 128 f. Tese – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2015.

Technological advances in recent years have led to a growth in the user base using communication networks, especially with multimedia services such as IPTV, video conferencing and VoIP. As these services require more resources and generate a great demand on the network infrastructure, each user must have their applications dealt with in accordance with certain priorities, to provide an acceptable level of service. Under this assumption, the communication networks for wireless broadband, based on the IEEE 802.16 standard, also known as WiMAX, can meet the various demands of end users, the need for access to data at all times and in any place and the connectivity efficient broadband, offering excellent tradeoff for the end user by enabling high data transmission capacity at a relatively low cost of deployment. The provision of quality services is a very important factor for the performance of communication networks for that, the scheduling mechanisms, admission control connections and traffic policing, must be present. However, the IEEE 802.16 standard specifies only policies of how these mechanisms should be implemented, but does not define how to implement them. This research aims to present a new architecture for uplink traffic, the WiMAX network base stations, with provision of quality services through the particle swarm optimization metaheuristic to calculate the frame duration, allowing to find an ideal value, providing a better allocation of network users.

**Keywords:** IEEE 802.16, WiMAX, QoS, PSO, Time Frame, Scheduling

## LISTA DE FIGURAS

FIGURA 1	– Interpretação geométrica do PSO	36
FIGURA 2	– Busca da melhor solução pelo <i>pbest</i> e <i>gbest</i>	37
FIGURA 3	– Passos de um algoritmo PSO	39
FIGURA 4	– Busca da melhor solução pelo <i>pbest</i> e <i>gbest</i> em problemas com restrições	40
FIGURA 5	– Topologia de uma rede WiMAX	44
FIGURA 6	– Modelo de referência do padrão IEEE 802.16	45
FIGURA 7	– Estrutura do símbolo OFDM no domínio do tempo	50
FIGURA 8	– Símbolo OFDM no domínio da frequência	51
FIGURA 9	– Estrutura de um quadro no esquema de transmissão TDD	52
FIGURA 10	– Formato do subquadro de <i>downlink</i>	53
FIGURA 11	– Formato do subquadro de <i>uplink</i>	54
FIGURA 12	– Técnicas de duplexação para alocação de banda no padrão IEEE 802.16	55
FIGURA 13	– Arquitetura de QoS nativa do padrão IEEE 802.16	58
FIGURA 14	– <i>Polling unicast</i>	60
FIGURA 15	– <i>Polling</i> Baseado em Contenção	60
FIGURA 16	– Mecanismo de alocação de concessões <i>uplink</i> UGS	61
FIGURA 17	– Mecanismo de alocação de concessões <i>uplink</i> rtPS	62
FIGURA 18	– Mecanismo de alocação de concessões <i>uplink</i> ertPS	62
FIGURA 19	– Mecanismo de alocação de concessões <i>uplink</i> nrtPS	63
FIGURA 20	– Mecanismo de alocação de concessões <i>uplink</i> BE	63
FIGURA 21	– Filas na disciplina de escalonamento PQ	67
FIGURA 22	– Arquitetura do escalonador <i>uplink</i> proposto por (JAKIMOSKI; JANEVSKI, 2013)	69
FIGURA 23	– Estrutura hierárquica de alocação de banda	72
FIGURA 24	– Algoritmos de escalonamento <i>uplink</i> para redes WiMAX	74
FIGURA 25	– Modelo da nova arquitetura - Estação Base - BS	77
FIGURA 26	– Modelo da nova arquitetura - Subquadros <i>Downlink</i> e <i>Uplink</i>	78
FIGURA 27	– Modelo de um escalonador <i>LR</i>	80
FIGURA 28	– Alocação dos parâmetros na estrutura de um quadro TDD	82
FIGURA 29	– Tráfego de entrada modelado por um balde de fichas	82
FIGURA 30	– Aproximação das duas restrições para o cálculo do <i>TQ</i> Ideal	86
FIGURA 31	– Prioridades de atendimento de requisições	91
FIGURA 32	– Cenário de avaliação de uma rede para o cálculo do <i>TQ</i> Ideal	101
FIGURA 33	– Duração do <i>TQ</i> Ideal para o tráfego de áudio	102
FIGURA 34	– Duração do <i>TQ</i> Ideal para o tráfego VBR	103
FIGURA 35	– Duração do <i>TQ</i> Ideal para o tráfego MPEG4	104
FIGURA 36	– Número de execuções para encontrar o <i>TQ</i> Ideal para o tráfego de áudio	105
FIGURA 37	– Número de execuções para encontrar o <i>TQ</i> Ideal para o tráfego de vídeo VBR	105
FIGURA 38	– Número de execuções para encontrar o <i>TQ</i> Ideal para o tráfego de vídeo MPEG4	106
FIGURA 39	– Tempo de execução para encontrar o <i>TQ</i> Ideal para o tráfego de áudio	107
FIGURA 40	– Tempo de execução para encontrar o <i>TQ</i> Ideal para o tráfego de vídeo VBR	107
FIGURA 41	– Tempo de execução para encontrar o <i>TQ</i> Ideal para o tráfego de vídeo MPEG4	108
FIGURA 42	– Atraso médio de 10 conexões UGS com diferentes cargas de tráfego ertPS	112

FIGURA 43	– Atraso médio de 10 conexões rtPS com diferentes cargas de tráfego ertPS .....	113
FIGURA 44	– Variação média do atraso de 10 conexões UGS e variação média do atraso de conexões ertPS de 5 para 30 estações clientes (SSs) .....	113
FIGURA 45	– Variação média do atraso de 10 conexões rtPS com diferentes cargas de tráfego ertPS ..	114
FIGURA 46	– Taxa útil média de 10 conexões UGS e taxa útil média de conexões ertPS variando de 5 para 30 estações clientes (SSs) .....	114
FIGURA 47	– Taxa útil média de 10 conexões rtPS com diferentes cargas de tráfego ertPS .....	115
FIGURA 48	– Taxa útil média de 10 conexões nrtPS e 10 conexões BE com diferentes cargas de tráfego ertPS .....	116
FIGURA 49	– Aumento do número de conexões ertPS .....	116

## LISTA DE TABELAS

TABELA 1	– <i>Resumo das interfaces físicas</i> .....	49
TABELA 2	– <i>Duração do tempo de um quadro OFDM</i> .....	52
TABELA 3	– <i>Resumo das classes de serviço para redes padrão IEEE 802.16</i> .....	64
TABELA 4	– <i>Prioridades das Classes de Serviço WiMAX proposto por (JAKIMOSKI; JANEVSKI, 2013)</i> .....	68
TABELA 5	– <i>Símbolos com suas unidades de medida para a nova arquitetura</i> .....	80
TABELA 6	– <i>Configuração utilizada para o exemplo da Figura 30</i> .....	85
TABELA 7	– <i>Parâmetros para o cálculo de um símbolo OFDM</i> .....	88
TABELA 8	– <i>Parâmetros do algoritmo PSO para o cálculo do TQ Ideal</i> .....	94
TABELA 9	– <i>Descrição de diferentes tráfegos</i> .....	101
TABELA 10	– <i>Parâmetros das camadas PHY e MAC</i> .....	101
TABELA 11	– <i>Parâmetros do balde de fichas</i> .....	102
TABELA 12	– <i>Tempos para calcular o quadro com PSO e com o método de busca exaustiva no tráfego de áudio</i> .....	106
TABELA 13	– <i>Tempos para calcular o quadro com PSO e com o método de busca exaustiva no tráfego de vídeo VBR</i> .....	108
TABELA 14	– <i>Tempos para calcular o quadro com PSO e com o método de busca exaustiva no tráfego de vídeo MPEG4</i> .....	108
TABELA 15	– <i>Principais parâmetros de configuração da rede</i> .....	109
TABELA 16	– <i>Características dos traces de vídeos utilizados nas simulações</i> .....	110
TABELA 17	– <i>Valor do TQ Ideal em relação ao tempo do quadro de (JAKIMOSKI; JANEVSKI, 2013)</i> .....	111
TABELA 18	– <i>Resumo do gráfico apresentado na Figura 49, com o valor do atraso médio</i> .....	117

## LISTA DE SIGLAS

ACO	<i>Ant Colony Optimization</i>
ADRR	<i>Adaptive Deficit Round Robin</i>
AEs	<i>Algoritmos Evolutivos</i>
AMC	<i>Adaptive Modulation and Coding</i>
Anatel	<i>Agência Nacional de Telecomunicações</i>
ATM	<i>Asynchronous Transfer Mode</i>
BE	<i>Best Effort</i>
BS	<i>Base Station</i>
BWA	<i>Broadband Wireless Access</i>
CBR	<i>Constante Bit Rate</i>
CID	<i>Connection Identifier</i>
CP	<i>Cyclic Prefix</i>
CPS	<i>Common Part Sublayer</i>
CS	<i>Convergence Sublayer</i>
DC	<i>Direct Current</i>
DFS	<i>Dynamic Frequency Selection</i>
DRR	<i>Deficit Round Robin</i>
DSL	<i>Digital Subscriber Line</i>
EDF	<i>Earliest Deadline First</i>
ertPS	<i>extended real-time Polling Service</i>
EVRC	<i>Enhanced Variable Rate Codec</i>
FDD	<i>Frequency Division Duplexing</i>
FDDI	<i>Fiber Distributed Data Interface</i>
FFT	<i>Fast Fourier Transform</i>
FIFO	<i>First In-First Out</i>
FTP	<i>File Transfer Protocol</i>
ICI	<i>Inter-Carrier Interference</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IE	<i>Information Element</i>
IFFT	<i>Inverse Fast Fourier Transform</i>
IPTV	<i>Internet Protocol TV</i>
ISI	<i>Inter-Symbol Interference</i>
MMS	<i>Multimedia Messaging Service</i>
MPEG	<i>Moving Pictures Experts Group</i>
MPLS	<i>Multi Protocol Label Switching</i>
MRTR	<i>Minimum Reserved Traffic Rate</i>
MSTR	<i>Maximum Sustained Traffic Rate</i>
NLoS	<i>Non-line-of-sight</i>
nrtPS	<i>non-real-time Polling Service</i>
NS-2	<i>Network Simulator - version 2</i>
OFDM	<i>Orthogonal Frequency Division Multiplexing</i>

OSI	<i>Open Systems Interconnection</i>
PDU <sub>s</sub>	<i>Protocol Data Units</i>
PHY	<i>Physical Layer</i>
PMP	<i>Point-to-Multipoint</i>
PS	<i>Physical Slots</i>
PSO	<i>Particle Swarm Optimization</i>
QoS	<i>Quality of Service</i>
RF	<i>Radiofrequência</i>
RR	<i>Round Robin</i>
RTG	<i>Receiver-transmitter Turnaround Gap</i>
rtPS	<i>real-time Polling Service</i>
SDU <sub>s</sub>	<i>Service Data Units</i>
SFID	<i>Service Flow Identifier</i>
SI	<i>Swarm Intelligence</i>
SMS	<i>Short Message Service</i>
SS	<i>Subscriber Station</i>
TDD	<i>Time Division Duplexing</i>
<i>TQ</i>	<i>Tempo do Quadro</i>
TTG	<i>Transmitter-receiver Turnaround Gap</i>
UGS	<i>Unsolicited Grant Service</i>
VoIP	<i>Voice over Internet Protocol</i>
WiMAX	<i>Worldwide Interoperability for Microwave Access</i>
WirelessMAN-OFDM	<i>Wireless Metropolitan Area Networks-OFDM</i>
WMAN <sub>s</sub>	<i>Wireless Metropolitan Area Networks</i>

## LISTA DE SÍMBOLOS

$A_i$	Tráfego de Entrada
$BR_i$	Taxa do Balde de Fichas
$BT_i$	Tamanho do Balde de Fichas
$BW_{request}$	Segmento de Requisição de Largura de Banda
$c_1$	Parâmetro Cognitivo
$c_2$	Parâmetro Social
$D_{bound}$	Limite de Atraso Superior
$\Delta$	Sobrecarga do Subquadro de <i>Uplink</i>
$D$	Dimensão
$D_i$	Atraso Máximo na Fila
$D_{max,i}$	Atraso Máximo Requerido pelo Usuário
$gbest^k$	<i>Global Best</i>
$i$	Solução Candidata
$k$	Contador de Iterações
$L_{max,i}$	Tamanho Máximo de um Pacote
$pbest_i^k$	<i>Personal Best</i>
$r_1$	Vetor Aleatório do Parâmetro Cognitivo
$r_2$	Vetor Aleatório do Parâmetro Social
$R$	Taxa Física do Canal
$r_i$	Taxa Alocada pelo Servidor
$RI_{inic}$	Segmento de Registro Inicial
$s$	Tamanho do Enxame
$T_{DL}$	Tempo do Subquadro de <i>Downlink</i>
$\theta_i$	Latência
$t$	Período de Tempo
$TQ$	Tempo do Quadro
$TQ_{max}$	Tempo do Quadro Máximo
$TQ_{min}$	Tempo do Quadro Mínimo
$T_{RTG}$	Intervalo de Tempo entre o Subquadro de <i>Uplink</i> e <i>Downlink</i>
$T_{TTG}$	Intervalo de Tempo entre o Subquadro de <i>Downlink</i> e <i>Uplink</i>
$TU_i$	Tempo que cada SS tem para Transmitir seus Pacotes de Dados
$T_{UL}$	Tempo do Subquadro de <i>Uplink</i>
$v_i^k$	Velocidade Atual da Partícula $i$
$w$	Fator de Inércia
$w_{max}$	Peso Inercial Máximo
$w_{min}$	Peso Inercial Mínimo
$x_i^k$	Posição Atual da Partícula $i$



## LISTA DE ALGORITMOS

ALGORITMO 1	– Pseudocódigo do algoritmo PSO .....	38
ALGORITMO 2	– Pseudocódigo do algoritmo PSO com restrições .....	41
ALGORITMO 3	– Algoritmo <i>Principal</i> .....	93
ALGORITMO 4	– módulo <i>TQIdeal</i> .....	93
ALGORITMO 5	– módulo <i>TQ_PSO</i> .....	95
ALGORITMO 6	– módulo <i>SimboloOFDM</i> .....	97
ALGORITMO 7	– módulo <i>PrimeiraRestricao</i> .....	97
ALGORITMO 8	– módulo <i>SegundaRestricao</i> .....	97
ALGORITMO 9	– módulo <i>Segmentos_SimboloOFDM</i> .....	98

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>18</b>
1.1	RELEVÂNCIA DO TEMA	19
1.2	CARACTERIZAÇÃO DO PROBLEMA	20
1.3	OBJETIVO	20
1.4	RELEVÂNCIA	22
1.5	CONTRIBUIÇÕES	23
1.6	PUBLICAÇÕES	23
1.7	ORGANIZAÇÃO DA TESE	24
<b>2</b>	<b>OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS - PSO</b>	<b>25</b>
2.1	OTIMIZAÇÃO E MÉTODOS DE OTIMIZAÇÃO	25
2.1.1	Algoritmos de Otimização	27
2.2	INTELIGÊNCIA DE ENXAME	30
2.3	OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS	32
2.3.1	Modelo Matemático do PSO	33
2.3.2	Algoritmo PSO	38
2.3.3	Algoritmo PSO com Restrição	39
2.4	CONCLUSÃO DO CAPÍTULO	41
<b>3</b>	<b>O PADRÃO IEEE 802.16 E TRABALHOS RELACIONADOS</b>	<b>43</b>
3.1	AS REDES NO PADRÃO IEEE 802.16	44
3.1.1	Topologia de uma Rede no Padrão IEEE 802.16	44
3.1.2	Modelo de Referência do Padrão IEEE 802.16	45
3.1.2.1	Camada de Controle de Acesso ao Meio (MAC)	46
3.1.2.2	Camada Física (PHY)	47
3.1.2.3	Faixas de Frequência Suportadas	48
3.1.2.4	Especificações da Interface Física	49
3.1.3	A Construção de um Símbolo OFDM	50
3.1.3.1	Símbolo OFDM no Domínio do Tempo	50
3.1.3.2	Símbolo OFDM no Domínio da Frequência	51
3.1.4	Estrutura de um Quadro no Padrão IEEE 802.16	52
3.1.4.1	Subquadro de <i>Downlink</i>	53
3.1.4.2	Subquadro de <i>Uplink</i>	53
3.1.4.3	Modos de Acesso ao Meio Físico	54
3.2	QUALIDADE DE SERVIÇO PARA REDES IEEE 802.16	55
3.2.1	Métricas e Parâmetros para Provisão de QoS em Redes Padrão IEEE 802.16	55
3.2.2	QoS para as Classes de Serviço das Redes Padrão IEEE 802.16	57
3.2.2.1	UGS	60
3.2.2.2	rtPS	61
3.2.2.3	ertPS	61
3.2.2.4	nrtPS	62
3.2.2.5	BE	63
3.3	TRABALHOS RELACIONADOS DE ESCALONAMENTO EM REDES PADRÃO	

IEEE 802.16 .....	64
3.3.1 Mecanismo de Escalonamento para Redes Padrão IEEE 802.16 .....	65
3.3.2 Classificação dos Algoritmos de Escalonamento .....	66
3.3.2.1 Algoritmos de Escalonamento Homogêneos .....	66
3.3.2.2 Algoritmos de Escalonamento Híbridos .....	72
3.4 CONCLUSÃO DO CAPÍTULO .....	74
<b>4 UMA NOVA ARQUITETURA PARA PROVISÃO DE QoS EM REDES WIMAX</b>	
<b>FIXAS .....</b>	<b>76</b>
4.1 MODELO DA NOVA ARQUITETURA .....	77
4.2 CAMADA 1 - CÁLCULO DA DURAÇÃO DO TEMPO DO QUADRO IDEAL ...	79
4.2.1 Modelo Analítico para o Cálculo do Tempo do Quadro ( $TQ$ ) .....	79
4.2.2 Otimização do Tempo do Quadro utilizando Enxame de Partículas .....	86
4.3 CAMADA 2 - ALOCAÇÃO DE RECURSOS .....	87
4.4 CAMADA 3 - CONTROLE DE ADMISSÃO DE CONEXÕES (CAC) .....	91
4.5 PSEUDOCÓDIGOS DA NOVA ARQUITETURA .....	92
4.6 CONCLUSÃO DO CAPÍTULO .....	98
<b>5 ANÁLISE DE DESEMPENHO DA NOVA ARQUITETURA .....</b>	<b>100</b>
5.1 CENÁRIOS E PARÂMETROS DO AMBIENTE DE AVALIAÇÃO E DE	
SIMULAÇÃO .....	100
5.1.1 Cenário 1 - Cálculo do TQ Ideal .....	100
5.1.1.1 Resultados Obtidos com a Primeira Avaliação .....	102
5.1.2 Cenário 2 - Utilização do $TQ$ Ideal para Validação da Nova Arquitetura .....	109
5.1.2.1 Resultados Obtidos com a Segunda Avaliação .....	111
5.2 CONCLUSÃO DO CAPÍTULO .....	117
<b>6 CONCLUSÕES .....</b>	<b>119</b>
6.1 TRABALHOS FUTUROS .....	121
<b>REFERÊNCIAS .....</b>	<b>122</b>
<b>Apêndice A – OBTENÇÃO DA EQUAÇÃO (9) .....</b>	<b>128</b>

## 1 INTRODUÇÃO

O acesso à Internet banda larga é um serviço de demanda globalizada e em plena expansão. Segundo Silva e Biondi (2012), o grande desafio é democratizar este acesso aos usuários de Internet. Diante de uma rede de comunicação em expansão e pelo fato da crescente integração de mídias que convergem para interagir em um único ambiente, a qualidade da conexão se tornou um parâmetro indispensável à vida moderna. Assim, a Internet banda larga, representa uma condição estratégica para todos, capaz de garantir que a experiência dos usuários ocorra de modo pleno e satisfatório.

O surgimento de dispositivos tecnológicos de última geração que utilizam o acesso sem fio, permitiu uma mudança no perfil de acesso às redes de computadores, gerando uma forte demanda por aplicações com tráfego heterogêneo (voz, vídeo e dados) (ROSA; GUARDIERO, 2011). Esses equipamentos são utilizados em todos os setores, tanto na indústria e no comércio, como na educação e na pesquisa, nas comunicações e, até mesmo, no lazer e entretenimento (ARHAIF, 2011). Com esses novos equipamentos e suas tecnologias de acesso sem fio à disposição, e o surgimento de um mercado de usuários corporativos e residenciais, a grande batalha é prover acesso de Internet banda larga de forma eficiente, com baixo custo de implantação e manutenção e possibilitar que o usuário tenha a informação disponível a qualquer hora e em qualquer lugar.

A solução tradicional para fornecer o acesso à Internet banda larga em redes de alta velocidade é a utilização de tecnologias de acesso com fio, tais como cabo, DSL (*Digital Subscriber Line*) e fibra óptica. Porém, o alto investimento em infraestrutura, em regiões mais isoladas ou com dificuldades de acesso, por vezes torna inviável a implantação destas tecnologias.

Uma solução possível e viável, para reduzir os custos de implantação de acesso à Internet banda larga em áreas onde a infraestrutura não está presente, é a utilização de tecnologias sem fio, conhecidas como BWA (*Broadband Wireless Access*) (rádio, WiMAX, LTE, satélite, 3G e 4G), que não se utilizam de cabos e reduzem o custo e o tempo de

implantação (DOSCIATTI et al., 2012). As redes sem fio são de uma importância fundamental, pois oferecem a implantação rápida e o acesso fácil, possuem alta escalabilidade e um baixo custo de infraestrutura, manutenção e atualização (PAPAPANAGIOTOU et al., 2009).

## 1.1 RELEVÂNCIA DO TEMA

As redes de computadores que utilizam a tecnologia WiMAX (*Worldwide Interoperability for Microwave Access*), e em tradução livre, **Interoperabilidade Global para Acesso por Microondas**, possibilitam o acesso à Internet banda larga de alta velocidade sem fio para as WMANs (*Wireless Metropolitan Area Network*), ou redes metropolitanas sem fio (TAGHIPOOR et al., 2012).

WiMAX é o nome comercial utilizado para agrupar um número de tecnologias sem fio que surgiram nos laboratórios do IEEE (*Institute of Electrical and Electronics Engineers*) com a finalidade de atender as várias demandas dos usuários finais, como a mobilidade, a necessidade de acesso a dados em todos os momentos e em qualquer lugar, e a conectividade de banda larga eficiente (ARHAIF, 2011).

O padrão IEEE 802.16 (IEEE, 2012) especifica políticas de escalonamento, políticas de controle de tráfego e políticas de controle de conexão, porém, não estabelece como implementar essas políticas para as redes WiMAX. Esta tecnologia foi escolhida para a implementação deste trabalho de pesquisa pelo fato de ser uma especificação aberta e não possuir uma arquitetura definida. Além disso, outros atrativos prevalecem nas redes WiMAX, como a infraestrutura básica para montagem da rede, a baixa latência, o avançado provisionamento de qualidade de serviço (QoS - *Quality of Service*) e a segurança sofisticada para o tráfego dos dados (KOLIAS et al., 2013).

Levando em consideração que os avanços ocorridos na Internet e, principalmente, nas comunicações móveis, resultaram em um crescimento acentuado na base de usuários, provocando uma tendência contínua de aumento da utilização de serviços de multimídia, tais como IPTV (*Internet Protocol TV*), videoconferência e VoIP (*Voice over Internet Protocol*). Estes serviços requerem enormes recursos e geram uma grande demanda sobre a infraestrutura da rede, criando uma necessidade de redes mais rápidas e mais confiáveis. A maioria das aplicações que trafegam na rede, são aplicações multimídia, pois envolvem a transferência de múltiplos tipos de mídia (dados, voz e vídeo) com requisito de tempo e sincronização para a sua operação com qualidade. Assim, a QoS nas redes WiMAX é um aspecto operacional fundamental para o desempenho fim-a-fim dessas aplicações, sendo importante o entendimento

dos seus princípios, parâmetros, mecanismos, algoritmos e protocolos desenvolvidos e utilizados para a obtenção de QoS.

Desta forma, uma característica incorporada pelo padrão IEEE 802.16 é o tratamento diferenciado do tráfego gerado pelos aplicativos, essenciais para o provisionamento de QoS para os diferentes fluxos de serviços, com diversos requisitos de QoS. As redes WiMAX são capazes de tratar as aplicações de acordo com as suas prioridades, oferecendo taxas de dados muito elevadas ao longo de grandes distâncias, incorporando uma estrutura bem definida de QoS (AHMED; HAMMA, 2011). Além disso, o padrão IEEE 802.16 requer políticas de escalonamento, de controle de tráfego e de controle de conexão para que a arquitetura de provisionamento de QoS seja completa. No entanto, como estes mecanismos não estão definidos na norma, a tecnologia WiMAX se torna um campo aberto para o desenvolvimento de pesquisas, tanto por parte da indústria, bem como pela comunidade acadêmica. Como os mecanismos não estão especificados na norma, diferentes soluções podem existir, desde que atendam aos requisitos de QoS.

## 1.2 CARACTERIZAÇÃO DO PROBLEMA

A grande motivação do IEEE é desenvolver uma tecnologia de QoS de ponta a ponta, com escalabilidade e elevadas taxas de transmissão de dados para oferecer serviços de qualquer tipo, a qualquer momento e em qualquer lugar. O escopo deste trabalho de pesquisa, é o desenvolvimento de uma nova arquitetura para a provisão de QoS para as redes WiMAX fixas.

## 1.3 OBJETIVO

Este trabalho de pesquisa tem como objetivo o desenvolvimento de uma nova arquitetura para o escalonador da estação base, no sentido do tráfego *uplink*. A função principal desta nova arquitetura é a provisão de QoS para as cinco classes de serviço especificadas pelo padrão IEEE 802.16, com a garantia de recursos para as classes com tráfego em tempo real e com garantia de recursos mínimos necessários para as classes com tráfego que não necessitem ser em tempo real. A nova arquitetura é composta por três camadas de processamento, executadas de forma hierárquica.

A *Camada 1* tem a função de identificar a melhor utilização da rede através do cálculo da duração do tempo do quadro ( $TQ$ ). Um quadro é formado por segmentos físicos, sendo que cada segmento físico tem um determinado tempo, assim, a soma dos segmentos físicos, determina o tempo de um quadro. Desta forma, toda vez que o algoritmo de escalonamento é

executado, se houver a necessidade de mudança do número de usuários na rede, por conta do tipo de tráfego de entrada, um novo cálculo do  $TQ$  é efetuado. Conseqüentemente, uma nova duração do  $TQ$  pode ser definida, tendo a possibilidade de que mais usuários sejam alocados no sistema, tornando a rede mais eficiente. O cálculo da duração do  $TQ$  é baseado na teoria do servidor *Latency-Rate* (LR) (STILIADIS; VARMA, 1998), utilizando o modelo analítico desenvolvido no trabalho de (DOSCIATTI, 2010). Para calcular o  $TQ$  Ideal, a metaheurística de otimização por enxame de partículas (*Particle Swarm Optimization* - PSO) (KENNEDY et al., 2001) é utilizada nesta arquitetura, por três fatores: (i) é um algoritmo simples, pois não existem muitos parâmetros para serem ajustados; (ii) é um algoritmo poderoso, pelo fato de ser mais rápido e lidar com vários tipos de problemas de otimização, onde as funções podem ter restrições; (iii) não tem um fator limitante no que diz respeito a função objetivo e as restrições, pois não precisa pré-processar o objetivo e as restrições do problema a ser resolvido. Com o  $TQ$  Ideal calculado, o sistema pode utilizar o quadro com mais eficácia, possibilitando a alocação de mais usuários em um mesmo quadro, evitando que seja transmitido com segmentos físicos vazios, ou seja, com segmentos que não foram alocados para nenhum usuário. A utilização do PSO para o cálculo do  $TQ$  Ideal tem um fator de extrema importância na diminuição do esforço computacional, ocorrendo uma redução da utilização de recursos computacionais durante o cálculo do  $TQ$  Ideal.

A *Camada 2* tem a função de alocar os recursos disponíveis para a transmissão de dados no sentido *uplink* em uma rede WiMAX. Esses recursos devem ser divididos considerando a capacidade de transmissão de cada segmento no subquadro de *uplink*. Para isso, esta camada executa o cálculo de um símbolo OFDM (*Orthogonal Frequency Division Multiplexing*), que é uma pequena porção de informação, sendo possível determinar o valor do segmento físico, que é a capacidade de transmissão da unidade de alocação utilizada no subquadro de *uplink*.

A *Camada 3* realiza o controle de admissão de conexões (CAC) para gerenciar a QoS de uma rede WiMAX, ou seja, define quais requisições devem ser atendidas. Para esta camada, uma nova arquitetura foi desenvolvida para o tratamento das classes de serviço nas filas de escalonamento da BS. Esta nova arquitetura é composta por 3 filas de escalonamento, com a agregação de prioridades para as cinco classes de serviços especificadas pelo padrão IEEE 802.16.

Para atingir o objetivo deste trabalho de pesquisa, as seguintes ações são executadas:

1. **Otimização do tempo do quadro com a utilização da metaheurística PSO.** Para otimizar o  $TQ$ , em cada iteração do sistema, quando for necessário, a nova arquitetura

calcula um tempo ideal de duração de um quadro, utilizando o algoritmo de otimização PSO.

2. **Definição de uma nova arquitetura e especificação dos algoritmos.** A nova arquitetura é baseada em uma estrutura hierárquica, dividida em três camadas, com um modelo de 3 filas de escalonamento, com a agregação de prioridades para cada classe de serviço definida pelo padrão IEEE 802.16. Os algoritmos para otimização do tempo do quadro utilizando enxame de partículas em redes WiMAX fixas com o objetivo de provisão de QoS são especificados e descritos.
3. **Definição da utilização de um mecanismo de admissão de conexão (CAC).** O modelo de controle de admissão de conexões é definido, baseado na iteração com a nova arquitetura desenvolvida nos itens descritos acima.
4. **Avaliação do desempenho da nova arquitetura com a utilização do PSO.** A nova arquitetura é avaliada em relação a utilização da metaheurística PSO no cálculo da duração do tempo do quadro, com o objetivo de provisão de QoS e a priorização do tráfego nas três filas de escalonamento na nova arquitetura desenvolvida. Essa avaliação tem o objetivo de demonstrar a eficiência da utilização do PSO quando comparado com o modelo de cálculo do tempo do quadro realizado através de busca exaustiva (“força bruta”). Também mostra que o algoritmo PSO reduz o esforço computacional e o tempo de execução para o cálculo do tempo do quadro.
5. **Comparação da avaliação do desempenho da nova arquitetura.** A nova arquitetura tem seu desempenho avaliado através de simulação, com a utilização de diferentes cenários e com diferentes cargas de tráfego. Os resultados obtidos são comparados com os resultados apresentados no trabalho de (JAKIMOSKI; JANEVSKI, 2013), que utiliza três filas de escalonamento com cinco níveis de prioridade, uma para cada classe de serviço.

#### 1.4 RELEVÂNCIA

A relevância do trabalho desenvolvido reside no fato do desenvolvimento de uma nova arquitetura para provisão de QoS para as redes WiMAX fixas. O modelo da nova arquitetura é dividido em três camadas hierárquicas, possibilitando ter um maior controle das filas do escalonador da estação base, utilizando um método para o cálculo do tempo do quadro, através da aplicação da metaheurística PSO, possibilitando, com isso, que mais usuários tenham acesso ao sistema, com um menor esforço computacional.



Este trabalho também possui conteúdo inovador no sentido de que mostra os resultados da aplicação inédita do PSO no modelo de arquitetura de escalonamento definido pelo padrão IEEE 802.16. Isso possibilita uma melhoria na alocação de usuários em cada quadro, evitando que ocorra o desperdício de banda, pelo fato de existir a possibilidade de transmissão de quadros sem segmentos físicos vazios, com isso, a provisão de QoS é garantida para todas as classes de serviço especificadas na norma.

## 1.5 CONTRIBUIÇÕES

As contribuições presentes neste trabalho são:

- Modificação do modelo do trabalho de (STILIADIS; VARMA, 1998), que, além de garantir o atraso requerido pelo usuário de tempo real, vai garantir a largura de banda mínima para os usuários de tempo não real.
- Utilização do PSO para calcular a duração ideal do tempo de um quadro, possibilitando que mais usuários tenham acesso ao sistema com um menor esforço computacional.
- Pesquisa bibliográfica para a apresentação do estado da arte relacionado ao desenvolvimento dos modelos de mecanismos de escalonamento no tráfego *uplink* em redes baseadas no padrão IEEE 802.16.
- Uma nova arquitetura para a estação base em redes baseadas no padrão IEEE 802.16, com a provisão de QoS para aplicações com necessidades distintas, dentro das cinco classes de serviço especificadas pela norma, tanto de tempo real, quanto de não tempo real, utilizando a metaheurística de enxame de partículas para o cálculo da duração do tempo de um quadro.
- Validação da arquitetura através de simulação computacional e da análise de eficiência do modelo desenvolvido, comparando com trabalhos relevantes da área.

## 1.6 PUBLICAÇÕES

- (Revista) Dosciatti, E.R.; Godoy Júnior, W.; Foronda, A., **TQ/PSO - A New Scheduler to Optimize the Time Frame with PSO in WiMAX Networks**. Latin America Transactions, IEEE (Revista IEEE America Latina), vol.13, no.1, pp.365-376, Jan. 2015, DOI: 10.1109/TLA.2015.7040671. ISSN: 1548-0992. Disponível em: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7040671>

- (Conferência) Dosciatti, E.R.; Foronda, A., **Time Frame Optimization using PSO to Guarantee QoS in IEEE 802.16 Networks**. SAI2015 (Science and Information Conference 2015), Jul. 28-30, 2015, London, UK. Página do evento: <http://thesai.org/SAIConference2015>.

## 1.7 ORGANIZAÇÃO DA TESE

Além deste capítulo, o presente trabalho está organizado conforme a estrutura descrita a seguir. O Capítulo 2 é dedicado a fundamentação teórica na área de otimização, onde está descrito o método de otimização, a definição de inteligência de enxame e uma descrição detalhada do PSO, com seu modelo matemático e como está definida a utilização do PSO na otimização do  $TQ$ .

No Capítulo 3 é descrita a visão geral do padrão IEEE 802.16 com sua arquitetura e suas camadas física e de acesso ao meio, além da especificação da técnica de modulação utilizada e a estrutura do quadro utilizada pelo padrão IEEE 802.16. Também é descrito o modelo de provisão de QoS para o padrão IEEE 802.16 com as classes de serviço utilizadas pela norma. Por fim, descreve uma abordagem sobre o mecanismo de escalonamento em redes WiMAX, as disciplinas existentes e a classificação dos algoritmos de escalonamento.

O Capítulo 4 apresenta uma nova arquitetura para provisão de QoS utilizando enxame de partículas em redes WiMAX fixas. Esta nova arquitetura para a BS é baseada na modulação OFDM, utilizando a topologia PMP (*Point to Multipoint*) e a duplexação por divisão de tempo (*Time Division Duplexing* - TDD). Ainda, a modelagem matemática da nova arquitetura é demonstrada, bem como suas características e como será calculado o tempo de cada segmento no quadro.

No Capítulo 5, é demonstrado que, com a utilização do PSO para otimização do cálculo do  $TQ$ , ocorre um ganho no tempo de execução. Também, resultados da implementação da arquitetura, com a utilização do simulador de redes NS-2 (*Network Simulator*) (NS-2, 2015), são demonstrados e comparados com os trabalhos relacionados no Capítulo 3.

No Capítulo 6, a conclusão do trabalho desenvolvido é relatada e os trabalhos futuros são relacionados.

## 2 OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS - PSO

A primeira parte deste trabalho de pesquisa fornece uma revisão de algumas definições básicas relacionadas à questão de otimização. Uma breve discussão de métodos de otimização tradicionais e estocásticos é realizada e, dentro dos métodos estocásticos, os algoritmos evolutivos são abordados por ser o foco do desenvolvimento deste trabalho de pesquisa. Com os fundamentos desta abordagem, os conceitos de inteligência de enxame (*Swarm Intelligence* - SI) são apresentados pois são utilizados, através da aplicação da metaheurística PSO, para calcular a duração do *TQ* Ideal, possibilitando a maximização da alocação de usuários dentro deste quadro.

### 2.1 OTIMIZAÇÃO E MÉTODOS DE OTIMIZAÇÃO

Otimização é um termo científico que trata da detecção de soluções ótimas para um determinado problema, entre diversas alternativas. Desta forma, o objetivo da otimização é buscar valores para um conjunto de parâmetros para maximizar ou minimizar uma determinada função objetivo, sujeita a certas restrições que são impostas pelo problema. As restrições farão com que ocorra uma redução no número de soluções possíveis. Se uma solução satisfaz todas as restrições, ela é chamada de solução viável. Dentre todas as soluções viáveis, o problema de otimização refere-se à detecção de uma ou mais soluções ótimas. Desta forma, vários problemas ou processos podem ser formulados através de uma função e solucionados como um problema de otimização (KENNEDY et al., 2001).

Segundo Spall (2003), a otimização de sistemas tem sido uma área de pesquisa muito ativa nas últimas décadas. O crescimento científico e tecnológico dos últimos anos tem solucionado uma infinidade de difíceis problemas de otimização, para isso, algoritmos mais eficientes foram desenvolvidos. Os problemas com uma maior complexidade para se chegar a uma solução de otimização, são os relacionados ao mundo real. A solução destes problemas enfrenta muitas dificuldades devido as seguintes razões:

- Presença de ruído na avaliação da solução;

- Dificuldades em distinguir uma solução globalmente ótima de soluções localmente ótimas;
- A “maldição da dimensionalidade”, ou seja, o crescimento exponencial do espaço de busca em relação a dimensão do problema;
- Dificuldades associadas com as restrições e as necessidades de especificar métodos para o problema;
- A falta de estacionaridade na solução, devido as mudanças de condições do problema, que ocorrem ao longo do tempo.

Em virtude disso e, devido a natureza e as diferentes características matemáticas dos problemas de otimização, existe uma necessidade de especialização de algoritmos para solução de problemas específicos que compartilham propriedades comuns, tais como: a não-linearidade, a convexidade, a diferenciabilidade, a continuidade e a precisão de avaliação (PARSOPOULOS; VRAHATIS, 2010). Além disso, as características intrínsecas de cada algoritmo podem torná-los mais adequados, tanto para problemas de otimização locais, como para problemas de otimização globais. Tais características incluem, entre outros, estocasticidade, potencial para paralelização em sistemas computacionais modernos e limitados requisitos computacionais.

A otimização pode ser aplicada em várias áreas, como, por exemplo, no projeto de sistemas ou componentes, planejamento e análise de operações, controle de sistemas dinâmicos, dentre outros. Como vantagens no uso de procedimentos de otimização podem ser citados os seguintes aspectos (TELES; GOMES, 2010):

- Diminuição do tempo dedicado ao projeto;
- Possibilidade de tratamento simultâneo de uma grande quantidade de variáveis e restrições, que podem ser de difícil visualização gráfica e/ou tabular;
- Possibilidade de obtenção de um resultado melhor, quando comparado ao resultado anterior para um mesmo problema;
- Obtenção de soluções não tradicionais;
- Custo computacional menor.

Avanços significativos na área de otimização tem sido alcançados pelos pesquisadores. As técnicas clássicas de otimização são confiáveis, mas podem apresentar algumas dificuldades

numéricas e problemas de robustez relacionados com a falta de continuidade das funções a serem otimizadas ou de suas restrições, o uso de funções não convexas, a necessidade de se trabalhar com valores discretos das variáveis e a existência de mínimos ou máximos locais (GOMES, 2009).

As soluções para estes problemas estão nas técnicas heurísticas de otimização, que proporcionam soluções boas, mas não garantem que as soluções encontradas sejam as ótimas. Assim, métodos baseados na busca aleatória controlada por critérios probabilísticos, tiveram um importante desenvolvimento nos últimos anos, principalmente devido ao avanço dos recursos computacionais, uma vez que esses métodos necessitam de um número elevado de avaliações da função objetivo. Dentre esses métodos, destaca-se a técnica PSO, desenvolvida por Kennedy e Eberhart (1995), que é utilizada neste trabalho e detalhada na Seção 2.3. O PSO foi escolhido pois o problema a ser resolvido é uma otimização não-linear.

PSO é uma técnica baseada no comportamento sócio biológico associado com grupos de pássaros. Fazendo uma analogia, cada pássaro seria correspondente a uma partícula (ou indivíduo) da população, a “área sobrevoada” é equivalente ao espaço de busca e o ato de encontrar o alimento ou o ninho corresponde a encontrar o resultado ótimo. Ao observar o comportamento do grupo, constatou-se que o bando encontra alimento por meio do esforço conjunto e não individual, o que pode ser inferido que cada informação é compartilhada dentro do grupo, e assim, o comportamento é influenciado pela experiência acumulada por cada pássaro (comportamento cognitivo), bem como pelo resultado da experiência acumulada pelo grupo (comportamento social) (GOMES, 2009).

### 2.1.1 ALGORITMOS DE OTIMIZAÇÃO

Os algoritmos de otimização buscam melhorar a performance em direção a algum ponto (ou pontos) ótimo(s). Esta definição pode ser dividida em: (i) localização do “ponto ótimo”, que se utiliza de métodos tradicionais que são exatos para encontrar soluções ótimas. A ideia é que, se um problema pode ser resolvido, então o algoritmo deve encontrar a melhor solução global. Um método exato se utiliza de pesquisa do tipo “força bruta” (ou exaustiva), que é um método de pesquisa em que o algoritmo tenta cada solução no espaço de busca, de modo que a solução global ótima tem a garantia de ser encontrada. Obviamente, quando o espaço de busca aumenta, o custo do algoritmo força bruta também aumenta (MICHALEWICZ; FOGEL, 2013); e, (ii) melhoramento da performance de alguma abordagem, que são os algoritmos de otimização estocásticos e são utilizados para encontrar soluções perto de ótimas, em tempo polinomial. Isto é conseguido ao assumir que soluções boas estão próximas umas das

outras no espaço de busca. Este pressuposto é válido para a maioria dos problemas do mundo real (SPALL, 2003; LOVBERG, 2002).

Um algoritmo de otimização estocástico tem o objetivo de encontrar uma solução quase ótima, eles podem não conseguir encontrar uma solução global ótima. Enquanto um algoritmo exato gera uma única solução após a execução ser completada, um algoritmo estocástico pode ser interrompido a qualquer momento durante a execução e, então, fornecer a melhor solução encontrada até este momento. (LOVBERG, 2002).

Algoritmos estocásticos de busca têm algumas vantagens (VENTER; SOBIESZCZANSKI-SOBIESKI, 2002):

- São fáceis de implementar;
- Podem ser utilizados de forma eficiente em um ambiente de múltiplos processadores;
- Não requerem que a função de definição do problema seja contínua;
- Na grande maioria dos casos encontram soluções ótimas ou quase ótimas;
- São apropriados para problemas discretos e combinatoriais.

Os três principais algoritmos estocásticos são: Subida da Montanha (*Hill-Climbing*), que pode ser encontrado em (MICHALEWICZ; FOGEL, 2013), Recozimento Simulado (*Simulated Annealing*), em (LAARHOVEN; AARTS, 1987) e Busca Tabu (*Tabu Search*), em (GLOVER, 1989) e (GLOVER, 1990).

Dentro da classe de algoritmos estocásticos, os Algoritmos Evolutivos (AEs) são métodos estocásticos de busca direta baseados na teoria da evolução. Os AEs diferem dos métodos de otimização, como o Subida de Montanha e Recozimento Simulado, pelo fato de que os AEs mantêm uma população (ou candidatos) de potenciais soluções para um problema, e não apenas uma solução (ENGELBRECHT, 2002). Nesta ideia, os AEs manipulam a população de soluções de tal forma que as piores soluções desapareçam e as melhores evoluem em busca da solução ótima, ou seja, o conceito principal é a sobrevivência do mais forte. Assim, podemos definir um AE como sendo um procedimento iterativo de busca (otimização) inspirado nos mecanismos evolutivos biológicos. A computação evolutiva se configura como a área de pesquisa, dentro da linha maior de inteligência computacional, que trata dos AEs.

A evolução admite uma interpretação na forma de um processo de otimização, que procura por melhores soluções dentro de um espaço de busca. Sendo assim, este processo procura identificar regiões mais promissoras, ou seja, regiões que contêm indivíduos com

melhores índices de aptidão. Sob uma visão matemática, a superfície de aptidão pode ser a função que se deseja maximizar ou minimizar. Assim, soluções ótimas locais se encontram em extremos desta superfície (GONCALVES, 2011).

Um AE padrão é composto por quatro elementos básicos:

1. Uma população de indivíduos, onde cada indivíduo corresponde a uma solução candidata;
2. Uma maneira de criar novos indivíduos, a partir de indivíduos já existentes, e inserir variabilidade nos descendentes;
3. Uma forma de medir a qualidade de cada indivíduo;
4. Um método de selecionar os melhores indivíduos, aplicando assim o princípio da seleção natural.

Diferentes classes de AEs foram desenvolvidas (ENGELBRECHT, 2002):

- **Algoritmos Genéticos:** é uma técnica de resolução de problemas com a interessante propriedade de buscar a solução de forma similar a teoria da evolução. São algoritmos de pesquisa baseados no mecanismo de genética e na seleção natural, utilizando operadores que emulam estes princípios para guiá-los até que a solução seja encontrada;
- **Programação Genética:** é uma técnica baseada em algoritmos genéticos, mas os indivíduos são programas e são representados como árvores;
- **Programação Evolutiva:** é uma técnica derivada a partir da simulação do comportamento adaptativo na evolução;
- **Estratégias Evolutivas:** são voltadas para a modelagem dos parâmetros estratégicos que controlam a variação da evolução, ou seja, a evolução da evolução;
- **Evolução Diferencial:** é uma técnica semelhante aos algoritmos genéticos, diferindo no mecanismo de reprodução utilizado;
- **Evolução Cultural:** modela a evolução da cultura de uma população e, de como a cultura influencia a evolução genética dos indivíduos;
- **Coevolução:** permite que indivíduos, inicialmente “burros”, evoluam através da cooperação, ou da competição com outros indivíduos, adquirindo as características necessárias para sobreviver.

Os AEs tem sido utilizados com sucesso em aplicações do mundo real, por exemplo, mineração de dados, otimização combinatória, diagnóstico de falhas, processamento de imagens, reconhecimento de padrões, escalonamento de tarefas, projeto de antenas, projeto de redes, aproximação de séries temporais, entre outros.

Neste trabalho de pesquisa, o foco é a otimização por enxame de partículas, um método de otimização, que faz parte dos AEs, e que será utilizado e descrito a seguir, antes, porém, uma introdução a inteligência de enxame é necessária.

## 2.2 INTELIGÊNCIA DE ENXAME

A inteligência de enxame, conhecida como SI (*Swarm Intelligence*), é o termo utilizado para designar sistemas baseados no comportamento coletivo dos indivíduos em uma população, sendo que estes indivíduos possuem a capacidade de perceber e modificar o seu ambiente de maneira local. O termo SI foi proposto no fim da década de 1980, quando se referia a sistemas robóticos compostos por uma coleção de agentes simples em um ambiente, interagindo de acordo com regras locais (ZUBEN; ATTUX, 2008). Esses sistemas podem ser auto-organizáveis, autônomos, flexíveis, dinâmicos e distribuídos. A SI pode ser natural ou artificial. Exemplos naturais de SI são colônia de formigas, cardume de peixes, bando de pássaros, enxame de abelhas, e assim por diante. Programas de computador para solução de problemas de otimização, para análise de dados e para estratégias de coordenação e controle de robôs são exemplos de SI artificiais.

As técnicas de SI mais bem sucedidas são o PSO e a otimização por colônia de formigas (do Inglês, *Ant Colony Optimization - ACO*) (DORIGO; BIRATTARI, 2007). No PSO, cada partícula voa através do espaço multidimensional e ajusta a sua posição, a cada passo, com a sua própria experiência e a de seus pares, em direção a uma solução ótima.

O termo “enxame” (ou população) é utilizado de forma genérica para se referir a qualquer coleção estruturada de indivíduos, que também podem ser chamados de agentes autônomos, capazes de interagir localmente entre si, regidos por um comportamento global, buscando a solução para problemas de forma distribuída (MILLONAS, 1994). Desta forma, um agente autônomo é um subsistema que interage com o ambiente, e que provavelmente consiste de outros agentes autônomos, mas age de forma relativamente independente de todos os outros agentes autônomos, não seguindo comandos de um líder, ou algum plano global (LIU; PASSINO, 2000). Apesar de cada agente autônomo possuir um espaço de ação muito limitado e não ter nenhum controle central, o comportamento global do enxame apresenta traços de



inteligência, ou seja, uma habilidade para reagir às mudanças ambientais e uma capacidade de tomada de decisões (PARSOPOULOS; VRAHATIS, 2010). Por exemplo, para um pássaro participar de um bando, é só ajustar seus movimentos para coordenar com os movimentos dos outros pássaros do bando, normalmente os “vizinhos” que estão próximos a ele no bando.

De acordo com Millonas (1994), os modelos para aplicação na vida artificial são desenvolvidos a partir de cinco princípios básicos, que podem ser considerados como princípios evolutivos de seleção, porém, não pretendem ser definitivos:

1. **Princípio da proximidade:** a população deve ser capaz de responder, através do seu comportamento, às variações dos estímulos ambientais de espaço e tempo. Neste caso, o comportamento pode ser entendido como uma resposta direta aos estímulos ambientais que, em algum momento, maximiza a utilidade para a população como um todo, de algum tipo de atividade. Embora os tipos de atividades possam ter uma maior variação, dependendo tanto do tipo e complexidade dos organismos, algumas atividades típicas incluem a busca e recuperação de alimentos, a construção de ninhos, a defesa da população, o movimento coletivo, e, no caso de organismos mais desenvolvidos, as interações necessárias para muitas funções sociais, ou seja, cada agente autônomo deve possuir a capacidade de interagir com toda a população;
2. **Princípio da qualidade:** a população deve ser capaz de responder aos fatores de qualidade do ambiente, que incluem a qualidade dos alimentos e a segurança local, desta forma, os agentes devem ser capazes de avaliar seus comportamentos;
3. **Princípio da diversidade:** a população não deve alocar seus recursos em áreas muito restritivas. Deve distribuir os recursos, prevenindo-se de perturbações do ambiente, com isso, os agentes devem permitir que o sistema possa reagir a situações inesperadas;
4. **Princípio da estabilidade:** a população não deve mudar seu modo de comportamento a cada mudança que ocorrer no ambiente. Isso significa que nem todas as variações ambientais devem afetar o comportamento de um agente;
5. **Princípio da adaptabilidade:** a população deve mudar o seu modo de comportamento quando, essa mudança, lhe proporcione benefícios, ou seja, ter a capacidade de adequação a variações ambientais.

Assim, pode-se resumir que a inteligência de enxame é uma propriedade emergente de um sistema coletivo que resulta de seus princípios de proximidade, qualidade, diversidade, estabilidade e adaptabilidade (ZUBEN; ATTUX, 2008).

A Seção 2.3 descreve o comportamento do algoritmo de otimização por enxame de partículas, além de suas principais características e terminologias.

### 2.3 OTIMIZAÇÃO POR ENXAME DE PARTÍCULAS

Os princípios fundamentais do PSO, apresentados abaixo, são importantes propriedades do comportamento social humano, que utiliza esses conceitos na adaptação a mudanças no ambiente e na resolução de problemas complexos (KENNEDY et al., 2001):

- **Avaliação dos estímulos.** Caracterizada como sendo tanto, positivo, negativo, de atração ou repulsão. A avaliação possivelmente representa a característica comportamental mais evidente dos seres vivos. A aprendizagem se torna possível quando o organismo é capaz de avaliar e distinguir características do ambiente. Portanto, a aprendizagem pode ser definida como a capacidade do organismo de melhorar a sua avaliação média do meio ambiente.
- **Comparação entre os indivíduos.** Inspirado por modelos de adaptações culturais e teorias sócio-psicológicas, o modelo adaptativo cultural adota como referência o comportamento de um indivíduo em relação aos outros indivíduos, ou seja, cada indivíduo se compara aos seus vizinhos e o imita apenas naquilo que julgar como sendo superior.
- **Imitação.** Como uma atividade relativamente restrita na natureza, a imitação é encontrada, de fato, somente entre algumas espécies animais, incluindo o ser humano, pelo fato de representar uma forma muito eficiente de aprendizado.

No PSO, a população, denominada de enxame, é composta por partículas. Esse enxame de partículas realiza buscas dentro de um espaço de busca multidimensional. Essas buscas são soluções candidatas, de cada partícula, para um determinado problema. Neste espaço de busca, em um determinado intervalo de tempo, cada partícula ocupa uma posição específica e movimenta-se com uma velocidade, a qual é modificada baseando-se na experiência adquirida pela própria partícula e na experiência das outras partículas do enxame. Fazendo uma analogia ao bando de pássaros, cada partícula age como se fosse uma ave do bando em busca de comida, utilizando as informações das melhores posições no espaço de busca encontradas por ela e pelas demais aves do bando.

A experiência de cada partícula inclui informações relevantes da sua trajetória no espaço de busca, de forma a armazenar a melhor posição já ocupada pela própria partícula, que corresponde à memorização do conhecimento associado às boas soluções obtidas durante

o processo iterativo. A posição e a velocidade de cada partícula são ajustadas a cada iteração, de acordo com a melhor posição já ocupada pela própria partícula e com a melhor posição já ocupada por uma partícula, dentre todas as partículas do enxame (ALRASHIDI; EL-HAWARY, 2009). O segredo do sucesso do PSO está, justamente, no procedimento da troca de experiências em que, a experiência de cada partícula, é continuamente comunicada a uma parte ou a todo enxame, fazendo, com isso, que todo o enxame se movimente para as áreas mais promissoras detectadas até agora no espaço de busca (PARSOPOULOS; VRAHATIS, 2010).

Seguindo a linha de raciocínio de (KENNEDY; EBERHART, 1995), cada indivíduo pode ser a solução para o problema que está sendo investigado, sendo atribuído para cada indivíduo um valor que está relacionado a adequação da partícula com a solução do problema, que é denominado de função de aptidão (*fitness*), e também uma variável velocidade que representa a movimentação do indivíduo (partícula). Com o desenvolvimento da solução, os indivíduos vão ajustando suas velocidades em relação à melhor solução (melhor aptidão) encontrada por ele próprio e também pela melhor solução do grupo (enxame), realizando este processo até que as partículas encontrem a melhor solução global ou próxima dela. O valor da função de aptidão é calculado com base na função objetivo que se deseja otimizar (maximizar ou minimizar). Assim, o valor da função de aptidão é definido pela natureza do problema de otimização e é computada por uma função objetivo que avalia a solução.

### 2.3.1 MODELO MATEMÁTICO DO PSO

No algoritmo PSO, cada partícula, tratada como um ponto no espaço de busca, representa uma potencial solução para o problema e segue as melhores posições encontradas até o momento pelas partículas do bando. Assim, na inicialização do algoritmo, as partículas são aleatoriamente inicializadas (vetores de posição e de velocidade), sendo que cada partícula  $i$  é representada por três vetores, sendo  $N$  a  $i$ -ésima posição do vetor:

- A posição atual da partícula num espaço de busca  $D$ -dimensional:  $X_i = (x_{i1}, x_{i2}, \dots, x_{iN})$ .
- A melhor posição que a partícula encontrou:  $pbest_i = (pbest_{i1}, pbest_{i2}, \dots, pbest_{iN})$ .
- A velocidade atual da partícula:  $V_i = (v_{i1}, v_{i2}, \dots, v_{iN})$ .

Uma característica fundamental da abordagem PSO é a sua simplicidade, pois o modelo é composto por apenas duas equações. Nestas equações, as partículas se movem no espaço de busca à procura da melhor solução possível. A cada iteração  $k$ , a velocidade é

atualizada, conforme a Equação (1).

$$v_i^{(k+1)} = (w * v_i^{(k)}) + (c_1 * r_1 * [pbest_i^{(k)} - x_i^{(k)}]) + (c_2 * r_2 * [gbest^{(k)} - x_i^{(k)}]) \quad (1)$$

Os componentes que fazem parte da Equação (1) estão definidos a seguir:

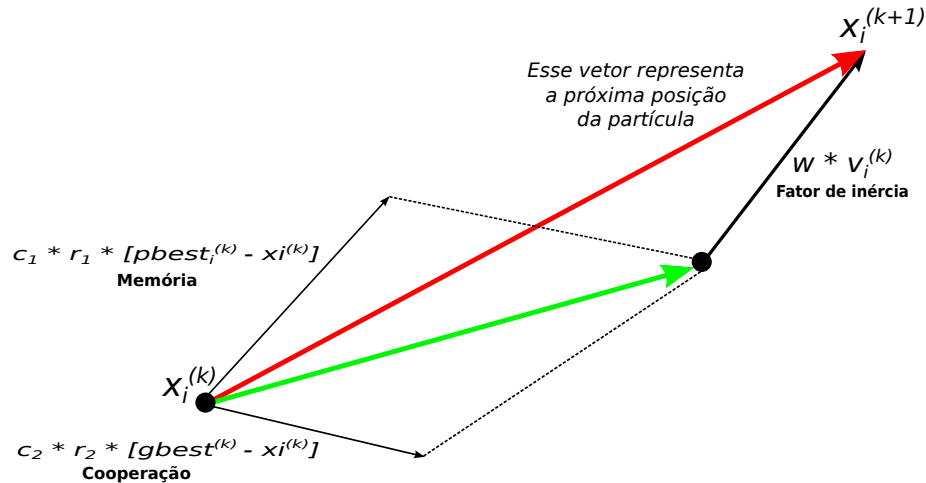
- $i$  é uma partícula que representa uma solução candidata;
- $k$  representa o contador de iterações;
- $v_i^k$  representa a velocidade atual da partícula  $i$ . A atualização da velocidade de cada partícula depende de parâmetros que devem ser ajustados a cada problema a ser otimizado. A velocidade de uma partícula, em cada dimensão  $D$ , é limitada a um valor máximo de velocidade, denominado de  $V_{max}$  e tem a função de auxiliar para que o sistema não extrapole o espaço de busca, pois determina a resolução que a região próxima às soluções atuais são procuradas. Um valor alto para  $V_{max}$ , faz com que o PSO tenha facilidade na busca global, enquanto que um valor pequeno de  $V_{max}$  facilita as buscas locais. Assim, se  $(v_i > V_{max})$ , então  $(v_i = V_{max})$ , senão, se  $(v_i < -V_{max})$ , então  $(v_i = -V_{max})$ .
- $x_i^k$  representa a posição atual da partícula  $i$ ;
- $c_1$  e  $c_2$  representam os parâmetros de confiança, também chamados de parâmetros **cognitivo** ( $c_1$ ), que é o quanto uma partícula confia em si, também chamado de **memória**, e **social** ( $c_2$ ), que é o quanto uma partícula confia no enxame, também chamado de **cooperação**. Esses parâmetros controlam a importância relativa da experiência particular de cada partícula versus a experiência social do enxame, controlando o movimento de cada partícula para a sua melhor posição individual versus a sua melhor posição global (VALLE et al., 2008). Como a velocidade modela a taxa de mudança dentro da posição da partícula, as mudanças induzidas pela Equação (1), representam a aceleração da partícula. Por este fato, as constantes  $c_1$  e  $c_2$ , também são chamadas de coeficientes de aceleração, e exercem influência no tamanho máximo do passo que uma partícula pode dar em uma única iteração. Segundo Ratnaweera et al. (2004), a redução do componente cognitivo ( $c_1$ ) e o aumento do componente social ( $c_2$ ), durante o início da otimização pelo PSO, pode ser uma alternativa viável, pois um valor maior do componente cognitivo e um valor menor do componente social, incentivam as partículas a moverem-se pelo espaço de busca em vez de moverem-se em direção à melhor partícula da população. Em contra partida, um menor valor de  $c_1$  e um maior valor de  $c_2$ , ao final da otimização, permitem que as partículas façam a convergência para um local ótimo;

- $w$  representa o fator de inércia ou peso inercial, que é um parâmetro utilizado para controlar a exploração e utilização do espaço de busca, ou seja, utilizado para controlar o impacto das velocidades anteriores sobre a velocidade atual.  $w$  tem a influência sobre o equilíbrio entre as habilidades de exploração global e local das partículas, sendo que um alto valor de  $w$  facilita um comportamento mais global, enquanto um valor de  $w$  baixo facilita um comportamento mais local. Assim, segundo Lee e Park (2006), para controlar o equilíbrio entre a exploração global e local, obtendo-se uma convergência rápida, e para chegar a uma solução ideal, o peso inercial, cujo valor decresce linearmente com o número de iterações deve ser:  $(w = w_{max} - (\frac{w_{max} - w_{min}}{k_{max}}) * k)$ , sendo que,  $(w_{max} > w_{min})$ , onde  $w$  é o peso inercial;  $w_{max}$  e  $w_{min}$  são os pesos de inércia máximo e mínimo, respectivamente,  $k_{max}$  é o número máximo de iterações; e  $k$  é o número da iteração atual;
- $r_1$  e  $r_2$  representam dois vetores com números aleatórios uniformemente distribuídos no intervalo  $[0, 1]$ .  $r_1$  é a variável aleatória da parte cognitiva e  $r_2$  é a variável aleatória da parte social. Como o PSO opera, em cada dimensão, de forma independente, cada dimensão tem o seu próprio número aleatório  $r$  (PARSOPOULOS; VRAHATIS, 2010);
- $pbest_i^k$  (*personal best*) representa a melhor posição já alcançada pela partícula  $i$  com base na sua própria experiência:  $pbest_i^k = [x_{i1}^{pbest}, x_{i2}^{pbest}, \dots, x_{iN}^{pbest}]$ . Também pode ser visto como a memória de cada partícula. Como a partícula se move através do espaço de busca, ela compara o valor da sua função de avaliação na posição atual com o melhor valor da função de aptidão atingido em qualquer momento até o momento atual;
- $gbest^k$  (*global best*) é a melhor posição das partículas encontrada pelo enxame, com base nas melhores posições de  $pbest$  obtidas até o momento atual.  $gbest$  utiliza a experiência global do enxame:  $gbest^k = [x_1^{gbest}, x_2^{gbest}, \dots, x_N^{gbest}]$ ;
- $s$  é o tamanho do enxame ou, simplesmente, o número de partículas no enxame. Quanto mais partículas no enxame, maior a diversidade inicial do enxame. O PSO tem a capacidade de encontrar as melhores soluções com enxame entre 10 a 30 partículas (ENGELBRECHT, 2002), considerados pequenos. O sucesso tem sido mesmo obtido com enxames com menos de 10 partículas (BERGH, 2002). Porém, o tamanho ideal do enxame é dependente do problema que está sendo modelado. Um conjunto predeterminado de partículas no enxame é utilizado, de acordo com o número de dimensões ( $D$ ) da posição, ou seja, o número de atributos de condição do problema a ser resolvido, que neste trabalho é de duas dimensões e definido por:  $(s = 10 + [2 * \sqrt{D}])$ . O tamanho do enxame afetará a velocidade de convergência do algoritmo.

A nova posição da partícula é determinada pela soma da sua posição atual e a nova velocidade, de acordo com a Equação (2).

$$x_i^{(k+1)} = x_i^{(k)} + v_i^{(k+1)} \quad (2)$$

O gráfico da Figura 1, mostra uma interpretação geométrica da movimentação de uma partícula dentro do espaço de busca.



**Figura 1: Interpretação geométrica do PSO**

Fonte: Autoria própria.

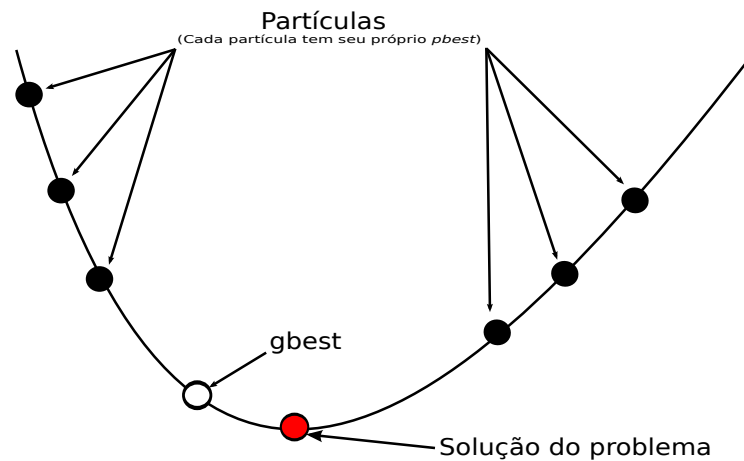
A memória dos movimentos exploratórios realizados pelas partículas, a cooperação entre as partículas e o fator de inércia relacionado a partícula, são armazenados por sua posição atual. Essas posições podem ser vistas como vetores, onde a soma vetorial resultante determina a nova posição de uma partícula  $x_i^{(k)}$ .

A definição dos parâmetros  $w$ ,  $c_1$  e  $c_2$  é de fundamental importância para a solução de um problema de otimização. A literatura propõe usar  $0 < w < 1.4$ ,  $c_1 = c_2 = 2$  com  $c_1 + c_2 \leq 4$ , para manter um equilíbrio entre a capacidade de busca global e local do algoritmo (PARSOPOULOS; VRAHATIS, 2010; ALRASHIDI; EL-HAWARY, 2009). Porém, usualmente, estes parâmetros são ajustados por heurísticas de tentativa e erro para encontrar o melhor conjunto de parâmetros para o algoritmo PSO, sempre dependente do problema a ser solucionado (VENTER; SOBIESZCZANSKI-SOBIESKI, 2002). Neste trabalho, que se utiliza da aplicação do PSO em redes WiMAX, os parâmetros utilizados estão definidos na Tabela 8 no Capítulo 4.

O algoritmo PSO clássico define que deve existir uma condição de parada para o processo. Essa condição de parada depende do tipo de problema a ser resolvido. Normalmente é determinado que o algoritmo seja executado um número fixo e predeterminado de iterações,

ou até alcançar um valor específico de erro, por exemplo, até que o desempenho do enxame não tenha melhora significativa.

Para cada iteração  $k$  e para cada partícula  $i$  do enxame, o algoritmo deve avaliar as Equações (1) e (2) e atualizar o vetor de posição da partícula  $i$  com os valores encontrados, calcular a função de aptidão da partícula  $i$  e atualizar os valores de  $pbest$  e  $gbest$ . Assim, o conceito do algoritmo PSO consiste de, a cada passo iterativo  $k$ , mudar a velocidade de cada partícula  $i$  em direção às localizações do  $pbest$  e do  $gbest$  (SHA; HSU, 2006). A partícula se move de acordo com a sua velocidade. As velocidades iniciais de cada partícula são geradas aleatoriamente para as posições  $pbest$  e  $gbest$ . A Figura 2 mostra a aproximação de  $pbest$  e  $gbest$  em direção à solução do problema.



**Figura 2: Busca da melhor solução pelo  $pbest$  e  $gbest$**   
**Fonte: Autoria própria.**

Em um resumo breve do funcionamento do PSO, pode-se dizer que cada partícula utiliza dois importantes tipos de informação no processo de decisão. A primeira informação é a própria experiência da partícula, ou seja, ela busca utilizar de suas próprias escolhas e sabe quais dessas foram melhores até a geração atual. Pode-se dizer que essa informação corresponde à aprendizagem individual (cognitiva). A segunda é baseada na experiência de outras partículas, ou seja, cada partícula tem conhecimento de como as outras ao seu redor se comportam. Elas sabem quais as escolhas das partículas, que estão em sua vizinhança, são as mais positivas até agora. Desta forma, esta segunda informação corresponde à transmissão cultural (social). Assim, no PSO, cada partícula faz a sua decisão de acordo com as suas próprias experiências e as experiências do enxame.

### 2.3.2 ALGORITMO PSO

No algoritmo PSO, a população tem partículas que representam soluções candidatas. O Algoritmo 1 apresenta um pseudocódigo do procedimento para a implementação do algoritmo básico PSO.

#### Algoritmo 1: Pseudocódigo do algoritmo PSO

Fonte: Autoria própria.

---

```

1: Atribuir os parâmetros para cada partícula do enxame de forma aleatória
2: para cada partícula  $i = 1$  até  $s$  faça ( $s =$  tamanho do enxame)
3:   Iniciar  $x_i$  com uma solução aleatória para o problema
4:   Iniciar  $v_i$  com uma velocidade aleatória  $< V_{max}$  ( $V_{max} =$  valor máximo da velocidade)
5:    $pbest_i \leftarrow gbest \leftarrow x_i$ 
6: fim para cada
7: enquanto não atingir a condição de parada faça (condição de parada = tamanho do enxame)
8:   para cada partícula  $i = 1$  até  $s$  faça
9:     se  $x_i$  for melhor que  $pbest_i$  então
10:       $pbest_i = x_i$ 
11:     fim se
12:     se  $pbest_i$  for melhor que  $gbest$  então
13:       $gbest = pbest_i$ 
14:     fim se
15:   fim para cada
16:   Atualizar a velocidade  $v_i$  da partícula  $i$ 
17:   Atualizar a posição  $x_i$  da partícula  $i$ 
18: fim enquanto
19: Retornar melhor  $gbest$ 

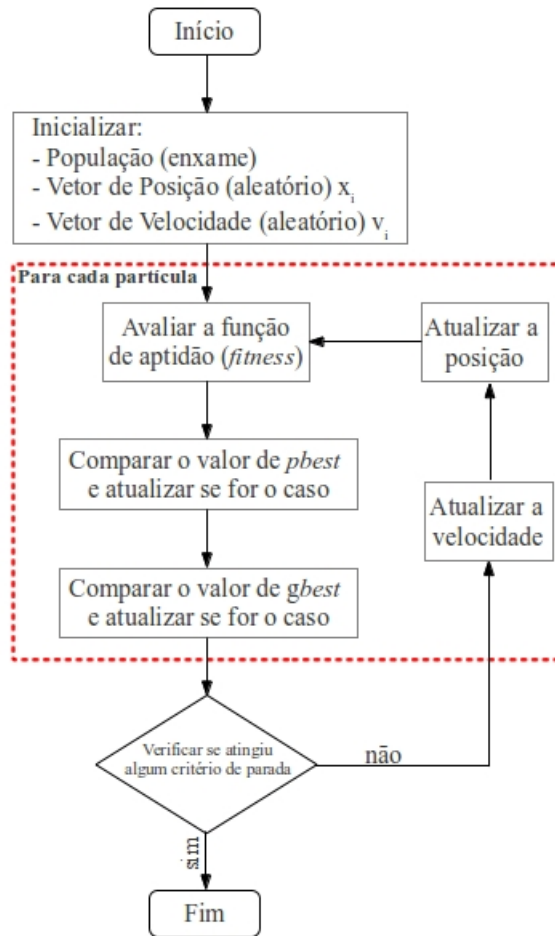
```

---

Neste algoritmo, os passos 2 a 6 inicializam aleatoriamente os vetores de posição  $x_i$  e velocidade  $v_i$ , sendo que a velocidade  $v_i$  é limitada por um valor máximo de velocidade, denominado por  $V_{max}$ , a fim de evitar que o sistema extrapole o espaço de busca, também inicializa os valores de  $pbest_i$ , que representa a melhor posição já alcançada pela partícula  $i$  com base na sua própria experiência e de  $gbest$ , que é a melhor posição das partículas encontrada pelo enxame, com base nas melhores posições de  $pbest$  obtidas até o momento atual, ou seja, utiliza a experiência global do enxame. Os passos 7 a 18 são executados até que um critério de parada seja encontrado, que neste caso será o tamanho do enxame ( $s$ ). Nos passos 9 a 11, é realizada a comparação da avaliação da função de aptidão da partícula  $i$  com o  $pbest$  desta partícula. Se o valor corrente é melhor que  $pbest$ , então o valor de  $pbest$  passa a ser igual ao valor da função de aptidão da partícula  $i$ , e a localização do  $pbest$  passa ser igual a localização atual no espaço  $D$ -dimensional. Nos passos 12 a 14, é realizada a comparação do valor obtido em  $pbest$  com o melhor valor global  $gbest$ . Se o valor de  $pbest$  for melhor, atualizar  $gbest$  com o novo valor. Os passos 16 e 17 atualizam os vetores de velocidade  $v_i$  e a posição  $x_i$ , respectivamente. O passo 19 retorna o valor de  $gbest$ .



Na Figura 3, os passos de um algoritmo PSO básico podem ser melhor visualizados.

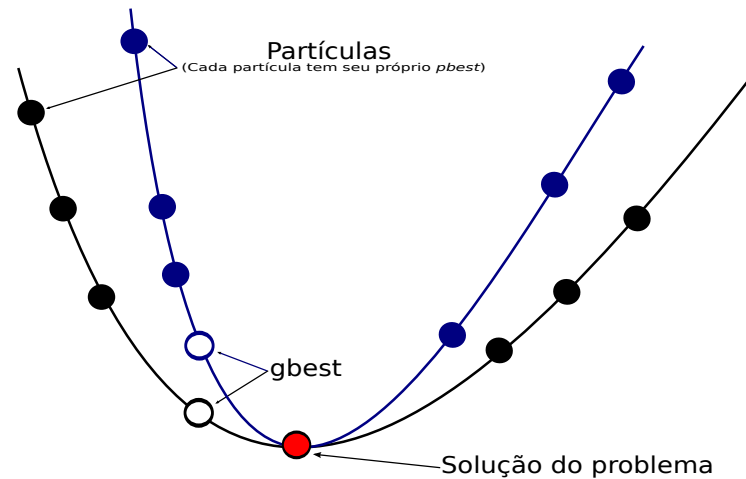


**Figura 3: Passos de um algoritmo PSO**  
**Fonte: Autoria própria.**

### 2.3.3 ALGORITMO PSO COM RESTRIÇÃO

Os problemas encontrados no mundo real possuem tipicamente propriedades de: (i) múltiplas metas; (ii) funções objetivo com ruídos e variantes no tempo; (iii) dados mal-estruturados; e (iv) restrições complexas. O tratamento de problemas com restrições afeta o espaço de busca, pois as restrições dividem o espaço de busca em soluções viáveis, que são aquelas que atendem as restrições, e inviáveis, que não atendem as restrições. A Figura 4 mostra o caso que é desenvolvido neste trabalho, onde duas condições de restrição devem ser satisfeitas para a busca da solução do problema.

O algoritmo básico do PSO, por tratar-se de um algoritmo natural, não trabalha diretamente com restrições, ou seja, foi definido para problemas sem restrições. Quando há restrições nos problemas de otimização, as partículas que desrespeitam alguma restrição se



**Figura 4: Busca da melhor solução pelo  $pbest$  e  $gbest$  em problemas com restrições**  
**Fonte: Autoria própria.**

enquadram em um grupo especial que merece um tratamento diferenciado. Uma vez que a maioria dos problemas de engenharia são limitados de uma forma ou de outra, é importante adicionar a capacidade de tratar problemas de otimização com restrições.

Como a restrição reduz o espaço de busca viável onde a solução para o problema possa ser encontrada e, algoritmos de otimização precisam garantir que uma solução viável seja encontrada, então, um algoritmo de otimização deve encontrar uma solução que optimize tanto a função objetivo quanto satisfaça todas as restrições (RINI et al., 2011).

Segundo Hu e Eberhart (2002), os problemas não-lineares de otimização com restrição, são constituídos por três componentes básicos: (i) conjunto de variáveis; (ii) função de aptidão para ser otimizada (minimizar ou maximizar); e (iii) conjunto de restrições que especificam os espaços viáveis das variáveis e, podem ser definidos como se segue:

$$\begin{aligned} \text{Otimize} \quad & f(X), \quad X = (x_1, x_2, \dots, x_n) \in \mathfrak{R}^N \\ \text{Sujeito a} \quad & g_i(X) \leq 0 \quad \text{para } i = 1, 2, \dots, p \\ & h_i(X) = 0 \quad \text{para } i = 1, 2, \dots, m, \end{aligned}$$

onde  $f(X)$  é a função objetivo a ser minimizada ou maximizada;  $x_i$  e  $i = 1, 2, \dots, n$  são variáveis de um problema de dimensão  $n$ ;  $g_i(X)$  e  $h_i(X)$  são, respectivamente, as  $p$  e  $m$  restrições de desigualdade e igualdade em um problema de otimização com restrições.

Assim, o objetivo dos problemas não-lineares de otimização com restrição é encontrar os valores das variáveis que otimizem a função objetivo, desde que satisfaçam as restrições (HU; EBERHART, 2002). Outras abordagens de PSO relacionadas com problemas com restrições podem ser encontradas em (HU; EBERHART, 2002; SCHOOF; NAUDTS, 2002; COATH;

HALGAMUGE, 2003; GAING, 2003; PARSOPOULOS; VRAHATIS, 2010).

O algoritmo PSO original precisa ser modificado para tratar das restrições. Algumas ideias de manipulação de restrições podem ser adotadas. O mais simples é o método baseado na preservação da viabilidade das soluções. A fim de encontrar o melhor no espaço possível, cada partícula pesquisa o espaço inteiro, mas só mantém rastreamento nas soluções viáveis. Para acelerar este processo, todas as partículas são inicializados como soluções viáveis. O Algoritmo 2 mostra o pseudocódigo com restrições do algoritmo PSO.

### Algoritmo 2: Pseudocódigo do algoritmo PSO com restrições

Fonte: Autoria própria.

---

```

1: Atribuir os parâmetros para cada partícula do enxame de forma aleatória
2: para cada partícula  $i = 1$  até  $s$  faça    ( $s =$  tamanho do enxame)
3:   Iniciar  $x_i$  com uma solução aleatória para o problema
4:   Iniciar  $v_i$  com uma velocidade aleatória  $< V_{max}$     ( $V_{max} =$  valor máximo da velocidade)
5:    $pbest_i \leftarrow gbest \leftarrow x_i$ 
6: fim para cada
7: enquanto não atingir a condição de parada faça (condição de parada = tamanho do enxame)
8:   para cada partícula  $i = 1$  até  $s$  faça
9:     se  $x_i$  for melhor que  $pbest_i$  E a partícula está no espaço viável então
10:       $pbest_i = x_i$ 
11:     fim se
12:     se  $pbest_i$  for melhor que  $gbest$  então
13:        $gbest = pbest_i$ 
14:     fim se
15:   fim para cada
16:   Atualizar a velocidade  $v_i$  da partícula  $i$ 
17:   Atualizar a posição  $x_i$  da partícula  $i$ 
18: fim enquanto
19: Retornar melhor  $gbest$ 

```

---

No algoritmo PSO com restrições, todas as soluções aleatoriamente inicializadas precisam estar localizadas no espaço de busca viável, ou seja, ao calcular os valores  $pbest_i$  e  $gbest$ , apenas as posições no espaço viável são computadas.

## 2.4 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou uma abordagem sobre otimização e seus métodos, tendo como ênfase o algoritmo de otimização PSO e, dentro dele, a otimização com o uso de restrições. O PSO é uma técnica originalmente inspirada no movimento de revoadas de pássaros em busca por alimentos, apresenta uma série de características extremamente desejáveis em mecanismos de otimização, como baixas exigências computacionais, robustez, flexibilidade e rápida convergência. Mesmo sendo uma técnica recente, quando comparado a metaheurísticas

similares, o PSO conseguiu alcançar uma grande relevância, sendo seu uso explorado em uma grande variedade de problemas, da otimização combinatória à composição musical.

Segundo Parsopoulos e Vrahatis (2010), o PSO tem algumas vantagens em relação a outros métodos heurísticos, tais como a existência de poucos parâmetros para serem ajustados, pois é um algoritmo simples. Também, pelo fato de ser mais rápido e aplicável a vários tipos de problemas de otimização, onde as funções podem ter restrições, se torna um algoritmo poderoso. E, por fim, tem a vantagem de não ter um fator limitante no que diz respeito a função objetivo e as restrições, pois não precisa pré-processar o objetivo e as restrições do problema a ser resolvido.

### 3 O PADRÃO IEEE 802.16 E TRABALHOS RELACIONADOS

A quantidade de usuários que utilizam as redes de comunicação tem aumentado significativamente, principalmente com a utilização de serviços de multimídia, como IPTV, videoconferência e VoIP, que requerem uma gama elevada de recursos e geram uma grande demanda sobre a infraestrutura da rede, necessitando de redes mais rápidas e mais confiáveis para suportá-los. Para isso, as redes precisam ser capazes de tratar as aplicações de acordo com as prioridades de cada usuário, oferecendo um nível aceitável de serviço.

Diante deste cenário, as redes de comunicação baseadas na tecnologia WiMAX são uma excelente opção para oferecer taxas de dados elevadas ao longo de grandes distâncias, pois tem como principal objetivo fornecer acesso de banda larga de alta velocidade sem fio para as WMANs, com garantia de QoS (TAGHIPOOR et al., 2012).

Este capítulo tem como objetivo proporcionar uma visão geral do padrão IEEE 802.16, com seus mecanismos de provisão de QoS e escalonamento. Para isso, na Seção 3.1, os aspectos da arquitetura de uma rede WiMAX são apresentados, como a sua topologia, o modelo de referência definido para uma rede WiMAX, com as camadas MAC e física, como é construído um símbolo OFDM e como é a estrutura de um quadro no padrão IEEE 802.16, utilizando a modulação OFDM.

A Seção 3.2 tem o objetivo de proporcionar um entendimento de como as redes WiMAX garantem o suporte a QoS, abordando os principais conceitos de QoS e especificando os mecanismos e políticas para a provisão de QoS. A relação das métricas e dos parâmetros para que uma rede possa garantir QoS para seus usuários é apresentada na Seção 3.2.1. Na Seção 3.2.2 é apresentada a arquitetura de QoS baseada nas cinco classes de serviços definidos pelo padrão IEEE 802.16.

Na Seção 3.3, o funcionamento de um escalonador padrão é apresentado, definindo os seus mecanismos, a classificação existente na literatura e uma abordagem sobre os principais algoritmos de escalonamento para redes WiMAX, com suas vantagens e desvantagens em relação a garantia de QoS.

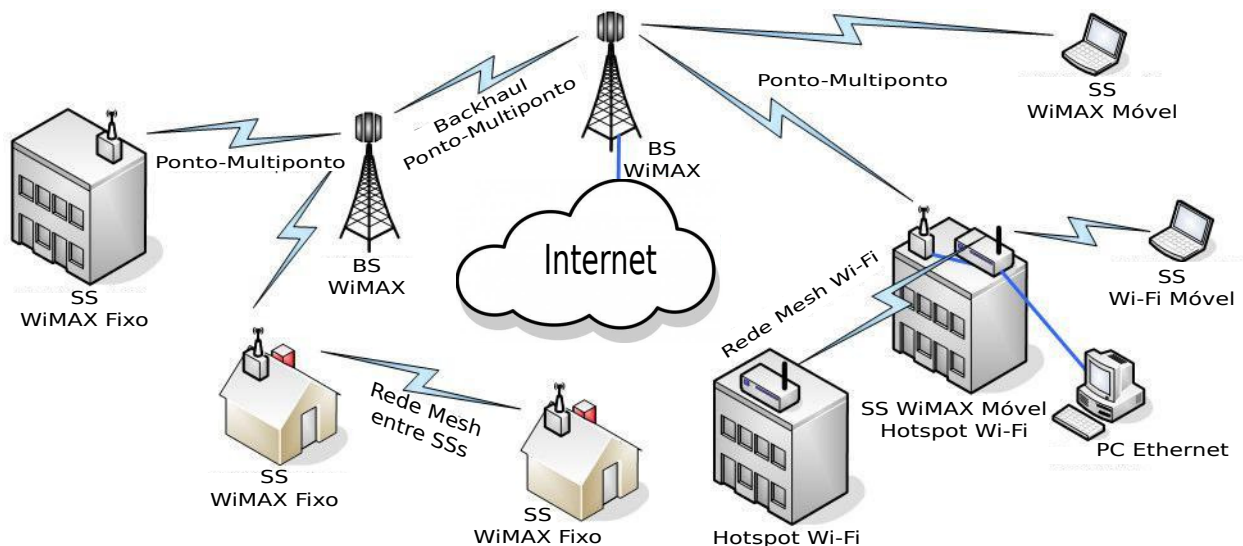
### 3.1 AS REDES NO PADRÃO IEEE 802.16

A ideia das redes normatizadas pelo padrão IEEE 802.16 é oferecer acesso de banda larga de alta velocidade sem fio em áreas metropolitanas, com uma área geográfica mais extensa do que as redes locais, e também suportar a interligação entre redes heterogêneas. O acesso de banda larga de alta velocidade deve atender aos anseios dos usuários finais, como a garantia de qualidade de serviço e uma conectividade de banda larga eficiente.

#### 3.1.1 TOPOLOGIA DE UMA REDE NO PADRÃO IEEE 802.16

Uma rede WiMAX, definida pelo padrão IEEE 802.16 (IEEE, 2012), é composta por duas entidades. As estações base (BS - *Base Station*) provêm o acesso para as estações clientes (SS - *Subscriber Station*) de uma determinada área, denominada célula.

Dependendo da configuração, uma rede WiMAX é composta por uma BS e uma ou várias SSs, conforme apresentado na Figura 5. A BS é o nó central que coordena todas as atividades envolvendo as SSs, ou seja, toda a comunicação que uma SS necessite, deve ser provida pela BS. As BSs são mantidas em torres e espalhadas de forma a otimizar a área de cobertura da rede. Cada uma das BSs que fazem parte da rede, são conectadas entre si por uma rede, e que, conectadas à Internet, permitem que as SSs também tenham acesso à Internet.



**Figura 5: Topologia de uma rede WiMAX**

Fonte: Autoria própria.

O padrão IEEE 802.16 define dois tipos de topologias de comunicação entre os componentes da rede, PMP (ponto-multiponto) e *Mesh*, sendo que a PMP é definida como obrigatória e a *Mesh* como opcional.

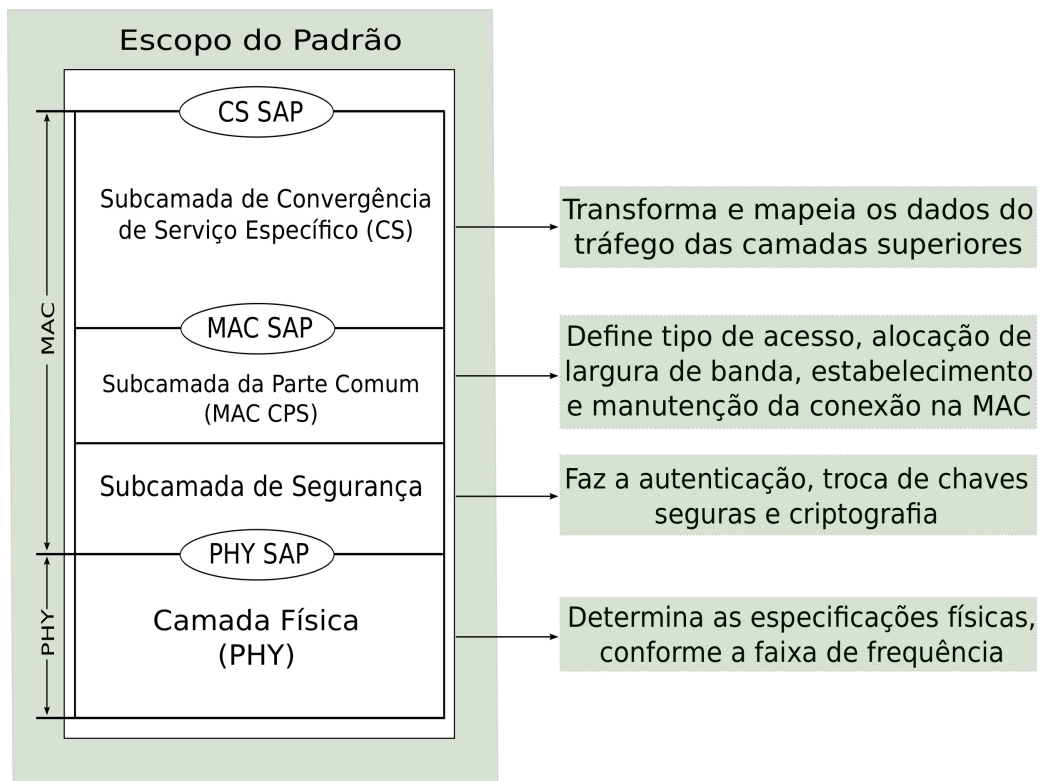
A topologia PMP, está estruturada de tal forma que toda a comunicação entre as SSs é proporcionada pela BS, quer o tráfego seja interno ou para que as SSs tenham acesso às redes externas. Assim, o tráfego flui apenas entre a BS e as SSs e vice-versa. Neste trabalho de pesquisa, esta topologia é utilizada, pois, por ser centralizada, permite que se tenha o controle dos parâmetros da rede, necessários para avaliar o desempenho da arquitetura proposta.

Na topologia *Mesh* existe a comunicação normal entre a BS e as SSs, porém, é permitido que todas as SSs troquem informações entre si, sem a necessidade de um nó central, ou seja, o tráfego pode ser roteado através das SSs e pode ocorrer diretamente entre duas SSs.

### 3.1.2 MODELO DE REFERÊNCIA DO PADRÃO IEEE 802.16

As redes WiMAX, definidas no padrão IEEE 802.16, se baseiam no modelo OSI (*Open Systems Interconnection*) e especificam uma camada física (PHY) e uma camada de controle de acesso ao meio (MAC) para as redes metropolitanas sem fio (WMANs) de banda larga.

O padrão IEEE 802.16, especifica um modelo de referência para as camadas MAC e PHY. A Figura 6 representa este modelo, apresentando o escopo de desenvolvimento do padrão.



**Figura 6: Modelo de referência do padrão IEEE 802.16**

Fonte: Figura adaptada de (IEEE, 2012).

A camada MAC tem a função de controle de acesso ao meio e da garantia do nível de

QoS na interface aérea, por meio de mecanismos dinâmicos de reserva de banda e priorização de tráfego (IEEE, 2012). O suporte a múltiplas interfaces para a camada física e como os mecanismos de sincronismo e segurança da informação funcionam, constituem o conjunto principal de funcionalidades da camada MAC, especificada no padrão IEEE 802.16.

Na camada física (PHY) são especificados o espectro de frequência, os esquemas de modulação, as técnicas de correção de erros, a taxa de dados utilizada, a estrutura de multiplexação e, também, como ocorre a sincronização entre o transmissor e o receptor.

### 3.1.2.1 CAMADA DE CONTROLE DE ACESSO AO MEIO (MAC)

Os serviços requeridos pelos usuários finais de um sistema, tem natureza variada (dados, VoIP, vídeo). A camada MAC, do padrão IEEE 802.16, deve acomodar tanto o tráfego contínuo quanto o tráfego em rajadas e garantir a qualidade destes serviços. O protocolo da camada MAC foi projetado para funcionar sob aplicações ponto-a-multiponto e para suportar taxas muito altas, tanto em direção à BS (tráfego *uplink*), quanto em direção às SSs (tráfego *downlink*) e os algoritmos de acesso e alocação de banda podem compartilhar recursos com vários usuários (IEEE, 2012).

A camada MAC tem algumas funções importantes, dentre elas: (i) o controle de acesso e transmissão dos dados, através dos algoritmos de acesso e alocação de banda, que permitem o compartilhamento do canal e determinam quais estações podem acessar o meio físico; (ii) a provisão de garantia de QoS através de mecanismos de alocação dinâmica de recursos; (iii) atribuição de prioridades de tráfego para o escalonamento do fluxo de dados; (iv) a garantia de segurança; (v) o ajuste adaptativo das técnicas de transmissão; (vi) a interoperabilidade das redes com suporte a várias topologias de redes; e (vii) a multiplexação dos fluxos de tráfego por conexão. Essas várias funções exercidas pela camada MAC possibilitam que ela se torne flexível o suficiente para suportar qualquer tipo de tráfego.

De acordo com o padrão IEEE 802.16, e conforme apresentado na Figura 6, a camada MAC é dividida em três subcamadas:

1. **Subcamada de convergência de serviço específico (CS - *Convergence Sublayer*):** esta subcamada realiza a interface entre a camada MAC e as camadas superiores, fazendo o mapeamento do tráfego proveniente da camada de rede. Os pacotes de dados recebidos da camada de rede, denominados de PDUs (*Protocol Data Units*), são classificados e transformados em dados *MAC SDUs* - (*Service Data Units*). A CS é responsável por associar as SDUs às conexões. Cada conexão possui um identificador único, denominado



CID (*Connection Identifier*), sendo este, relacionado com um determinado fluxo de serviço, identificado como SFID (*Service Flow Identifier*). O padrão estabelece dois tipos de convergência na CS, a **Convergência de Pacotes**, que é definida para mapear serviços baseados em pacotes, como IPv4, IPv6 e Ethernet e a **Convergência ATM**, definida para serviços ATM (*Asynchronous Transfer Mode*). A CS é responsável por realizar todas as operações que dependem da natureza das camadas superiores, como compressão de cabeçalhos e mapeamento de endereços de identificadores.

2. **Subcamada da Parte Comum (CPS - Common Part Sublayer)**: a CPS é considerada a principal subcamada da camada MAC, pelo fato de executar diversas funcionalidades, tais como: o controle de acesso ao sistema, o gerenciamento das conexões (estabelecimento e manutenção), a construção dos PDUs, o suporte à camada física através do ajuste adaptativo das técnicas de transmissão, a multiplexação dos fluxos de tráfego em conexões, a provisão e garantia de QoS através do mecanismo de requisição de largura de banda, o mecanismo de controle de admissão de conexão e o mecanismo de escalonamento e alocação dinâmica de recursos.
3. **Subcamada de Segurança (Security Sublayer)**: esta subcamada provê as funcionalidades de segurança e privacidade dos dados. É responsável pela autenticação para o acesso à rede, troca de chaves de segurança, criptografia dos dados, encriptação, autorização e troca de chaves de criptografia entre os componentes da rede.

Em termos resumidos, a camada MAC é responsável pelo controle de acesso ao meio, pelo estabelecimento e manutenção das conexões e pela provisão de QoS através dos mecanismos de controle de admissão, dos mecanismos de requisição de largura de banda e dos mecanismos de escalonamento (IEEE, 2012). É na camada MAC que são especificadas as funções associadas aos serviços oferecidos aos usuários, que incluem, desde a transmissão de dados em quadros, até o controle do acesso ao meio sem fio compartilhado.

### 3.1.2.2 CAMADA FÍSICA (PHY)

O padrão IEEE 802.16 suporta múltiplas especificações na camada física, devido ao conceito de uso da modulação. Por modulação, entende-se a técnica onde as características da portadora (sinal que é modulado) são modificadas com a finalidade de transmitir as informações, sendo feitas as alterações combinadas de frequência, amplitude ou fase.

É na camada física que estão definidos vários requisitos do sistema, dentre eles (IEEE, 2012): (i) as faixas de frequências; (ii) as interfaces físicas suportadas de acordo com sua

aplicabilidade; (iii) as técnicas de modulação e codificação adaptativas para a transmissão digital; (iv) as técnicas de duplexação e correção de erros; e (v) a construção dos quadros e subquadros.

### 3.1.2.3 FAIXAS DE FREQUÊNCIA SUPOSTADAS

O WiMAX opera em faixas de frequência que podem ser licenciadas ou isentas de licença, descritas a seguir:

- **Faixas Licenciadas de 10-66 GHz:** Nestas faixas de frequência, devido ao menor comprimento de onda, é requerida uma linha de visada (LoS), resultando na negligência dos efeitos da propagação multipercurso. O padrão se compromete a oferecer taxa de dados acima de 120 Mbps nesta faixa de frequência, com a modulação de portadora única *WirelessMAN-SC* (IEEE, 2012). Canais com esta faixa de frequência utilizam larguras de banda típicas de 25 MHz ou 28 MHz. Essas faixas são ideais para o acesso na topologia PMP (*Point-to-Multipoint*) para os usuários domésticos ou em ambientes com volume pequeno de aplicações.
- **Faixas Licenciadas e Isentas de Licença de 2-11 GHz:** Nestas faixas de frequência, em ambas, licenciadas e isentas de licença, foram introduzidas funcionalidades físicas adicionais para operar em ambientes NLoS (*Non-line-of-sight*) e para atenuar os efeitos da propagação multipercurso. A maioria dos recursos mais vantajosos da camada física do padrão IEEE 802.16 são encontrados nesta faixa de frequência. Mecanismos de endereçamento das camadas MAC e física, como DFS (*Dynamic Frequency Selection*) atuam em faixas de frequência isentas de licença, principalmente entre 5 e 6 GHz, para detectar, anular e prevenir as interferências que venham a prejudicar outros usuários.

No Brasil as bandas licenciadas pela Anatel (Agência Nacional de Telecomunicações) (ANATEL, 2015), estão nas faixas de 2,5 GHz, 3,5 GHz e 10,5 GHz. A banda com 2,5 GHz é considerada como a melhor frequência disponível para a tecnologia WiMAX no Brasil. Por ter uma frequência mais baixa que outras, possui maior alcance. Já as bandas isentas de licença estão nas faixas de 2,4 GHz e 5,8 GHz e podem ser utilizadas por qualquer operadora de serviços.

Como o padrão WiMAX é projetado para transmissões em longas distâncias, a banda de 5,8 GHz é a preferida pelo mercado, por sofrer menos interferências que a 2,4 GHz (que também é utilizada no padrão 802.11a). Entretanto, utilizar bandas isentas de licença

pode ter algumas desvantagens, como interferências, maior concorrência e potência limitada (OLIVEIRA et al., 2012).

### 3.1.2.4 ESPECIFICAÇÕES DA INTERFACE FÍSICA

Existem cinco interfaces físicas especificadas pela norma vigente do padrão IEEE 802.16. A Tabela 1 mostra um resumo das especificações das interfaces e apenas a especificação WirelessMAN-OFDM (*Wireless Metropolitan Area Networks - Orthogonal Frequency Division Multiplexing*) será detalhada, pois é utilizada na arquitetura proposta.

**Tabela 1: Resumo das interfaces físicas**

Fonte: Autoria própria.

Designação	Faixas de Frequência	Técnica de Duplexação	Características
WirelessMAN-SC	10-66 GHz	TDD, FDD	Modulação com portadora única. Transmissões em rajadas.
WirelessMAN-SCa	2-11 GHz Faixas Licenciadas	TDD, FDD	Modulação com portadora única. Suporta operações NLoS.
WirelessMAN-OFDM	2-11 GHz Faixas Licenciadas	TDD, FDD	Modulação OFDM. Suporta operações NLoS.
WirelessMAN-OFDMA	2-11 GHz Faixas Licenciadas	TDD, FDD	Modulação OFDM dividido em subgrupos para fornecer acesso múltiplo em uma faixa de frequência única.
WirelessHUMAN	2-11 GHz Faixas Isentas de Licença	TDD	Pode ser SC, OFDM ou OFDMA. Deve incluir DFS para redução de interferências.

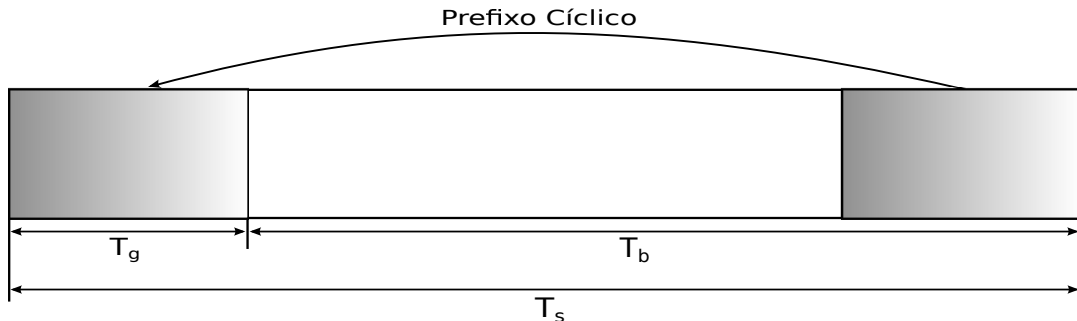
**WirelessMAN-OFDM:** A especificação WirelessMAN-OFDM é uma técnica de modulação multiportadora, baseada na modulação OFDM, tendo como ideia a divisão de um fluxo de dados, com alta taxa de *bits*, em vários fluxos de dados com baixa taxa de *bits*, todos em paralelo, mas cada fluxo modulando portadoras separadas, que são denominadas de subportadoras.

Seu modelo de funcionamento consiste em dividir o espectro de frequências em subportadoras ortogonais, igualmente espaçadas, que transmitem uma pequena porção da informação, denominada de símbolo OFDM. O símbolo OFDM é formado por 256 subportadoras o que permite suportar múltiplas SSs. O *WiMAX Forum* (WIMAX, 2015) adotou esta especificação física para as BWAs. Considerando que as frequências são ortogonais, elas podem sobrepor-se e, por conseguinte, as interferências são minimizadas e a eficiência do espectro é maximizada.

Quando um símbolo OFDM é transmitido, através de um canal de dispersão no tempo, dois tipos de dificuldades surgem. Em primeiro lugar, a dispersão do canal destrói a ortogonalidade entre subportadoras e causa interferência interportadora (ICI - *Inter-Carrier Interference*) no sinal. Em segundo lugar, em determinados momentos, o sistema pode transmitir múltiplos símbolos OFDM em série, causando interferência intersimbólica (ISI -

*Inter-Symbol Interference*) entre os símbolos OFDM sucessivos. Intervalos de guarda foram propostos como uma solução para estas dificuldades. Um intervalo de guarda é definido como sendo um espaço vazio entre dois símbolos OFDM, que serve como uma área de armazenamento para a reflexão de múltiplas trajetórias.

Com a utilização de redes sem fio, a propagação de multicaminhos<sup>1</sup> é uma consequência, causando atrasos na propagação do canal e, com isso, aumentando a duração do símbolo OFDM. Para minimizar este problema, o padrão IEEE 802.16 define a utilização de um intervalo de guarda, denominado de prefixo cíclico (CP - *Cyclic Prefix*), que é uma cópia da última parte do símbolo OFDM adicionada no início do bloco de dados do símbolo OFDM (HOYMANN, 2005), como apresentado na Figura 7.



**Figura 7: Estrutura do símbolo OFDM no domínio do tempo**  
**Fonte: Figura adaptada de (IEEE, 2012).**

### 3.1.3 A CONSTRUÇÃO DE UM SÍMBOLO OFDM

A modulação OFDM é implementada pela utilização da transformada inversa rápida de *Fourier* (IFFT - *Inverse Fast Fourier Transform*)<sup>2</sup>. Um símbolo OFDM é representado em função do domínio do tempo e em função do domínio da frequência. Na seção 4.3, quando da apresentação da nova arquitetura proposta, a construção de um símbolo OFDM é descrita com mais detalhes.

#### 3.1.3.1 SÍMBOLO OFDM NO DOMÍNIO DO TEMPO

A Figura 7, que representa um símbolo OFDM no domínio do tempo, apresenta como o CP é utilizado. Neste esquema,  $T_b$  representa o tempo útil do símbolo OFDM e  $T_g$  corresponde ao tempo de guarda, que, como descrito, é formado por uma cópia do último período do tempo útil do símbolo OFDM.  $T_s$  é o tempo total de duração de um símbolo OFDM.

<sup>1</sup> Propagação de multicaminhos é o fenômeno pelo qual um sinal chega à antena do receptor por caminhos múltiplos, devido à reflexão.

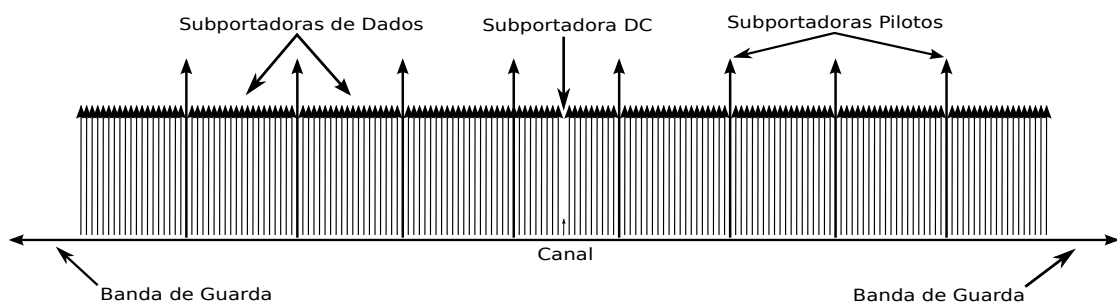
<sup>2</sup> A IFFT converte uma função no domínio da frequência para uma função no domínio do tempo.

Embora os símbolos gerados sejam sempre ortogonais, a inserção do CP tem seu próprio custo. Parte da energia do símbolo OFDM é perdida, uma vez que não carrega nenhuma informação. A vantagem obtida através da introdução do CP é a anulação dos efeitos de ICI e ISI, mesmo que parte da energia do símbolo seja perdida. Assim, o CP combate os dois problemas principais da transmissão dos símbolos OFDM, remove o efeito de ISI, e, através da manutenção de ortogonalidade, remove completamente a ICI (PRASAD; VELEZ, 2010).

### 3.1.3.2 SÍMBOLO OFDM NO DOMÍNIO DA FREQUÊNCIA

O padrão IEEE 802.16 especifica que um símbolo OFDM é criado sobre as subportadoras que são determinadas pelo tamanho da FFT (*Fast Fourier Transform*) utilizada. No caso da especificação *WirelessMAN-OFDM*, 256 subportadoras são determinadas, divididas em 192 subportadoras de dados, 8 subportadoras para pilotos e 56 subportadoras nulas. O propósito da subportadora piloto é servir como uma referência para minimizar os erros de fase e de frequência, e também para estimar a resposta do canal. As 56 subportadoras nulas são usadas, uma para subportadora DC (*Direct Current*) e 27 e 28 subportadoras, para os intervalos de guarda, nos extremos da banda, para apontar o início e o fim do espectro de frequência. (ANDREWS et al., 2007).

A Figura 8 apresenta um símbolo OFDM no domínio da frequência, com os três tipos de subportadoras: (i) **Subportadora de dados**, para transmissão de dados; (ii) **Subportadora Nulas**, para as bandas de guarda e para a subportadora DC; e (iii) **Subportadoras pilotos**, para vários propósitos, como estimação, controle de potência e sincronismo.



**Figura 8: Símbolo OFDM no domínio da frequência**

Fonte: Figura adaptada de (IEEE, 2012).

Com as definições expostas acima, é possível especificar a maneira como as redes WiMAX trabalham.

### 3.1.4 ESTRUTURA DE UM QUADRO NO PADRÃO IEEE 802.16

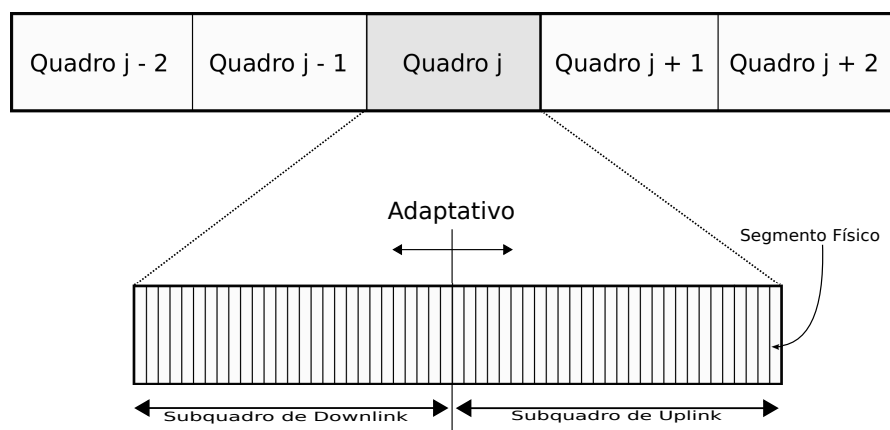
No padrão IEEE 802.16, a camada física opera com base em quadros (*frames*). A norma especifica que os quadros tenham a duração no domínio do tempo entre 2,50 *ms* e 20,00 *ms*. Também define que quando a BS estabelece a duração do quadro, este valor não deve ser alterado, pois se for alterado, todas as SSs devem novamente se sincronizar com a BS (IEEE, 2012). A Tabela 2 apresenta os valores padrões da duração de um tempo de quadro e a quantidade de quadros por segundo, especificados pela norma.

**Tabela 2:** Duração do tempo de um quadro OFDM  
Fonte: (IEEE, 2012).

Duração do Quadro ( <i>ms</i> )	Quadros por Segundo
2,50	400
4,00	250
5,00	200
8,00	125
12,50	80
20,00	50

Neste trabalho de pesquisa, a duração do tempo do quadro é calculada de acordo com os parâmetros da rede, desta forma, em cada iteração do sistema, se ocorreu algum tipo de alteração nos parâmetros da rede, a duração do tempo do quadro é recalculada, utilizando um algoritmo de otimização baseado em PSO e as SSs são sincronizadas novamente com a BS.

A Figura 9 mostra a estrutura de um quadro. A duração dos subquadros de *downlink* e *uplink* podem apresentar tamanhos variáveis e adaptativos, conforme os parâmetros utilizados pelo sistema.



**Figura 9:** Estrutura de um quadro no esquema de transmissão TDD

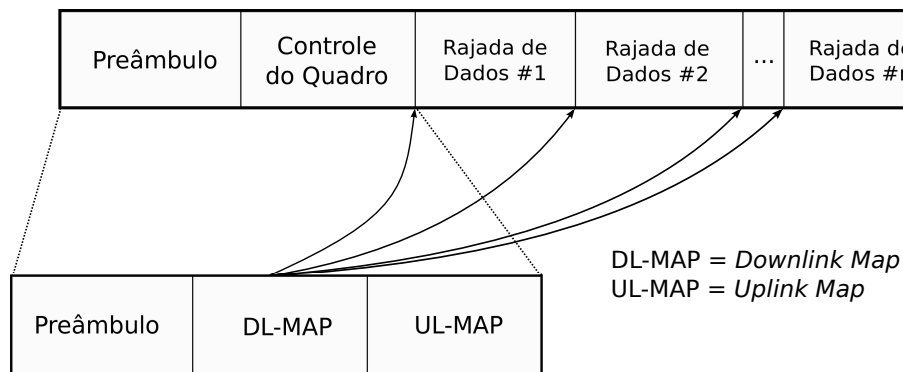
Fonte: Figura adaptada de (IEEE, 2012).

Os quadros são constituídos por intervalos de tempo denominados de segmentos físicos (*physical slots* - PS), sendo que o número de segmentos físicos existentes em um quadro é uma função da duração de um símbolo OFDM e do tempo do quadro. Cada quadro é composto

por dois subquadros, denominados de subquadro de *downlink* e subquadro de *uplink*. A BS se utiliza do subquadro de *downlink* para enviar dados e informações de controle para as SSs. Por outro lado, o subquadro de *uplink* é utilizado pelas SSs para enviar dados e mensagens de requisição de largura de banda para a BS.

#### 3.1.4.1 SUBQUADRO DE *DOWNLINK*

A transmissão no subquadro de *downlink* é realizada em *broadcast*. Como na comunicação *downlink*, somente a BS transmite, então, todas as SSs receberão os dados. Somente para a SS a qual os dados são destinados é que será permitido o acesso. A Figura 10 ilustra graficamente o formato de um subquadro de *downlink*.



**Figura 10: Formato do subquadro de *downlink***

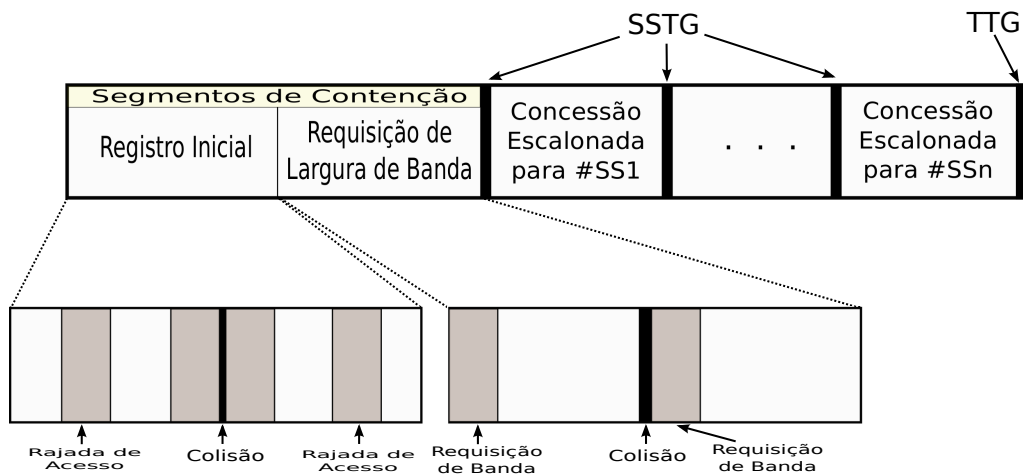
Fonte: Figura adaptada de (IEEE, 2012).

#### 3.1.4.2 SUBQUADRO DE *UPLINK*

De acordo com o padrão IEEE 802.16, três tipos de rajadas podem ser transmitidas pelas SSs:

1. Rajadas que são transmitidas em segmentos de contenção reservados para o processo de registro inicial (*initial ranging*);
2. Rajadas que são transmitidas em segmentos de contenção reservados para mensagens de requisição de largura de banda;
3. Rajadas que são transmitidas em intervalos definidos e alocados para SSs individuais para transmissão de dados no sentido *uplink*.

A Figura 11 mostra a estrutura de um subquadro de *uplink*, que tem como principal objetivo fazer com que as SSs se comuniquem com a BS para enviar seus dados.



**Figura 11: Formato do subquadro de *uplink***

Fonte: Figura adaptada de (IEEE, 2012).

Assim, o formato de um subquadro de *uplink* consiste em um segmento de contenção para o registro inicial de uma SS na rede, outro segmento de contenção para requisição de banda pelos fluxos de menor prioridade (*bandwidth request*), e várias rajadas *uplink*, denominadas de concessão escalonada, cada uma transmitida por uma SS para envio de dados e requisição de banda (Figura 11).

Entre as concessões escalonadas, existe um intervalo denominado de SSTG (*Subscriber Station Transition Gap*), que permite a finalização completa do sinal anterior e a sincronização da BS com a nova SS.

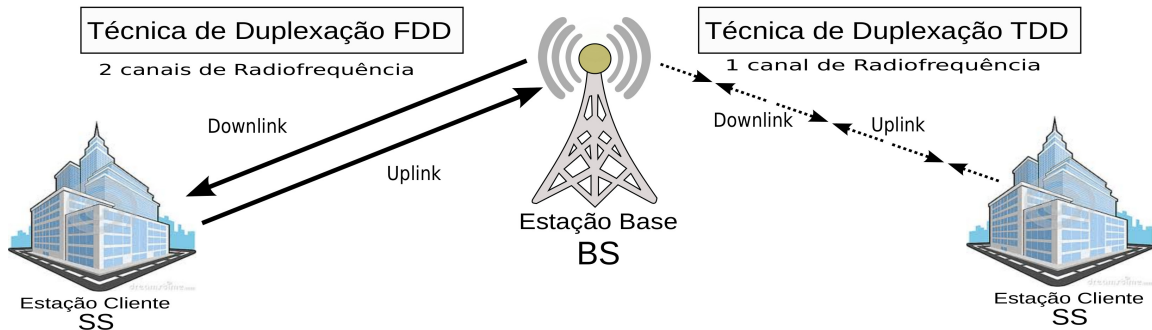
Ao final de cada subquadro, existe um intervalo de tempo para que a BS possa alterar o seu modo de operação, de transmissão para recepção, e vice-versa. Este intervalo é denominado de TTG (*Transmitter-receiver Turnaround Gap*) quando o modo de operação é alterado de transmissor (*downlink*) para receptor (*uplink*) e, RTG (*Receiver-transmitter Turnaround Gap*), quando o modo de operação do quadro é alterado de receptor (*uplink*) para transmissor (*downlink*).

#### 3.1.4.3 MODOS DE ACESSO AO MEIO FÍSICO

O padrão IEEE 802.16 especifica dois modos de acesso ao meio físico, duplexação por divisão de tempo (TDD - *Time Division Duplexing*) e duplexação por divisão de frequência (FDD - *Frequency Division Duplexing*). O modo de acesso FDD utiliza os subquadros de *downlink* e *uplink* simultaneamente no tempo, porém em frequências de portadoras e canais diferentes. Já no modo de acesso TDD, os dados compartilham um mesmo canal de transmissão, ou seja, a mesma frequência, tanto no sentido *downlink* quanto no sentido *uplink*,



porém, os dados são transmitidos em tempos diferentes. A Figura 12 ilustra a estrutura do quadro operando nos modos TDD e FDD.



**Figura 12: Técnicas de duplexação para alocação de banda no padrão IEEE 802.16**  
**Fonte: Autoria própria.**

### 3.2 QUALIDADE DE SERVIÇO PARA REDES IEEE 802.16

O grande objetivo vislumbrado pelas redes WiMAX, é a possibilidade da existência de redes eficientes e escalonáveis para dados, vídeo e voz. Para conseguir isso, um sistema de garantia de QoS precisa estar presente no momento da transmissão dos dados dos usuários pelas redes de comunicação. Nas redes de acesso sem fio, um grande número de questões técnicas precisam ser solucionadas e, a garantia de QoS, é a principal delas, porque impõe muitos desafios relacionados a fatores como a limitação da largura de banda disponível (MORAES, 2013).

Para atender as necessidades dos usuários, no que diz respeito a provisão de QoS, um conjunto de parâmetros é especificado e estabelecidos nos contratos de nível de serviço e que expressam os anseios dos usuários. Esses parâmetros são definidos através de requisitos de desempenho do sistema, tais como, atraso ou latência (*delay*), vazão (*throughput*) e variação do atraso (*jitter*) (IEEE, 2012). Assim, as redes WiMAX necessitam de mecanismos que garantam a provisão de QoS solicitados pelos usuários, para que possam gerenciar esses parâmetros.

#### 3.2.1 MÉTRICAS E PARÂMETROS PARA PROVISÃO DE QoS EM REDES PADRÃO IEEE 802.16

O objetivo de uma rede de computadores, no que diz respeito a provisão de QoS, é fornecer um serviço que atenda adequadamente as necessidades dos usuários, maximizando a utilização de recursos da rede (MA; LU, 2009). Assim, o provimento de QoS é de fundamental

importância em redes de computadores, quando se utiliza aplicações em tempo real, pelo fato de que elas demandam largura de banda e serviços diferenciados, podendo necessitar, em algumas aplicações, que a transmissão de dados seja realizada sem interrupção ou perda de pacotes. Ainda, segundo Ma e Lu (2009), para que uma rede possa oferecer um nível de QoS que satisfaça os desejos dos usuários, ela deve gerenciar seus próprios recursos, através da análise dos requisitos da aplicação, do gerenciamento de recursos e da aplicação dos vários mecanismos de escalonamento disponíveis. Para isso, algumas métricas são definidas e devem estar dentro de limites bem definidos (valor mínimo, valor máximo):

- **Vazão** (*throughput*): é a quantidade de dados transmitidos com sucesso durante um determinado intervalo de tempo, definida em *bps* (*bits* por segundo);
- **Atraso ou latência** (*delay*): é o intervalo de tempo, em segundos, compreendido entre o envio de um pacote pelo nó de origem e o seu recebimento no nó de destino;
- **Variação do Atraso** (*jitter*): é a variação do atraso entre os pacotes de dados que são transmitidos de forma consecutiva;
- **Taxa de Erros**: é a razão entre a quantidade de pacotes recebidos com erros pela quantidade de pacotes transmitidos;
- **Taxa de perda de pacotes**: é a razão entre a quantidade de pacotes perdidos e a quantidade de pacotes enviados, considerando os diversos tipos de perdas na rede, como descarte nas filas intermediárias e perdas causadas por fenômenos eletromagnéticos no meio físico.

Levando estes parâmetros em consideração, o padrão IEEE 802.16 define mecanismos de sinalização e funções para controle de transmissões de dados entre a BS e as SSs, para os tráfegos no sentido *uplink* e *downlink*, que são implementados na camada MAC. A QoS é implementada por processos de solicitação e concessão de largura de banda, que consiste em associar pacotes que passam pela camada MAC a um fluxo de serviço que fornece um transporte unidirecional aos dados. Essa associação permite fornecer um tratamento diferenciado aos tráfegos que são gerados pelas aplicações. Durante o estabelecimento da conexão, os fluxos de serviço são criados e ativados pela BS e pelas SSs. Assim, cada fluxo de serviço é caracterizado por um conjunto de parâmetros de QoS, definidos pelo padrão IEEE 802.16 da seguinte forma:

- **Minimum Reserved Traffic Rate (MRTR)**: determina a taxa de tráfego mínima que foi reservada para um determinado fluxo de serviço. De acordo com o padrão, se a quantidade

de banda requisitada for menor que o valor de MRTR, a BS deve alocar a banda excedente para outras conexões. O valor do parâmetro MRTR garante uma vazão mínima para cada aplicação;

- *Maximum Sustained Traffic Rate (MSTR)*: determina a taxa de pico máxima do serviço no sentido de tráfego *downlink* e *uplink*. A SS deve controlar o serviço no sentido *uplink* para que a taxa seja igual ao valor definido pelo parâmetro, enquanto que na BS no sentido *downlink*, é assumido que o tráfego é controlado na entrada da rede, não sendo necessário realizar um controle adicional;
- *Maximum Latency*: define a latência máxima entre a recepção do pacote na interface de rede da BS ou da SS e a transmissão do pacote pela interface de radiofrequência (RF);
- *Tolerated Jitter*: define a variação permitida para a latência da conexão;
- *Traffic Priority*: especifica a prioridade atribuída ao tráfego. Dados dois fluxos de serviço idênticos em todos os parâmetros de QoS exceto na prioridade, o fluxo de serviço com maior prioridade deve ter latência menor e prioridade maior no processo de armazenamento no *buffer*;
- *Request/Transmission Policy*: tem a capacidade de especificar alguns atributos para um fluxo de serviço associado, como por exemplo, a restrição sobre os tipos de opções de requisição de largura de banda que podem ser utilizados.

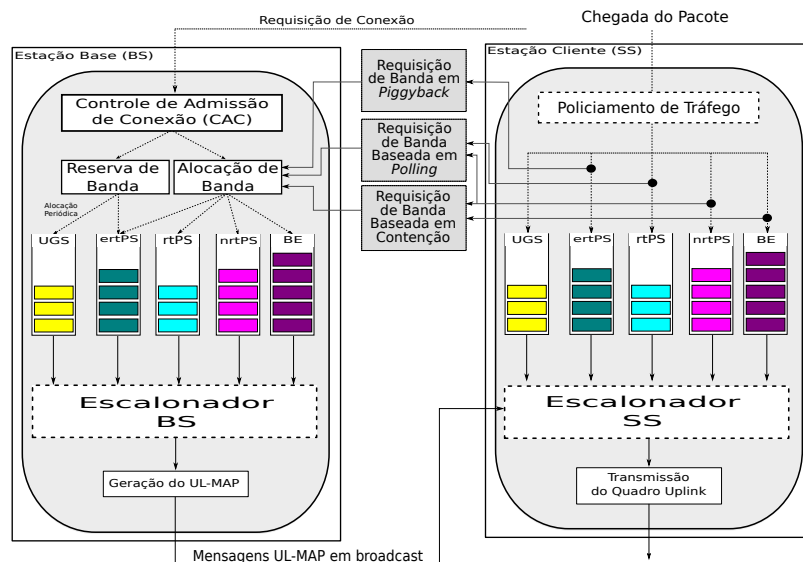
### 3.2.2 QoS PARA AS CLASSES DE SERVIÇO DAS REDES PADRÃO IEEE 802.16

A provisão de QoS depende da classificação dos pacotes gerados pelas aplicações das camadas superiores que passam pela camada MAC e da associação do fluxo de serviço a uma determinada conexão. Assim, na arquitetura de QoS do padrão IEEE 802.16, as classes de serviços representam um conjunto de parâmetros para a entrega de dados em um determinado fluxo de dados, tendo, como finalidade, o compartilhamento de largura de banda entre os diferentes tipos de usuários. Desta forma, nas redes IEEE 802.16, o serviço de escalonamento representa o mecanismo de transporte de dados em uma determinada conexão. Cada conexão é associada a somente um tipo de serviço. Cada serviço é associado a um conjunto de parâmetros de QoS que define o comportamento da conexão.

A fim de fornecer suporte a QoS eficiente para os usuários finais, o seguinte conjunto de componentes são necessários (TANG et al., 2010):

- **Controle de Admissão de Conexão (CAC):** responsável por realizar o controle da quantidade de conexões com base na largura de banda total disponível e a largura de banda solicitada por cada conexão. O CAC faz a restrição do número de usuários simultaneamente ativos na rede, com o objetivo de evitar a saturação dos recursos e, conseqüentemente, a degradação do desempenho dos parâmetros de QoS dos fluxos ativos;
- **Escalonamento de Pacotes:** aloca intervalos de tempo, denominados de segmentos, para diferentes tipos de conexões de acordo com os requisitos de QoS;
- **Policimento de Tráfego:** controla o volume de tráfego enviado em uma rede em um determinado período. É útil para proteger a rede contra tráfegos mal comportados que violam parâmetros negociados durante a fase de estabelecimento da conexão.

O funcionamento destes mecanismos é realizado através dos requisitos informados pelo conjunto de tráfego das classes de serviços que compõem a arquitetura de QoS. O padrão IEEE 802.16 define uma arquitetura de QoS nativa para as redes WiMAX, que é composta por um conjunto de mecanismos, conforme pode ser observado na Figura 13. A nova arquitetura desenvolvida neste trabalho vai atuar somente no escalonador BS e é discutida na Seção 4.1 do Capítulo 4.



**Figura 13: Arquitetura de QoS nativa do padrão IEEE 802.16**

Fonte: Figura adaptada de (WANG et al., 2010).

O padrão IEEE 802.16 especifica que o CAC é necessário para garantir que o tráfego agregado não resulte em uma sobrecarga na rede ou na degradação de serviço para o tráfego existente. Também especifica que o processo de requisição e concessão de largura de banda

consiste em definir e reservar uma quantidade adequada de segmentos de tempo para cada conexão transmitir seus dados.

A alocação de recursos para as SSs é realizada sob demanda. Quando uma SS precisa de largura de banda para uma conexão, ela envia uma mensagem de requisição de banda para a BS. Uma requisição de banda pode ser enviada como um pacote, cujo cabeçalho indica a presença de um pedido de banda, ou pode ser enviada juntamente com um pacote de dados (*piggyback*<sup>3</sup>). Todas as requisições de banda devem indicar o número de *bytes* necessários para transmitir os dados (*payload*<sup>4</sup>) e o cabeçalho MAC.

A requisição de largura de banda pode ser incremental ou agregada. Uma requisição incremental indica a largura de banda adicional que a SS precisa, enquanto uma requisição agregada indica a largura de banda total necessária para a SS.

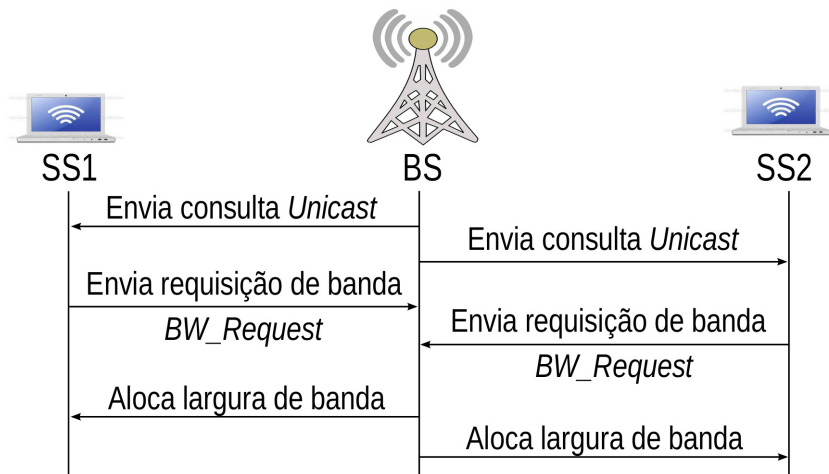
Para a SS, a requisição de banda sempre se refere a uma determinada conexão, enquanto que um *grant* (concessão) alocado pela BS é destinado a uma SS e não a uma conexão em particular. Dessa forma, a SS pode utilizar o *grant* recebido para uma conexão diferente daquela para a qual a requisição foi solicitada.

A alocação de *grants* para o envio de requisições de banda pode ser para uma SS particular ou para um grupo de SSs e recebe o nome de *polling*. O padrão define dois mecanismos de *polling* (FERREIRA, 2011):

- ***Polling Unicast***: A BS consulta cada uma das SSs para identificar se elas necessitam utilizar o meio de transmissão e, em caso afirmativo, a BS aloca largura de banda para as SSs enviarem suas mensagens de requisição da banda (*BW\_Request*), como apresentado na Figura 14. A desvantagem desse tipo de *polling* é que se a SS não tem dados aguardando para serem transmitidos, ela não responderá a consulta, acarretando em desperdício de banda.
- ***Polling Baseado em Contenção***: A BS efetua uma consulta via *multicast* ou *broadcast*. Essa técnica é empregada quando a BS não dispõe de recursos suficientes para consultar cada SS individualmente. Desta forma, a BS disponibiliza uma largura de banda onde várias SSs disputam a oportunidade de enviar a sua requisição de largura de banda, como pode ser observado na Figura 15.

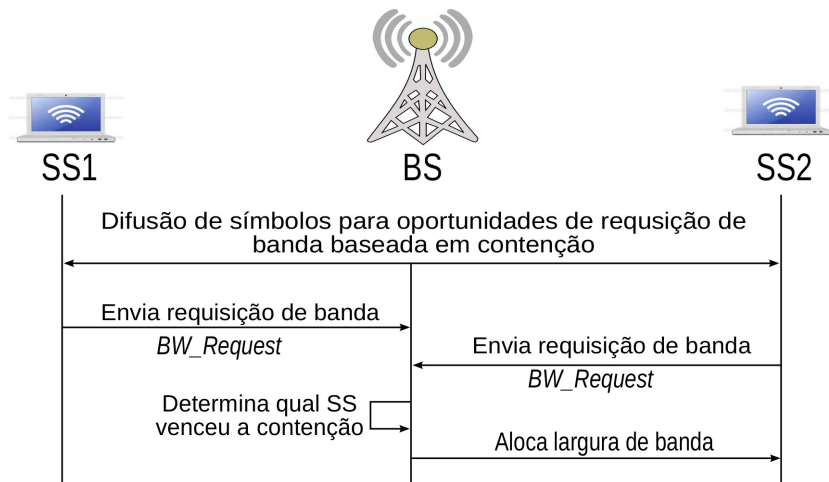
<sup>3</sup> *Piggyback* é quando uma SS envia um pedido de requisição de banda junto com o pacote de dados, o que significa que a requisição está indo de “carona” com o pacote de dados enviado pela SS.

<sup>4</sup> *Payload* se refere a parte dos dados transmitidos, excluindo as informações enviadas com ela, como cabeçalhos.



**Figura 14: Polling unicast**

Fonte: Figura adaptada de (FERREIRA, 2011).



**Figura 15: Polling Baseado em Contenção**

Fonte: Figura adaptada de (FERREIRA, 2011).

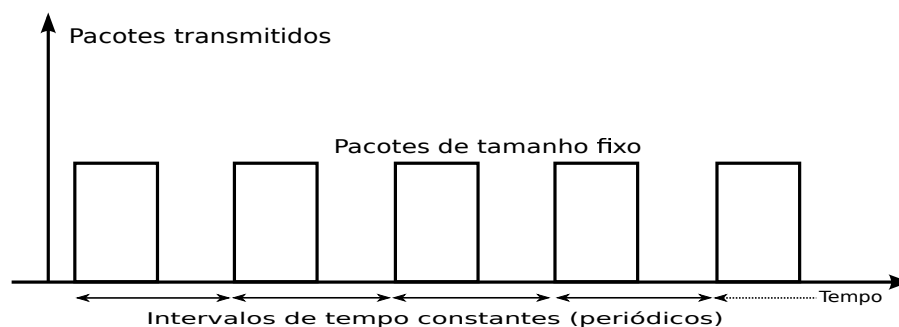
Para atender os requisitos de QoS de cada usuário, o padrão IEEE 802.16 especifica cinco classes de serviço, tratadas de forma diferenciada pelo mecanismo de escalonamento e que definem diferentes níveis de QoS: UGS (*Unsolicited Grant Service*), rtPS (*real-time Polling Service*), ertPS (*extended-real-time Polling Service*), nrtPS (*non-real-time Polling Service*) e BE (*Best Effort*).

### 3.2.2.1 UGS

A classe de serviço UGS é projetada para suportar fluxos de serviço em tempo real que geram pacotes de dados de tamanho fixo em intervalos periódicos. Assim, a UGS é destinada principalmente para serviços com tráfego CBR (*Constante Bit Rate*), tais como T1/E1 e VoIP

sem supressão de silêncio, o que significa que a obtenção de baixa latência e baixo *jitter* é muito importante (AWAL; BOUKHATEM, 2009).

Fluxos UGS oferecem concessões para envio de pacotes de tamanho fixo em intervalos regulares, fazendo com que ocorra o mínimo possível de latência e *jitter* nas requisições das SSs, garantindo sempre que as concessões estarão disponíveis para os fluxos de tempo real. Assim, a SS tem um segmento dedicado, de tamanho fixo, e nunca necessita solicitar largura de banda para este serviço, pois a BS oferece periodicamente concessões de banda não solicitadas para a transmissão de dados, eliminando a sobrecarga na rede. A Figura 16 representa graficamente esta classe de serviço.



**Figura 16: Mecanismo de alocação de concessões *uplink* UGS**

Fonte: (OLIVEIRA, 2011).

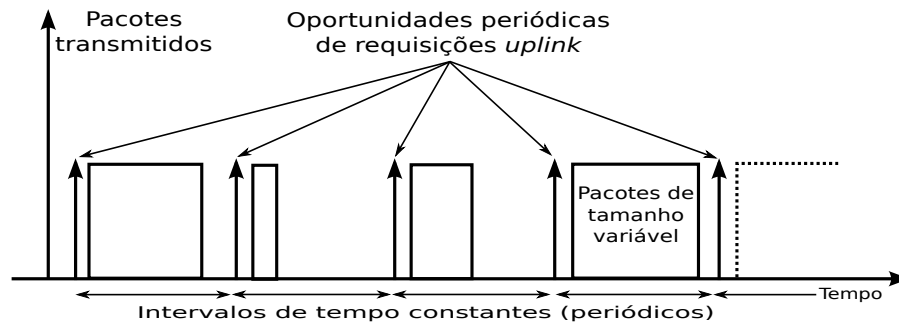
### 3.2.2.2 rtPS

A classe de serviços rtPS foi projetada para dar suporte às aplicações em tempo real que transportem pacotes de dados de tamanho variável em períodos constantes, como, por exemplo, vídeos com codificação MPEG (*Moving Pictures Experts Group*) e *streaming* de áudio.

Esta classe de serviço oferece, em tempo real, oportunidades periódicas de requisição de banda através de *polling unicast*, e permite às SSs informar a quantidade de largura de banda desejada, fazendo com que as necessidades dos fluxos de tempo real (largura de banda e atraso) sejam atendidas satisfatoriamente, pois permite que cada SS especifique o tamanho da concessão desejada. A Figura 17 representa graficamente a classe de serviço rtPS.

### 3.2.2.3 ertPS

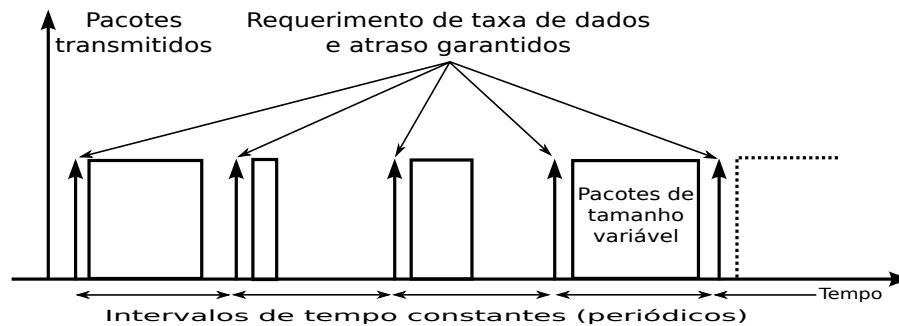
Esta classe de serviço foi projetada para dar suporte as aplicações em tempo real que transportem pacotes de dados de tamanho variável e com requerimentos de taxa de dados e atraso garantidos, como VoIP com supressão de silêncio. A Figura 18 representa graficamente



**Figura 17: Mecanismo de alocação de concessões uplink rtPS**

Fonte: (OLIVEIRA, 2011).

a classe de serviço ertPS.



**Figura 18: Mecanismo de alocação de concessões uplink ertPS**

Fonte: Autoria própria.

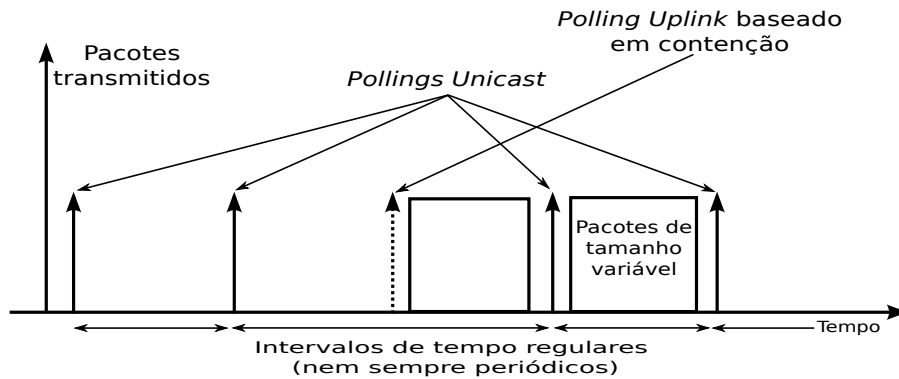
De maneira semelhante a classe de serviço rtPS, na classe de serviço ertPS a BS é responsável pela alocação da largura de banda por conexão, oferecendo, em tempo real, oportunidades periódicas de requisição de banda através de *polling unicast*, de uma forma espontânea, como na classe de serviço UGS, evitando a latência em uma requisição de largura de banda. No entanto, enquanto os fluxos UGS são de tamanho fixo, os fluxos ertPS são variáveis. A BS pode fornecer alocações no sentido *uplink* de forma periódica, que podem ser utilizadas para solicitar largura de banda ou para a transferência de dados.

#### 3.2.2.4 nrtPS

A classe de serviços nrtPS foi projetada para suportar fluxos de serviço que requerem garantia de taxa de dados, mas são tolerantes ao atraso, como a transferência de arquivos. Desta forma, o tráfego de dados não é em tempo real e gera pacotes de tamanho variável em intervalos periódicos, como, por exemplo, o tráfego gerado pelo protocolo FTP (*File Transfer Protocol*), serviço de *e-mail*, serviços de mensagens de texto (SMS - *Short Message Service*) e os serviços de mensagens multimídia (MMS - *Multimedia Messaging Service*). Nesta classe de serviço,



existe uma oferta periódica de oportunidades de requisição *polling unicast*, com intervalos de tempo mais espaçados do que na classe de serviço rtPS, assegurando que o fluxo de serviço seja recebido em oportunidades de requisição, mesmo ocorrendo congestionamento na rede. A Figura 19 representa graficamente esta classe de serviço.

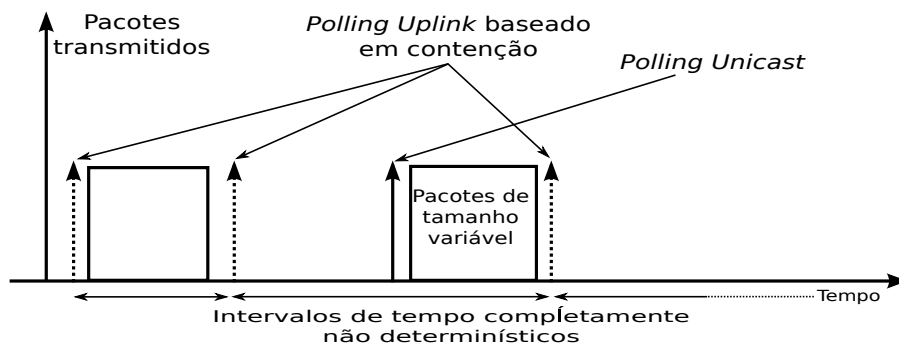


**Figura 19: Mecanismo de alocação de concessões uplink nrtPS**

Fonte: (OLIVEIRA, 2011).

### 3.2.2.5 BE

Esta classe de serviço foi projetada para suportar serviços que não requerem garantia de taxa de dados ou atraso. As SSs podem utilizar pedidos de oportunidades de segmentos de contenção e segmentos *unicast* para requisitar largura de banda, conforme apresentado na Figura 20.



**Figura 20: Mecanismo de alocação de concessões uplink BE**

Fonte: (OLIVEIRA, 2011).

Esta classe de serviço é indicada para tráfego *Web* que possui pacotes de tamanho variável e não é sensível ao atraso, ou seja, seu uso é indicado para aplicações que não possuem nenhum requisito de atraso específico. Para esta classe de serviço, os parâmetros de QoS são escolhidos de tal forma que possam fornecer suporte aos serviços de escalonamento de fluxos

de dados, para os quais nenhuma alocação mínima de recurso foi concedida. Para a classe de serviço BE não existe garantia de QoS.

A Tabela 3 apresenta um resumo das classes de serviços suportadas pelas redes WiMAX, com sua definição, aplicações típicas e os parâmetros que cada classe oferece para as aplicações.

**Tabela 3:** *Resumo das classes de serviço para redes padrão IEEE 802.16*

**Fonte:** *Autoria própria.*

<b>Classe QoS</b>	<b>Resumo</b>	<b>Aplicações Típicas</b>	<b>Parâmetros de QoS</b>
UGS	Fluxo de dados de tempo real compreendendo pacotes de dados de tamanho fixo e intervalos periódicos	VoIP, T1/E1, ATM e CBR	MRTR, MSTR, <i>Maximum Latency</i> , <i>Tolerated Jitter</i> e <i>Request/Transmission Policy</i>
rtPS	Fluxo de dados de tempo real consistindo de pacotes de dados de tamanho variável que são gerados periodicamente	Fluxo de Áudio e Vídeo	MRTR, MSTR, <i>Maximum Latency</i> e <i>Request/Transmission Policy</i>
ertPS	Fluxo de serviço de tempo real que geram pacotes de dados de tamanho variável em um período básico	Voz com Supressão de Silêncio	MRTR, MSTR, <i>Maximum Latency</i> , <i>Traffic Priority</i> e <i>Request/Transmission Policy</i>
nrtPS	Fluxo de dados tolerante a atrasos consistindo de pacotes de dados de tamanho variável para os quais a taxa dados mínima é requerida	<i>File Transfer Protocol</i> (FTP)	MRTR, MSTR, <i>Traffic Priority</i> e <i>Request/Transmission Policy</i>
BE	Fluxo de dados em que nenhum nível de serviço é requerido	Transferência de Dados, Navegador da Web	MSTR, <i>Traffic Priority</i> e <i>Request/Transmission Policy</i>

### 3.3 TRABALHOS RELACIONADOS DE ESCALONAMENTO EM REDES PADRÃO IEEE 802.16

Na tecnologia WiMAX, o escalonamento é um componente crítico e tem um impacto significativo em seu desempenho, já que a tecnologia foi desenvolvida para ter altas taxas de transmissão e requisições de QoS predefinidas, possibilitando a existência de redes eficientes e escalonáveis para dados, vídeo e voz. Porém, o padrão IEEE 802.16 não especifica um mecanismo de escalonamento que garanta a QoS exigida pelas aplicações multimídia, somente define as políticas de escalonamento que devem ser implementadas pelos fabricantes de equipamentos.

Esquemas de escalonamento oferecem um importante auxílio na garantia de serviços às classes heterogêneas de tráfego, onde existe uma variedade de diferentes requerimentos para provisão de QoS. Além do escalonamento, as requisições de largura de banda, controle de admissão e mecanismos de alocação de largura de banda também desempenham um papel fundamental na provisão de QoS para as redes WiMAX.

Um escalonador também deve ser capaz de proteger a rede contra os fluxos com comportamento inadequado e fornecer limites para que a vazão e o atraso tenham um desempenho satisfatório e não comprometam os parâmetros de qualidade da rede. Portanto, os algoritmos de escalonamento constituem o principal mecanismo responsável pelas garantias de desempenho para as diferentes aplicações existentes na rede.

O escalonador é o componente principal da camada MAC, que ajuda a garantir a QoS para várias aplicações (BACIOCCOLA et al., 2010). Assim, a escolha do algoritmo de escalonamento para os sistemas WiMAX é muito importante.

### 3.3.1 MECANISMO DE ESCALONAMENTO PARA REDES PADRÃO IEEE 802.16

Nas redes WiMAX, os pacotes de dados trafegam através de três escalonadores. Dois escalonadores estão implementados na BS, um no sentido de tráfego *downlink*, isto é, da BS para as SSs, e outro no sentido de tráfego *uplink*, ou seja, das SSs para a BS. O terceiro escalonador é implementado na SS, no sentido de tráfego *uplink* (BELGHITH; NUAYMI, 2009). Para a alocação dos recursos, os escalonadores utilizam as informações obtidas sobre os requisitos de QoS e como está a situação atual em cada uma das filas das conexões.

No escalonamento no sentido *downlink*, a BS é a única a transmitir durante o quadro de *downlink* e tem acesso direto às filas, tendo, com isso, o conhecimento completo das condições de cada fila, permitindo montar e enfileirar as rajadas de acordo com os parâmetros de QoS dos fluxos, transmitidos através de conexões *broadcast* para todas as SSs, porém, as SSs somente utilizam os pacotes destinados para elas.

O escalonamento no sentido *uplink*, localizado na BS, é mais complexo do que o escalonamento no sentido *downlink*. Por não ter acesso direto às filas de entrada, que estão localizadas nas SSs, a BS não tem qualquer informação sobre o tempo de chegada dos pacotes nas filas das SSs. Para que a BS se mantenha informada do estado das filas de cada conexão, o escalonador é dependente das mensagens de requisição de banda (*BW\_Request*) enviadas pelas SSs, para que a necessidade de banda seja estimada e a largura de banda seja alocada no quadro *uplink*. Isso faz com que se tenha a necessidade de coordenar a troca de mensagens entre a BS e cada uma das SSs, individualmente.

Segundo (BORIN; FONSECA, 2009), para a implementação de um algoritmo de escalonamento no sentido *uplink* na BS, foco da pesquisa deste trabalho, alguns aspectos, definidos no padrão IEEE 802.16, devem ser levados em consideração:

1. Baseado nos parâmetros de QoS das conexões e das requisições de largura de banda

enviadas pelas SSs, a BS deve distribuir a largura de banda no sentido *uplink* de tal modo que os requisitos de QoS, de cada conexão, sejam satisfeitos;

2. A BS deve alocar largura de banda, não somente para a transmissão de dados, mas também para a transmissão das requisições de largura de banda de acordo com o mecanismo de requisição definido para cada tipo de fluxo de serviço;
3. O escalonador no sentido *uplink* deve garantir todos os requisitos de QoS definidos pelo padrão.

Para abordar o primeiro e terceiro aspectos de implementação, o número de conexões simultâneas no sistema deve ser restringido, a fim de evitar a saturação do canal sem fio. Isto significa que um mecanismo de controle de admissão de conexões (CAC) deve trabalhar em conjunto com o escalonador *uplink* para garantir os requisitos de QoS. O padrão IEEE 802.16 menciona a necessidade de um mecanismo de CAC, porém, não define qualquer política, apenas elenca os parâmetros que devem ser seguidos para que se possa implementá-lo. Além disso, a implementação de um escalonador deve utilizar, de forma eficiente, a largura de banda disponível, possibilitando que uma quantidade maior de usuários sejam alocados na rede, aumentando o nível de utilização do sistema.

### 3.3.2 CLASSIFICAÇÃO DOS ALGORITMOS DE ESCALONAMENTO

Os algoritmos de escalonamento para redes IEEE 802.16, abordados na literatura, foram desenvolvidos em conformidade com algumas características definidas na norma, sendo classificados, no que diz respeito às características da camada MAC IEEE 802.16 e da camada física OFDM, em algoritmos homogêneos, algoritmos híbridos e algoritmos oportunistas.

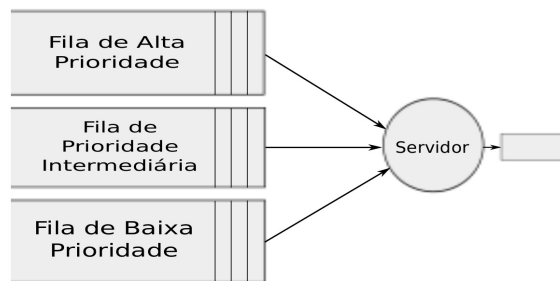
Os algoritmos homogêneos e híbridos utilizam políticas de escalonamento clássicas, sendo que a categoria dos algoritmos híbridos emprega vários desses algoritmos em uma tentativa de satisfazer os requisitos de QoS do tráfego multiclasse em redes WiMAX. A categoria de algoritmos oportunistas refere-se a algoritmos que exploram variações nas condições do canal de comunicação em redes WiMAX, incorporando os requisitos de QoS na sua concepção de escalonamento (ZHANG, 2009).

#### 3.3.2.1 ALGORITMOS DE ESCALONAMENTO HOMOGÊNEOS

Os algoritmos de escalonamento homogêneos, também conhecidos como clássicos ou tradicionais, têm como característica principal a utilização de apenas uma disciplina

de escalonamento<sup>5</sup> para tratar de todas as classes de serviços de QoS. As disciplinas de escalonamento *Priority Queuing* (PQ), *Round Robin* (RR), *Weighted Round Robin* (WRR), *Deficit Round Robin* (DRR), *Earliest Deadline First* (EDF) e *Weighted Fair Queuing* (WFQ) são utilizadas pelos algoritmos homogêneos. Algumas das disciplinas de escalonamento abordadas são descritas a seguir com trabalhos desenvolvidos utilizando esses algoritmos.

• **Disciplina de Escalonamento *Priority Queuing* (PQ):** A disciplina de escalonamento PQ permite atribuir prioridades aos diferentes tráfegos gerados na rede. Após ser classificado pelo sistema, de acordo com sua classe de serviço, um determinado pacote é armazenado em uma fila correspondente ao seu nível de prioridade, como, por exemplo, filas de alta prioridade, de prioridade intermediária e baixa prioridade, conforme mostrados na Figura 21.



**Figura 21: Filas na disciplina de escalonamento PQ**  
**Fonte: Autoria própria.**

A disciplina PQ pode operar de duas formas: (i) *strict priority* onde os pacotes das filas de menor prioridade só serão transmitidos quando as filas de maior prioridade estiverem totalmente vazias; e, (ii) *rate-controlled* onde todos os pacotes da fila de maior prioridade só serão entregues antes das demais filas de menor prioridade se a quantidade de tráfego na fila de maior prioridade estiver abaixo de um patamar pré-determinado.

Quando existe a necessidade de garantia de atraso máximo fim-a-fim para um tipo específico de tráfego, a disciplina PQ se torna útil, pois promove um tratamento diferenciado para o escalonamento entre classes de serviço. Porém, pode provocar um bloqueio no tráfego de menor prioridade. Nas filas de maior prioridade, determinados fluxos podem sofrer com atrasos inadequados caso um dos fluxos armazenados nessa fila tenha um comportamento inadequado (FERREIRA, 2011).

No trabalho de Borin (2010) o escalonador utiliza a política de escalonamento PQ, priorizando os fluxos de dados em tempo real. Nele os fluxos de dados são classificados e agrupados em três filas de acordo com os requisitos de QoS. A fila de baixa prioridade armazena

<sup>5</sup> A disciplina de escalonamento decide a ordem pela qual os pacotes são servidos numa conexão.

as requisições de largura de banda do serviço BE. Na fila intermediária estão as requisições de largura de banda enviadas por conexões rtPS e nrtPS. Essas requisições podem migrar para a fila de alta prioridade no momento em que precisarem ser atendidas para garantir a provisão de QoS. A fila de alta prioridade armazena as concessões periódicas para o envio de dados e as concessões periódicas para o envio de requisição de largura de banda que devem ser atendidas no próximo quadro.

O trabalho desenvolvido em (FERREIRA, 2011), é uma extensão do trabalho em (BORIN, 2010), mas com quatro filas, separando as filas de atendimento dos tipos de tráfego de tempo real e não tempo real para aumentar o isolamento entre os tráfegos, e aumentar a eficiência do enlace *uplink*, pré-allocando os recursos restantes, após o escalonador atender as filas, diretamente para a transmissão de dados.

Em (JAKIMOSKI; JANEVSKI, 2013), o escalonador suporta cinco diferentes classes de serviço, a fim de oferecer suporte a diferentes necessidades dos usuários móveis com diferentes critérios de QoS. O esquema de escalonamento proposto é baseado no sentido do tráfego *uplink*, com alocação de prioridades para as classes de serviço, o que melhora o desempenho de QoS das cinco classes de serviço WiMAX, especialmente da classe de serviço ertPS. Experimentos de simulação e análise foram realizados com a ferramenta de simulação NS-2. A carga de tráfego das conexões ertPS é iniciada com um valor baixo e alterada passo-a-passo, até chegar em um valor máximo, a fim de realizar a análise de desempenho detalhada dos resultados. Os resultados apresentados no atraso médio, *jitter* médio e taxa de transferência média são avaliados e comparados com o mecanismo de escalonamento *uplink* desenvolvido por Borin (2010). Os resultados mostram que o regime de escalonamento proposto, com conexões ascendentes, com base na prioridade das classes de serviço, melhora o desempenho de QoS, especialmente em cenários com uma carga mais elevada.

É assumido, no modelo proposto por Jakimoski e Janevski (2013), que durante a implementação do esquema de escalonamento com prioridades, cada SS carrega fluxos de serviço únicos, para facilitar a execução do escalonador na BS. As prioridades das classes de serviço propostas nesta solução, para as conexões nas SSs, é mostrada na Tabela 4.

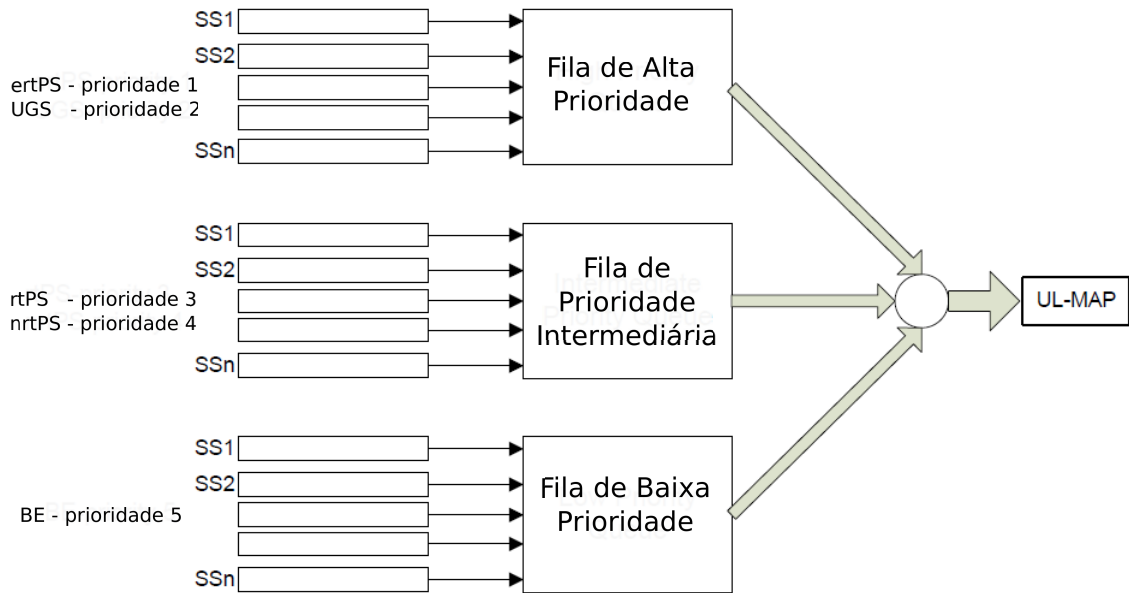
**Tabela 4:** Prioridades das Classes de Serviço WiMAX proposto por (JAKIMOSKI; JANEVSKI, 2013)

Fonte: Adaptada de (JAKIMOSKI; JANEVSKI, 2013).

Classes de Serviço WiMAX	Prioridade Atribuída
ertPS	1
UGS	2
rtPS	3
nrtPS	4
BE	5

De acordo com a proposta de (JAKIMOSKI; JANEVSKI, 2013), a classe de serviço

ertPS tem a prioridade mais alta e é servida em primeiro lugar. A solução proposta para o escalonamento da conexão ertPS ascendente nas SSs é implementada com base no mecanismo de simulação descrita no trabalho de (BORIN; FONSECA, 2009). A arquitetura do escalonador proposto é apresentado na Figura 22.



**Figura 22:** Arquitetura do escalonador *uplink* proposto por (JAKIMOSKI; JANEVSKI, 2013)

Fonte: Adaptada de (JAKIMOSKI; JANEVSKI, 2013).

O mecanismo de escalonamento de (JAKIMOSKI; JANEVSKI, 2013) utiliza três filas, de alta prioridade, prioridade intermediária e baixa e prioridade. A fila com baixa prioridade é usado para as solicitações de largura de banda da classe de serviço BE, a fila intermediária é reservada para armazenar os pedidos enviados por largura de banda das classes de serviço rtPS e nrtPS e a fila de alta prioridade é reservada para as classes de serviço UGS e ertPS. Solicitações de largura de banda enviadas por ambas as conexões rtPS e nrtPS podem migrar para a fila de alta prioridade para atender suas necessidades de QoS.

A diferença do trabalho de (JAKIMOSKI; JANEVSKI, 2013) em relação as prioridades utilizadas no mecanismo de escalonamento da nova arquitetura é em relação as classes de serviço UGS, que na nova arquitetura tem prioridade 1, e ertPS que na nova arquitetura tem prioridade 2, conforme pode ser visto na Figura 31 do Capítulo 4.

No Capítulo 5 é realizada uma análise de desempenho, através de simulações com o uso da ferramenta de simulação de redes NS-2, da nova arquitetura com o modelo proposto no trabalho de (JAKIMOSKI; JANEVSKI, 2013).

- **Disciplina de Escalonamento *Round Robin* (RR):** Na disciplina de escalonamento RR, os pacotes são rearranjados em classes, que são atendidas de maneira cíclica, sendo uma de cada

vez. Em um ciclo do escalonador RR é transmitido um pacote de cada fila. Caso uma fila esteja vazia, o escalonador passará para a próxima fila até que todas as filas fiquem vazias. Esta disciplina não atribui nenhum tipo de prioridade aos pacotes, porém acaba favorecendo os de tamanho maior (KUROSE; ROSS, 2010).

No trabalho de Cheng et al. (2010), uma avaliação do algoritmo RR é discutida. O algoritmo RR tem uma fácil implementação, pois distribui os recursos do canal de comunicação de maneira igualitária entre todas as SSSs, porém, quando existe diferentes tipos de tráfego e diferentes níveis de prioridade para este tráfego, o algoritmo de escalonamento RR não é prático. O trabalho de Sayenko et al. (2008) chegou a conclusão de que um algoritmo de escalonamento não deve perder muito tempo realizando o escalonamento e, em virtude disso, utiliza o algoritmo RR na solução proposta, compreendendo três etapas: (i) **Alocação de um número mínimo de segmentos (*slots*) de um quadro**. Nesta etapa, a BS calcula o número mínimo de segmentos para cada conexão para garantir os requisitos básicos de QoS. (ii) **Alocação dos segmentos não utilizados**. Nesta etapa, caso existam segmentos que não foram utilizados, a BS distribui esses segmentos entre as conexões rtPS, nrtPS e BE, a fim de evitar o comportamento *non-work-conserving*<sup>6</sup>. (iii) **Ordenação dos segmentos**. Nesta terceira etapa, a BS pode atribuir segmentos para vários tipos de rajadas de dados para assegurar que os requisitos de tempo sejam garantidos. Apesar das etapas descritas acima, este algoritmo não oferece garantias de atraso máximo limitado.

No trabalho de Dosciatti (2010), apesar de utilizar três diferentes tipos de tráfego, a disciplina de escalonamento utilizada é a RR. A alocação de cada usuário no sistema, se comporta tendo como parâmetro a duração do tempo de cada quadro. Uma quantidade  $n$  de usuários é alocada, sendo o usuário 1, alocado no período 1, o usuário 2, alocado no período 2, e assim por diante, até atender todos os usuários alocados no tamanho de duração do quadro.

- **Disciplina de Escalonamento *Weighted Round Robin* (WRR)**: Na disciplina WRR os pacotes são primeiramente classificados entre várias classes de serviço e então atribuídos a uma fila que é dedicada exclusivamente para esta classe de serviço. A cada fila é atribuído um peso que corresponde a uma porcentagem da largura de banda do canal. As filas são servidas em uma ordem que obedece a disciplina RR e caso a fila esteja vazia, o escalonador passará para a próxima fila. Nesta disciplina, pelo menos um pacote é removido de cada fila durante o ciclo de escalonamento independente do tamanho do pacote.

No trabalho em (CICCONETTI et al., 2007), o algoritmo WRR é avaliado para o

---

<sup>6</sup> Um pacote só pode ser transmitido quando este se tornar elegível, isto é, quando termina o tempo necessário para ele se manter em espera. Se apenas pacotes não elegíveis se encontram em espera, então o servidor manter-se-á inativo.



tráfego no sentido *uplink*, onde a largura de banda é distribuída para cada SS conforme o peso especificado e pode ser atribuído de modo a refletir a sua prioridade relativa, conforme a classe de tráfego que será transmitida. A conclusão do trabalho, realizada através de simulação, é de que, quando o tráfego for realizado com pacotes de tamanho variável, o escalonador não tem um desempenho satisfatório.

- **Disciplina de Escalonamento *Earliest Deadline First* (EDF):** A disciplina EDF é muito utilizada para o escalonamento do tráfego de tempo real. Nela, atribui-se um prazo de validade (*deadline*) a cada pacote e o escalonador serve os pacotes pela ordem dos seus prazos de validade. Assim, um pacote, ao qual é atribuído um prazo mais próximo do instante de chegada, sofre um atraso menor na fila de espera do que outro ao qual tenha sido atribuído um prazo mais distante. Dependendo da carga, pode não ser possível servir todos os pacotes antes de se atingir os respectivos prazos que lhes foram atribuídos.

Em (TSAI et al., 2006) é apresentada uma proposta de escalonamento no sentido *uplink* utilizando a disciplina de escalonamento EDF, trabalhando em conjunto com um algoritmo de controle de admissão de conexão baseado em *token bucket*. Também em (RUANGCHAIJATUPON et al., 2006) é avaliado o desempenho do algoritmo EDF. Um algoritmo EDF é do tipo *work-conserving* e foi proposto originalmente para aplicações em tempo real em redes de longa distância (FERRARI; VERMA, 1990).

- **Disciplina de Escalonamento *Weighted Fair Queuing* (WFQ):** A disciplina WFQ oferece uma distribuição justa de largura de banda para pacotes de tamanho variável. O WFQ tem seu comportamento similar a um escalonador RR *bit-a-bit* com pesos. Assim, cada *bit* dos pacotes na saída de cada fila são transmitidos separadamente, obedecendo à disciplina WRR, que leva em conta o tamanho do pacote. No WFQ, os pacotes são primeiramente classificados em fluxos pelo sistema e então atribuídos a filas que são especificamente dedicadas para estes fluxos. Dados a taxa de bits da porta de saída, o número de filas ativas, o peso relativo atribuído a cada fila e o tamanho de cada um dos pacotes em cada fila é possível calcular e atribuir o tempo final em cada um dos pacotes que chegarem. Este tempo final corresponde ao tempo em que esse pacote deveria deixar a fila. O escalonador então seleciona e encaminha o pacote com o menor tempo final entre todos os pacotes enfileirados de todas as filas.

O algoritmo WFQ é avaliado em (RUANGCHAIJATUPON et al., 2006) e, quando utilizado com pacotes de tamanho variável, tem um bom desempenho, porém, a sua complexidade de implementação, pelo fato de que o tempo final do serviço de cada pacote deve ser calculado, torna-se uma desvantagem em relação aos outros algoritmos.

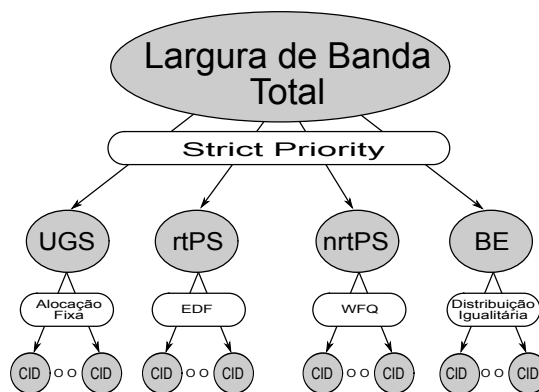
- **Disciplina de Escalonamento *Latency-Rate* (LR):** O escalonador *LR* permite calcular um

atraso máximo de cada sessão individual (STILIADIS; VARMA, 1998). Neste trabalho é utilizada esta disciplina, juntamente com a PQ, para que seja possível calcular os limites justos para o atraso fim-a-fim de sessões individuais suportando uma rede arbitrária de escalonadores. Mais detalhes da utilização da disciplina de escalonamento *LR* são abordados na Seção 4.2.1.

### 3.3.2.2 ALGORITMOS DE ESCALONAMENTO HÍBRIDOS

Nos algoritmos de escalonamento híbridos são empregadas mais de uma das disciplinas de escalonamento descritas na Seção 3.3.2.1, conforme a classe de serviço do tráfego e os requisitos de QoS solicitados.

Em (WONGTHAVARAWAT; GANZ, 2003) é proposto um algoritmo de escalonamento hierárquico que combina as disciplinas de escalonamento *Strict Priority*, EDF e WFQ. Neste trabalho, a alocação de banda é realizada em duas fases. Na primeira fase, a disciplina *Strict Priority* é utilizada e a largura de banda é distribuída entre as diferentes classes de serviço. Na segunda fase, em cada SS, a banda recebida é distribuída entre as conexões, sendo que a disciplina EDF é utilizada para escalonar o tráfego de tempo real (rtPS). Para o tráfego que não é de tempo real (nrtPS) as SSs utilizam a disciplina de escalonamento WFQ e, utilizam a distribuição igualitária para escalonar o tráfego da classe de serviço com melhor esforço (BE). As conexões com tráfego UGS recebem alocação de banda fixa. Uma desvantagem da disciplina *Strict Priority* é que as conexões de maior prioridade podem utilizar toda a largura de banda e as conexões de prioridade mais baixa podem não ser atendidas, também não vai garantir o atraso máximo. O algoritmo proposto é avaliado através de simulação, mas apenas para os tráfegos rtPS e BE. A estrutura hierárquica da alocação de banda é mostrada na Figura 23.



**Figura 23: Estrutura hierárquica de alocação de banda**  
**Fonte: (WONGTHAVARAWAT; GANZ, 2003).**

Em (VINAY et al., 2006), as SSs utilizam a combinação dos algoritmos de

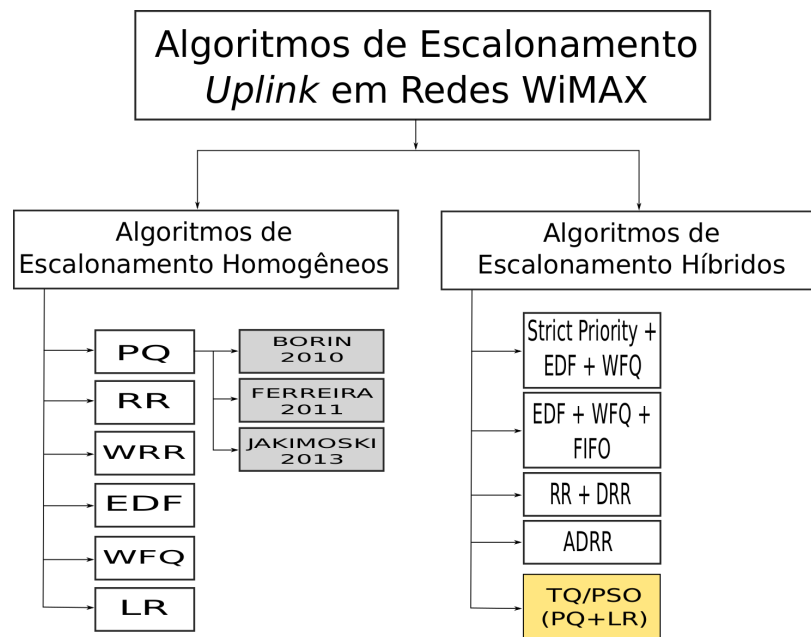
escalonamento EDF, WFQ e FIFO (*First In-First Out*). O mecanismo de alocação de banda é idêntico ao trabalho em (WONGTHAVARAWAT; GANZ, 2003), a diferença é a utilização da disciplina de escalonamento FIFO para as conexões com tráfego BE.

Em (SETTEMBRE et al., 2006) um algoritmo de escalonamento híbrido foi desenvolvido utilizando os algoritmos RR e WRR com um mecanismo de *strict priority* para a alocação da largura de banda total. A escolha das disciplinas WRR e RR é eficiente no que diz respeito à complexidade computacional e aplicação, permitindo para o escalonador superar as limitações em termos de equidade quando os pacotes de comprimento variável são escalonados. Nos estágios iniciais do algoritmo, um rigoroso esquema de prioridades é utilizado para alocar largura de banda, entre as SSs. O algoritmo WRR é responsável por alocar a largura de banda entre as classes de serviço rtPS e nrtPS até que elas sejam satisfeitas. A largura de banda residual é distribuída entre as SSs para as classes de serviço BE, utilizando o algoritmo RR. Para as conexões com tráfego UGS a alocação de banda é fixa.

No trabalho de Alshag et al. (2011) um algoritmo foi desenvolvido para ser utilizado na primeira fase do modelo de escalonamento desenvolvido em (WONGTHAVARAWAT; GANZ, 2003). O algoritmo é nominado por ADRR (*Adaptiv Deficit Round Robin*) e é baseado no algoritmo de escalonamento DRR. Neste trabalho a classe de serviço ertPS (*extended-real-time Polling Service*) é inserida na estrutura hierárquica. Na segunda etapa, para as conexões com serviço ertPS e rtPS, o algoritmo de escalonamento EDF é utilizado, o algoritmo de escalonamento WFQ é utilizado para a classe de serviço nrtPS e para a classe de serviço BE o algoritmo de escalonamento RR é utilizado.

A Figura 24 apresenta graficamente a estrutura dos algoritmos de escalonamento. Na disciplina PQ, os trabalhos em (BORIN, 2010), (FERREIRA, 2011) e (JAKIMOSKI; JANEVSKI, 2013), em destaque, estão relacionados com esta pesquisa, pois especificam uma tentativa de garantir o atraso e são os que mais se aproximam com o modelo desta nova arquitetura, mesmo que utilizem o tempo do quadro fixo, mas possuem filas de escalonamento com prioridades.

Nos algoritmos de escalonamento híbridos, a disciplina *TQ/PSO*, em destaque na Figura 24, foi desenvolvida e utilizada na arquitetura deste trabalho. Esta nova disciplina se utiliza das disciplinas de escalonamento homogêneas, *PQ* e *LR*. A disciplina *PQ* vai garantir prioridades para as classes de serviço, tratados na Camada 3 da nova arquitetura, que realiza o controle de admissão de conexões e está descrito na Seção 4.4. A disciplina *LR* vai garantir o atraso máximo requisitado pelos usuários para as classes de serviço que são sensíveis ao atraso (UGS, ertPS e rtPS). Além disso, a nova disciplina *TQ/PSO* acrescenta a garantia de uma banda



**Figura 24:** Algoritmos de escalonamento *uplink* para redes WiMAX  
 Fonte: Autoria própria.

mínima para as classes de serviço que não são sensíveis ao atraso (nrtPS e BE).

### 3.4 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou uma descrição da arquitetura de uma rede padrão IEEE 802.16, detalhando o modelo de referência que o padrão especifica para as camadas PHY e MAC, com suas principais características, os mecanismos de provisão de QoS e escalonamento.

Dentro da camada PHY são definidas várias características do sistema, dentre elas, as faixas de frequência suportadas, as especificações da interface física, a modulação utilizada pelas redes WiMAX fixas e como consiste o seu funcionamento. Também detalha como é realizada a transmissão das informações dentro dos quadros e subquadros utilizados para o tráfego dos dados entre os usuários da rede. A camada PHY é baseada na modulação OFDM, que é uma técnica eficaz para contornar o efeito de multipercurso, tendo técnicas avançadas para aumentar a confiabilidade do canal. Já na camada MAC, existe uma flexibilidade para acomodar uma variedade de tipos de tráfego, incluindo voz, vídeo e multimídia, com garantia de QoS.

Um aspecto importante deste capítulo foi a apresentação da maneira como as redes WiMAX garantem QoS. Os principais conceitos envolvendo a QoS foram apresentados e a sua importância para que as redes de computadores sejam eficientes e escalonáveis para dados, vídeo e voz. Assim, para provisão de QoS, alguns parâmetros como, atraso, vazão e variação do

atraso devem ser definidos através de requisitos do sistema. E, para que esses requisitos possam ser gerenciados pelas redes WiMAX, existem mecanismos que devem garantir a provisão de QoS dentro das solicitações dos usuários. O provimento de QoS é de grande importância para as redes de computadores, quando se utiliza aplicações em tempo real, pois elas demandam largura de banda e serviços diferenciados. Existem aplicações em que a transmissão de dados deve ser realizada sem interrupção ou perda de pacotes.

Também foi destacado que a QoS é implementada por processos de solicitação e concessão de largura de banda, que consiste em associar pacotes que passam pela camada MAC a um fluxo de serviço que fornece um transporte unidirecional aos dados. Essa associação fornece um tratamento diferenciado aos tráfegos que são gerados pelas aplicações. Assim, foi descrito que as classes de serviços representam um conjunto de parâmetros para a entrega de dados em um determinado fluxo de dados, tendo, como finalidade, o compartilhamento de largura de banda entre os diferentes tipos de usuários. Desta forma, nas redes IEEE 802.16, o serviço de escalonamento representa o mecanismo de transporte de dados em uma determinada conexão. Cada conexão é associada a somente um tipo de serviço. Cada serviço é associado a um conjunto de parâmetros de QoS que define o comportamento da conexão.

A fim de fornecer suporte a QoS eficiente para os usuários finais, o padrão especifica que os mecanismos de controle de admissão de conexão, de escalonamento de pacotes e de policiamento de tráfego formam o conjunto de componentes necessários para a garantia de QoS, porém, não são especificados pela norma e são fontes de pesquisa por parte da indústria e do universo acadêmico. Esses mecanismos funcionam através dos requisitos informados pelo conjunto de tráfego das cinco classes de serviços que compõem a arquitetura de QoS (UGS, rtPS, ertPS, nrtPS e BE).

O algoritmo de escalonamento é o principal mecanismo responsável pela obtenção da QoS nas redes IEEE 802.16. Assim sendo, o desempenho das aplicações, principalmente aquelas de tempo real, depende diretamente da disciplina de escalonamento utilizada. Este capítulo discutiu o escalonamento em redes WiMAX, regidas pelo padrão IEEE 802.16, definindo como os mecanismos de escalonamento devem ser utilizados, porém nenhum tipo de mecanismo é definido pelo padrão.

Na Seção 3.3.2, os principais algoritmos de escalonamento foram listados, seguindo a classificação por tipo e desenvolvidos segundo as características definidas pelo padrão. Atualmente, várias pesquisas vêm sendo desenvolvidas sobre escalonamento em redes IEEE 802.16. Entretanto, estudos mostram que o desenvolvimento de um escalonador justo e eficiente é uma área de pesquisa em desenvolvimento (MSADAA et al., 2010).

#### 4 UMA NOVA ARQUITETURA PARA PROVISÃO DE QoS EM REDES WIMAX FIXAS

Dos tipos de escalonamento definidos pelo padrão IEEE 802.16 (IEEE, 2012), o escalonamento no sentido *uplink*, localizado na BS, é o mais complexo de ser implementado. Como as SSs transmitem dados com perfis de rajadas diferentes e o algoritmo de escalonamento deve garantir QoS para todas as requisições das SSs, a função do algoritmo de escalonamento é fornecer QoS para todas as SSs, que podem estar localizadas em diferentes distâncias da BS. Assim, a implementação de um escalonador deve utilizar a largura de banda disponível de maneira eficiente, possibilitando que uma quantidade maior de usuários sejam alocados na rede, aumentando o nível de utilização do sistema.

Na nova arquitetura proposta, um escalonador no sentido *uplink*, na BS, é executado, caso for necessário, a cada quadro (*frame*), com o objetivo de otimizar a utilização da rede através do cálculo da duração do tempo do quadro ( $TQ$ ) para as classes de serviço sensíveis ao atraso. Com isso, existindo a necessidade, conforme os parâmetros da rede utilizados naquele momento, um novo cálculo da duração do  $TQ$  é realizado, possibilitando que mais usuários sejam alocados no sistema, tornando a rede mais dinâmica e mais eficiente.

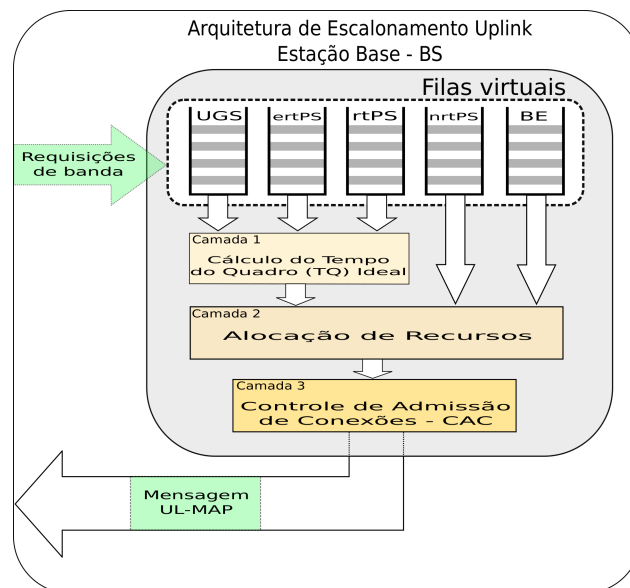
Este capítulo apresenta uma nova arquitetura para garantia de QoS com a utilização da otimização por enxame de partículas (PSO) em redes WiMAX fixas, atuando na BS, no escalonamento do tráfego no sentido *uplink*, baseado na modulação OFDM, utilizando a topologia PMP e o esquema de transmissão TDD.

O restante deste capítulo apresenta, na Seção 4.1, o modelo da nova arquitetura, enfatizando o objetivo de prover QoS para as cinco classes de serviços definidas pelo padrão IEEE 802.16, com garantia do limite de atraso para as sensíveis ao atraso. A nova arquitetura está descrita na Seção 4.1, demonstrando que uma importante modificação é introduzida, através do cálculo da duração do  $TQ$  em cada iteração do escalonador, utilizando a metaheurística de otimização PSO, para as classes de serviço sensíveis ao atraso. As Seções 4.2, 4.3 e 4.4 apresentam as camadas de desenvolvimento do algoritmo de escalonamento da nova arquitetura. A Seção 4.5 apresenta os algoritmos que fazem a composição para o cálculo da duração do  $TQ$ .

#### 4.1 MODELO DA NOVA ARQUITETURA

Um algoritmo de escalonamento deve possuir uma baixa complexidade computacional e, ao mesmo tempo, proteger a rede contra fluxos com comportamento inadequado, fornecendo limites para que a vazão e o atraso tenham um desempenho satisfatório e não comprometam os parâmetros de QoS da rede. Baseado nisto, este trabalho de pesquisa tem o objetivo de apresentar uma nova arquitetura para garantir QoS para as cinco classes de serviços definidas pelo padrão IEEE 802.16, garantindo o limite de atraso para as classes de serviços sensíveis ao atraso e oferecendo uma banda mínima para os outros tipos de tráfego.

O modelo da nova arquitetura é apresentado nas Figuras 25 e 26. As requisições de banda, definidas através das mensagens de requisição de largura de banda, enviadas pelas SSs, são recebidas do subquadro *uplink* e armazenadas em filas virtuais na BS, de acordo com as cinco classes de serviço especificadas pelo padrão IEEE802.16. Assim, a BS deve estar sincronizada com várias SSs para gerenciar adequadamente as operações de requisição e alocação de largura de banda.



**Figura 25: Modelo da nova arquitetura - Estação Base - BS**

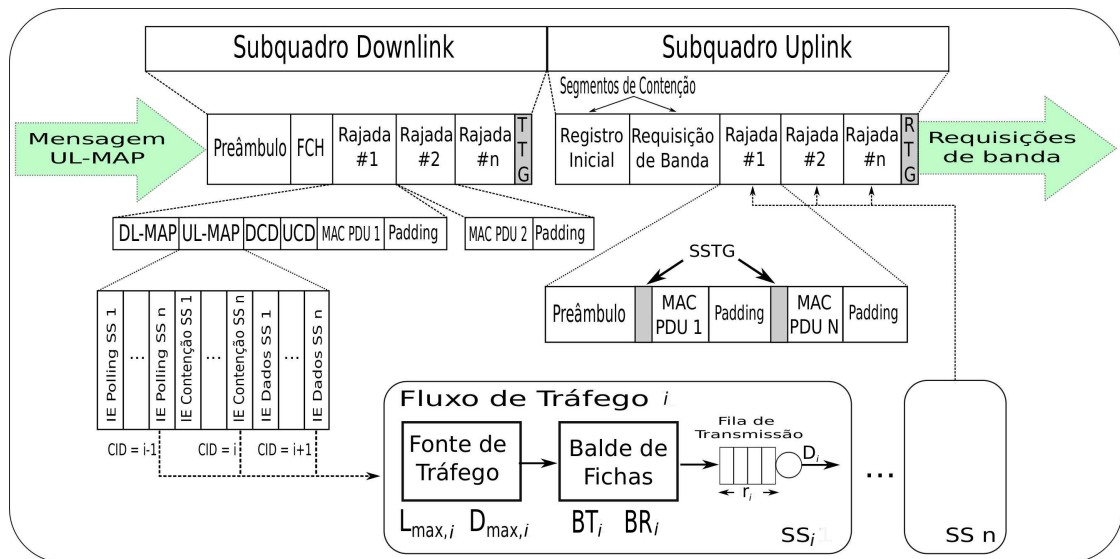
Fonte: Autoria própria.

A nova arquitetura vai utilizar uma disciplina de escalonamento híbrida, denominada de *TQ/PSO*, composta pelas disciplinas de escalonamento *PQ* e *LR*. A disciplina *PQ* vai estabelecer prioridades para o atendimento das requisições para as 5 classes de serviço e a disciplina *LR* vai garantir o atraso máximo solicitado pelo usuário, nas classes de serviço UGS, ertPS e rtPS, e a requisição de taxa mínima, nas classes de serviço nrtPS e BE. Esse modelo da nova arquitetura é genérico, pois, dependendo do atraso solicitado, o cálculo do *TQ* será

efetuado ou não, conforme os parâmetros de QoS que foram especificados na Seção 3.2.1.

A estratégia utilizada para o desenvolvimento da nova arquitetura é uma abordagem hierárquica, com três camadas, denominadas de **Camada 1 - Cálculo da Duração do Tempo do Quadro Ideal ( $TQ$ )**, **Camada 2 - Alocação de Recursos** e **Camada 3 - Controle de Admissão de Conexões (CAC)**, conforme pode ser observado na Figura 25. A execução das camadas é realizada em ordem sequencial, dependendo do tipo de tráfego que está sendo escalonado, das funcionalidades e com o objetivo de cada camada, tendo a possibilidade de modificações de parâmetros de forma independente em cada camada. Os detalhes do funcionamento das três camadas são discutidas nas Seções 4.2, 4.3 e 4.4.

Após a execução das três camadas, as necessidades de transmissão de cada uma das classes são organizadas para atendimento nos próximos subquadros, fornecendo uma mensagem, denominada de *UL – MAP*, que especifica as regiões de dados alocadas para cada estação no próximo subquadro *uplink*, localizada no subquadro de *downlink*, conforme mostrado na Figura 26).



**Figura 26: Modelo da nova arquitetura - Subquadros *Downlink* e *Uplink***  
Fonte: Autoria própria.

A mensagem *UL – MAP* é formada por um conjunto de informações denominado como IE (*Information Element*), que se refere às regiões de dados dos subquadros de *downlink* e de *uplink*. Cada IE contém a localização, duração, tipo e configuração AMC (*Adaptive Modulation and Coding*) de uma rajada. As IEs são classificadas em: (i) **polling**, que contém oportunidades de requisição de banda concedidas por *polling*; (ii) **contenção**, que contém oportunidades de requisição de banda concedidas em resposta a códigos informados no segmento de contenção do registro inicial de quadros anteriores; e (iii) **dados**, que contém



oportunidades de transmissão de dados concedidas às SSs.

Cada SS tem a responsabilidade de tratar as IEs destinados ao seu código de identificação (*CID*), fornecendo os dados para transmissão ou uma requisição de banda, de acordo com o tipo de oportunidade recebida. Assim, o próximo subquadro UL é formado por informações fornecidas pelas SSs, transmitidas nas oportunidades indicadas pela UL-MAP. Esse processo é cíclico e é repetido para a geração de cada quadro.

## 4.2 CAMADA 1 - CÁLCULO DA DURAÇÃO DO TEMPO DO QUADRO IDEAL

A nova arquitetura segue uma tendência das soluções mais recentes, como os trabalhos de Sayenko et al. (2008), Masri et al. (2009), Borin (2010), Ferreira (2011) e Jakimoski e Janevski (2013), para o escalonamento que atuam na BS, no tráfego no sentido *uplink*. A diferença entre a nova arquitetura em relação a estes trabalhos, descritos no Capítulo 3, é que a nova arquitetura implementa o cálculo para encontrar um valor ideal para a duração do tempo do quadro (*TQ*).

O *TQ* Ideal é calculado com a utilização da metaheurística PSO, para garantir o atraso ao tráfego das classes de serviço UGS, ertPS e rtPS, também garantindo a banda mínima para as classes de serviço nrtPS e BE. Com o cálculo da duração do *TQ* Ideal, o sistema é otimizado pelo fato de que mais usuários podem transmitir seus dados dentro de um mesmo quadro. Esse cálculo também evita que os quadros sejam transmitidos com segmentos físicos vazios, ou seja, segmentos que não foram alocados para nenhuma SS.

A seguir, o modelo analítico para o cálculo do *TQ* é discutido e as equações de restrição, para tornar o escalonamento mais eficiente, são apresentadas. Para isso, a Tabela 5 descreve os símbolos, com suas unidades de medida, utilizados na modelagem da nova arquitetura.

### 4.2.1 MODELO ANALÍTICO PARA O CÁLCULO DO TEMPO DO QUADRO (*TQ*)

A garantia de QoS em uma rede de pacotes requer o uso de algoritmos de escalonamento de tráfego, em cada roteador, para selecionar o pacote a ser transmitido. Como é improvável que as redes de computadores sejam homogêneas em relação ao tipo de algoritmo de escalonamento utilizado por um roteador individual, um modelo geral para a análise de algoritmos de escalonamento torna-se um instrumento valioso no projeto e análise dessas redes. Assim, Stiliadis e Varma (1998) desenvolveram um modelo, denominado de *Latency-Rate (LR)*, que tem como objetivo estudar o comportamento do pior caso de sessões individuais em uma

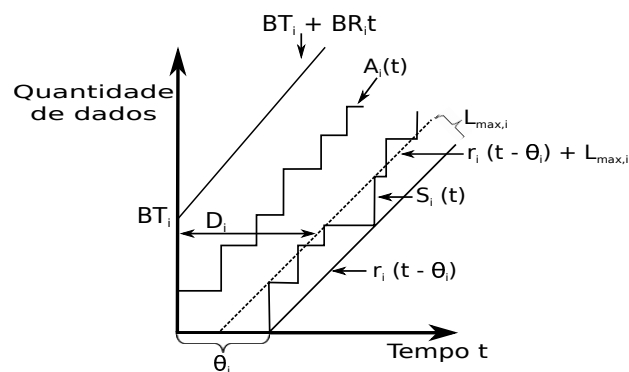
Tabela 5: Símbolos com suas unidades de medida para a nova arquitetura

Fonte: Autoria própria.

Símbolo	Grandeza ou descrição	Unidade de Medida
$i$	índice da sessão	inteiro
$t$	tempo	$s$
$\theta_i$	latência	$ms$
$r_i$	taxa alocada pelo servidor	$bps$
$S_i$	taxa de serviço	$bps$
$A_i$	tráfego de entrada	$bps$
$D_i$	atraso máximo na fila	$ms$
$D_{max,i}$	atraso máximo requerido pelo usuário	$ms$
$D_{bound}$	limite de atraso superior	$ms$
$BT_i$	tamanho do balde de fichas	$bits$
$BR_i$	taxa do balde de fichas	$bps$
$L_{max,i}$	tamanho máximo de um pacote	$bits$
$TQ$	tempo do quadro	$ms$
$TQ_{min}$	tempo do quadro mínimo	$ms$
$TQ_{max}$	tempo do quadro máximo	$ms$
$TU_i$	tempo que cada SS tem para transmitir seus pacotes de dados	$ms$
$R$	taxa física do canal	$Mbps$
$T_{TTG}$	intervalo de tempo entre o subquadro de <i>downlink</i> e <i>uplink</i>	$\mu s$
$T_{RTG}$	intervalo de tempo entre o subquadro de <i>uplink</i> e <i>downlink</i>	$\mu s$
$RI_{inic}$	segmento de registro inicial	$\mu s$
$BW_{request}$	segmento de requisição de largura de banda	$\mu s$

rede de escalonadores que utiliza diferentes algoritmos de escalonamento, permitindo calcular um atraso, para cada sessão individual, em uma rede com vários tipos de escalonadores.

O modelo de funcionamento de um escalonador, baseado no servidor *LR*, apresentado na Figura 27, é determinado por dois parâmetros: a latência  $\theta_i$  e a taxa alocada pelo servidor  $r_i$ . A latência de um servidor *LR* é o pior caso de atraso  $D_i$  medido pelo primeiro pacote do período ocupado em uma sessão, que é um pacote chegando quando a fila da sessão está vazia. O atraso  $D_i$  gerado, é limitado pela curva do atraso do balde de fichas ( $BT_i + BR_i t$ ) e pela curva do atraso do escalonador *LR* ( $r_i (t - \theta_i)$ ), que é dependente do valor de  $\theta_i$  e da taxa alocada  $r_i$ .

Figura 27: Modelo de um escalonador *LR*

Fonte: Figura adaptada de (STILIADIS; VARMA, 1998)

Desta forma, a latência de um algoritmo de escalonamento, depende de seus parâmetros internos, da sua taxa de transmissão no enlace de saída e das taxas alocadas nas diversas sessões, sendo que, para cada arquitetura de rede (MPLS, ATM, FDDI, IEEE 802.11, LTE, WiMAX), o modelo de cálculo da latência é diferente. No entanto, o máximo atraso fim-a-fim experimentado por um pacote, em uma rede de escalonadores, pode ser calculado somente pelas latências dos escalonadores individuais no caminho da sessão, e pelos parâmetros de tráfego da sessão que gerou o pacote.

O atraso máximo solicitado por um escalonador aumenta proporcionalmente a sua latência. Por este fato, o modelo de escalonamento *LR* demonstra a importância do uso de escalonadores de baixa latência para alcançar um atraso fim-a-fim baixo.

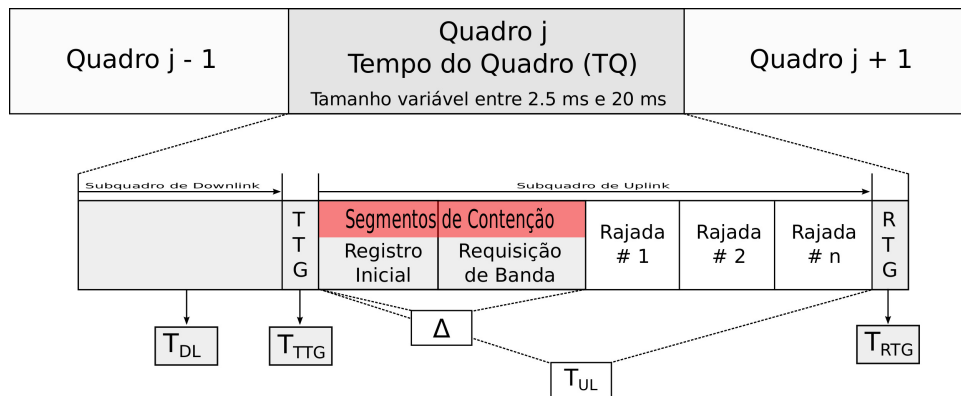
Na nova arquitetura, uma modificação no escalonador *LR* (STILIADIS; VARMA, 1998) é realizada, pois o escalonador *LR* só garante o atraso, já a nova arquitetura garante o atraso e a banda mínima para as classes de serviço que a necessitam. Assim, a nova arquitetura garante o atraso ao tráfego das classes de serviço UGS, ertPS e rtPS, e também garante a banda mínima para as classes de serviço nrtPS e BE. Isso é um fator de extrema importância para as redes, pois o seu tráfego vai necessitar de limites de atraso máximo e também de garantia de banda mínima.

Para controlar a taxa de pacotes que são injetados na rede, ou seja, para modelar o tráfego de entrada na rede, a nova arquitetura utiliza o algoritmo do balde de fichas (*token bucket*) (TSAI et al., 2006) para as classes de serviço sensíveis ao atraso. A ideia é que o balde de fichas estabeleça um limite para a entrada de todo o tráfego da rede e o escalonador *LR* estabeleça uma taxa de alocação para cada usuário. Desta forma, se a taxa alocada pelo escalonador *LR* é maior que a taxa do balde de fichas, um atraso máximo  $D_i$  é calculado.

O comportamento da nova arquitetura, é determinado por dois parâmetros para cada sessão  $i$  ( $i = 1, \dots, N$ ): latência  $\theta_i$  e taxa alocada pelo servidor  $r_i$ . Para cada tipo de arquitetura de rede, a latência tem um modelo diferente, assim, nas redes WiMAX, padrão IEEE 802.16, para cada sessão  $i$ , a latência é um período do tempo dos subquadros de *downlink* e *uplink* ( $T_{DL} + T_{UL}$ ), do tempo para transmitir um pacote de tamanho máximo ( $\frac{L_{max,i}}{R}$ ) e dos parâmetros dos intervalos de separação dos subquadros de *downlink* e *uplink*,  $T_{TTG}$  e  $T_{RTG}$ , e é expressa por

$$\theta_i = T_{DL} + T_{UL} + \frac{L_{max,i}}{R} + T_{TTG} + T_{RTG} . \quad (3)$$

A alocação dos parâmetros na estrutura de um quadro TDD é apresentado na Figura 28.



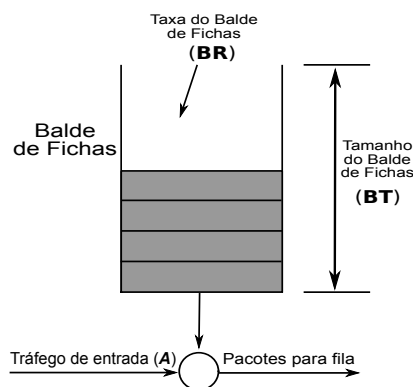
**Figura 28: Alocação dos parâmetros na estrutura de um quadro TDD**

Fonte: Autoria própria.

Como um escalonador *LR* pode prover um limite de atraso se o tráfego de entrada for modelado por um balde de fichas, o modelo apresentado pela nova arquitetura, considera, para cada pacote, o cabeçalho determinado pelo padrão IEEE 802.16. Dessa forma, a quantidade de fichas do balde é diminuída pelo tamanho do pacote e pelo seu cabeçalho.

Quando o tráfego de entrada for pertencente às classes de serviço *ertPS* e *rtPS*, a aplicação declara o tamanho máximo do pacote  $L_{max,i}$  e requisita um atraso máximo que é permitido, representado por  $D_{max,i}$ , e que é utilizado pelo escalonador na nova arquitetura para calcular a taxa de serviço de cada sessão para garantir o atraso requisitado e otimizar o número de estações na rede. Assim, o tráfego de entrada  $A_i$ , em um determinado período de tempo  $t$ , da sessão  $i$  ( $i = 1, \dots, N$ ), passa pelo balde de fichas de uma SS durante o intervalo  $(0, t)$ , conforme mostrado na Figura 29. Essa passagem do tráfego de entrada pelo balde de fichas é limitada pelo tamanho do balde de fichas  $BT_i$  e pela taxa do balde de fichas  $BR_i$  no período  $t$ , expressa por

$$A_i(t) \leq BT_i + BR_i t . \quad (4)$$



**Figura 29: Tráfego de entrada modelado por um balde de fichas**

Fonte: Autoria própria.

O pacote fica na fila da estação, até que ele acesse o meio sem fio e transmita. O atraso na fila é medido desde quando o último bit do pacote é recebido e enfileirado na estação até que ele acesse o meio sem fio e transmita. Na nova arquitetura, o atraso na fila depende dos parâmetros do balde de fichas, da latência da rede e da taxa alocada pelo servidor. Desta forma, como o tráfego de entrada  $A_i(t)$  é modelado por um balde de fichas e o escalonador aloca uma taxa do servidor  $r_i$ , um atraso máximo  $D_i$  pode ser provido por

$$D_i \leq \frac{BT_i}{r_i} + \theta_i - \frac{L_{max,i}}{r_i}, \quad (5)$$

onde  $\frac{L_{max,i}}{r_i}$  é a diferença entre a inclinação superior e inferior na Figura 27, ou seja, a menor inclinação do servidor ( $r_i(t - \theta_i)$ ) é o menor limite válido para o serviço em qualquer ponto do período ocupado e a maior inclinação do servidor ( $BT_i + BR_i t$ ) é válida no ponto em que um pacote sai do sistema. Dessa forma, o atraso pode ser calculado baseado nesta inclinação superior, cujo resultado é um limite de atraso máximo na rede, denominado de limite de atraso superior  $D_{bound}$ , que é representado por  $\frac{BT_i}{r_i} + \theta_i - \frac{L_{max,i}}{r_i}$ . Esse limite de atraso superior deve ser menor ou igual ao atraso máximo que é requerido pelo usuário e que pode ser admitido pelo sistema e é definido por

$$\frac{BT_i}{r_i} + \theta_i - \frac{L_{max,i}}{r_i} \leq D_{max,i}. \quad (6)$$

Desta forma, existem três diferentes tipos de atraso definidos no trabalho. O primeiro é o atraso máximo  $D_i$ , o segundo é o limite de atraso superior  $D_{bound}$  e o terceiro é o atraso máximo que é requerido pelo usuário e que pode ser admitido pelo sistema  $D_{max,i}$ . A relação entre eles é  $D_i \leq D_{bound} \leq D_{max,i}$ .

Para cada usuário que entra no sistema e vai fazer a transmissão de seus dados, uma sobrecarga (*overhead*) deve ser considerada. Na nova arquitetura, essa sobrecarga é composta pelos segmentos de contenção no subquadro de *uplink* (ver Figura 28) e expresso por

$$\Delta = RI_{inic} + BW_{request}, \quad (7)$$

onde  $RI_{inic}$  é reservado para manutenção inicial, oportunidade na qual podem ser atendidas requisições de ingresso na rede, bem como de alteração do perfil de rajada. Após isso, com  $BW_{request}$  se inicia o período no qual as SSs podem realizar solicitações de banda para transmissão no próximo subquadro de *uplink*.

Conforme a *Camada 1* desta nova arquitetura (veja Figura 25), o *TQ Ideal* será calculado se houver alterações nos parâmetros da rede, ou seja, quando novas SSs requisitarem acesso para a BS, se SSs deixarem a rede, ou se uma SS trocar seus parâmetros de QoS, mas

somente quando, nas filas virtuais, existir tráfego para as classes de serviço UGS, ertPS e rtPS.

A taxa alocada pelo servidor para escalonar cada usuário é definida como sendo  $r_i$ . Porém, levando em consideração a sobrecarga  $\Delta$  em (7), deve-se acrescentar este valor para a transmissão de cada usuário, determinado por

$$r'_i = r_i + \frac{\Delta * R}{TQ}. \quad (8)$$

Então, de (3), (5) e (8), o atraso máximo  $D_i$ , na nova arquitetura, é limitado por

$$D_i \leq \frac{(BT'_i - L'_{max,i}) * TQ}{r'_i * TQ - \Delta * R + L'_{max,i}} + T_{DL} + T_{UL} + \frac{L'_{max,i}}{R} + T_{TTG} + T_{RTG}, \quad (9)$$

onde  $L'_{max,i}$  é  $L_{max,i}$  com a sobrecarga de (7) e  $BT'_i$  é o tamanho do balde de fichas com a mesma sobrecarga. Como obter (9) é demonstrado no Apêndice A.

Com isso, a primeira condição de restrição para o cálculo do  $TQ$  Ideal, relacionada ao atraso, é definida por

$$\frac{(BT'_i - L'_{max,i}) * TQ}{r'_i * TQ - \Delta * R + L'_{max,i}} + T_{DL} + T_{UL} + \frac{L'_{max,i}}{R} + T_{TTG} + T_{RTG} \leq D_{max,i}, \quad (10)$$

onde  $R$ ,  $L'_{max,i}$  e  $BT'_i$  são parâmetros declarados pela aplicação. Entretanto,  $TQ$  e  $r'_i$  são parâmetros que devem ser calculados para satisfazer (10).

A Equação (11) é a segunda condição de restrição para que a nova arquitetura possa efetuar o cálculo da duração do  $TQ$  Ideal e da taxa alocada pelo servidor. A taxa do balde de fichas, mais a taxa para transmitir a sobrecarga e o tamanho máximo de um pacote, devem ser menor do que o total da taxa alocada pelo servidor. Então, a segunda condição de restrição é obtida por

$$BR_i + \frac{\Delta * R + L'_{max,i}}{TQ} \leq r'_i. \quad (11)$$

Os trabalhos abordados na Seção 3.3.2, que tratam dos algoritmos de escalonamento, não especificam qualquer tipo de mecanismo para calcular o valor da duração do tempo de um quadro. Cada escalonador requisita um tempo de quadro que será fixo durante todo o procedimento de escalonamento. Porém, a duração de um  $TQ$  é importante, pois existe um equilíbrio entre um pequeno e um grande valor para o  $TQ$ . Um quadro com uma pequena duração reduz o atraso máximo, no entanto, aumenta o valor do cabeçalho. Por outro lado, um quadro com duração maior, tem a tendência de diminuir o valor do cabeçalho, mas o atraso máximo se eleva. A nova arquitetura calcula um valor para a duração do  $TQ$  que seja ideal,

alocando um número máximo de SSs sob as condições de restrição apresentadas nas Equações (10) e (11).

Assim, o número máximo de SSs é alcançado quando o total da taxa alocada pelo servidor  $r'_i$ , para cada SS com tráfego UGS, ertPS ou rtPS, for a mínima para garantir o  $D_{bound}$ . Esse valor é encontrado por (12).

$$f(TQ) = \frac{\sum_{i=1}^n TU_i}{TQ} \rightarrow \min, \quad (12)$$

onde  $TU_i$  é o tempo que cada SS tem para transmitir os pacotes de dados e é calculado pela Equação (13). Neste cálculo, o valor da sobrecarga  $\Delta$  não é considerado, pois somente a carga de dados é o que interessa para o sistema.

$$TU_i = \frac{r'_i * TQ}{R} - \Delta \quad (13)$$

Os valores de  $TQ$  e  $TU_i$  são calculados novamente, quando: (i) novas SSs requisitarem acesso para a BS; (ii) se SSs deixarem a rede; ou (iii) se uma SS trocar seus parâmetros de QoS.

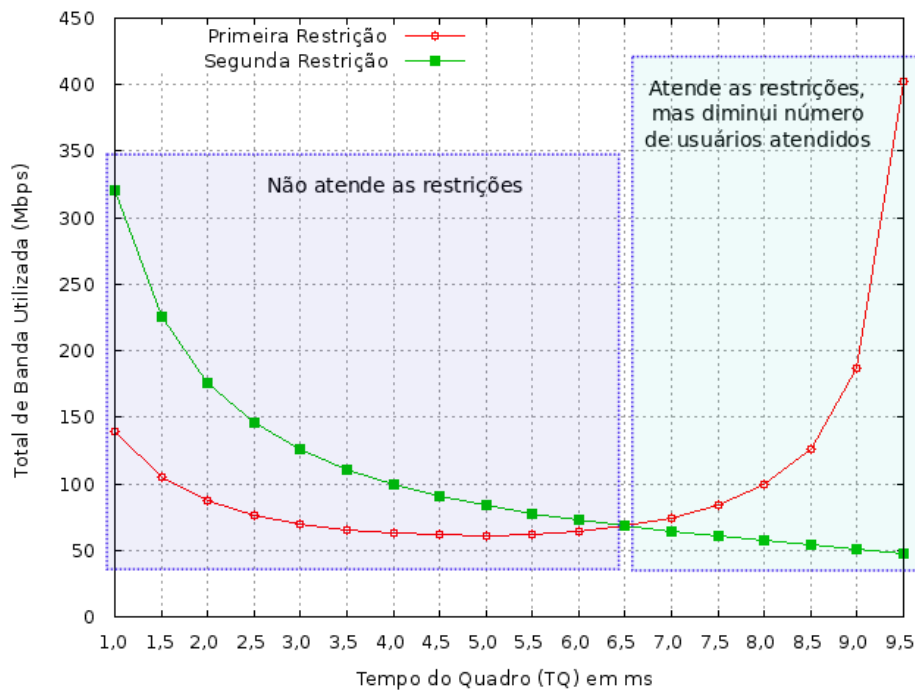
A Figura 30 mostra a aproximação da primeira com a segunda restrição, para o cálculo do valor do  $TQ$  Ideal, sendo que quando o valor da primeira restrição, calculado pela Equação (10), for menor do que a segunda restrição, calculado pela Equação (11), e a taxa física de ambas for menor do que a taxa física alocada pelo sistema ( $R$ ), o valor de  $TQ$  Ideal será encontrado e a quantidade de usuários aptos a transmitir seus pacotes neste quadro é definida em razão dos parâmetros utilizados.

A Tabela 6 descreve a configuração utilizada para a geração do gráfico da Figura 30 e o valor calculado do  $TQ$  Ideal. Maiores detalhes sobre o cálculo do  $TQ$  Ideal são discutidos na Seção 4.3.

**Tabela 6: Configuração utilizada para o exemplo da Figura 30**

Fonte: Autoria própria.

Descrição	Valor
Duração do Símbolo OFDM	13.89 $\mu s$
Atraso Requerido pelo Usuário ( $D_{max,i}$ )	10 ms
$\Delta \rightarrow$ 9 símbolos OFDM	125.10 $\mu s$
TTG + RTG $\rightarrow$ 1 símbolo OFDM	13.89 $\mu s$
Taxa Física	70Mbps
Valor de $TQ$ Ideal calculado	6.5 ms



**Figura 30: Aproximação das duas restrições para o cálculo do  $TQ$  Ideal**

**Fonte: A autoria própria.**

#### 4.2.2 OTIMIZAÇÃO DO TEMPO DO QUADRO UTILIZANDO ENXAME DE PARTÍCULAS

Para calcular o  $TQ$  Ideal, diversas técnicas de otimização, descritas no Capítulo 2, podem ser utilizadas para encontrar a solução deste problema. Nesse trabalho de pesquisa, a otimização por enxame de partículas (PSO) (KENNEDY; EBERHART, 1995) foi a abordagem escolhida por se tratar de um método de otimização heurística, que tem como objetivo a busca de valores, dentro de um conjunto de parâmetros, para maximizar ou minimizar uma determinada função objetivo, sujeita a restrições que são impostas pelo problema. As restrições fazem com que ocorra uma redução no número de soluções possíveis, sendo que, se uma solução satisfaz todas as restrições, ela é chamada de solução viável. E, dentre todas as soluções viáveis, o problema de otimização refere-se à detecção de uma solução ótima.

Assim, para otimizar o  $TQ$ , em cada iteração do sistema, um valor para a duração do  $TQ$  Ideal é calculada. Com a utilização da metaheurística PSO, haverá uma redução no número de execuções, e o tempo de execução será menor, reduzindo a complexidade computacional, como demonstrado através de simulações no Capítulo 5. Isso se deve ao fato de que somente as soluções viáveis são analisadas para o cálculo da duração do  $TQ$  Ideal, como descrito na Seção 2.3.3, do Capítulo 2. A formulação matemática do problema de otimização é descrita a seguir:



**Função Objetivo:**

$$\text{Minimizar } f(TQ) = \frac{\sum_{i=1}^n TU_i}{TQ} \quad (14)$$

**Sujeito a:**

$$\begin{aligned} TQ_{min} &\leq TQ \leq TQ_{max} \\ 0 &\leq \sum r'_i \leq R \end{aligned} \quad (15)$$

$$\frac{(BT'_i - L'_{max,i}) * TQ}{r'_i * TQ - \Delta * R + L'_{max,i}} + T_{DL} + T_{UL} + \frac{L'_{max,i}}{R} + T_{TTG} + T_{RTG} \leq D_{max,i} \quad (17)$$

$$BR_i + \frac{\Delta * R + L'_{max,i}}{TQ} \leq r'_i \quad (18)$$

O cálculo do  $TQ$  Ideal é uma característica muito importante da nova arquitetura e é um diferencial em relação aos escalonadores apresentados na Seção 3.3.2. A duração do tempo de um quadro, naqueles trabalhos, é declarado com um valor fixo, no início do escalonamento, assim, a duração do tempo que cada usuário tem para transmitir seus dados, por estar relacionado ao valor do tempo do quadro, também se torna fixo durante todo o processo de escalonamento.

#### 4.3 CAMADA 2 - ALOCAÇÃO DE RECURSOS

Esta camada trata do problema de alocação dos recursos disponíveis para transmissões no sentido *uplink*. Esses recursos devem ser divididos considerando a capacidade de transmissão de cada segmento no subquadro de UL e, então, alocados para oportunidades de requisição de banda ou de transmissão de dados pelas SSs. Assim, a *Camada 2* da nova arquitetura é responsável pela organização das rajadas a serem transmitidas no subquadro UL. Estas rajadas são alocadas considerando os símbolos OFDM e a configuração AMC utilizada na transmissão do quadro. A construção de um símbolo OFDM foi inicialmente abordada na Seção 3.1.3 e é detalhada a seguir.

Um símbolo OFDM é uma pequena porção de informação, resultante da divisão do espectro de frequências em subportadoras ortogonais, característica do funcionamento da modulação OFDM. Do ponto de vista do domínio do tempo, um símbolo OFDM é dividido em duas partes. Uma parte representa o tempo útil do símbolo OFDM, que é o tempo de transmissão

efetiva. Outra parte corresponde ao tempo de guarda, que é uma cópia da porção final do tempo útil do símbolo OFDM e tem como objetivo a geração do prefixo cíclico (CP - *Cyclic Prefix*).

O CP é responsável pela sincronização dos símbolos OFDM no receptor e por minimizar os efeitos de fenômenos físicos tais como, desvanecimento, atraso por multipercurso e interferência entre símbolos adjacentes.

A Tabela 7 descreve todos os parâmetros que são utilizados para o cálculo de um símbolo OFDM.

**Tabela 7: Parâmetros para o cálculo de um símbolo OFDM**

**Fonte: Tabela adaptada de (IEEE, 2012).**

Parâmetros	Descrição
$R$	Largura de Banda do Canal.
$N_{used}$	Número de subportadoras utilizadas para dados. Na modulação OFDM é 200.
$N_{FFT}$	Número total de subportadoras utilizadas pela IFFT. Deve ser maior do que $N_{used}$ . Na modulação OFDM é 256.
$n$	Fator de amostragem. Utilizado para determinar o espaçamento entre as subportadoras e o tempo útil do símbolo. Valores possíveis: - Largura de banda do canal múltiplo de 1.75 MHz: $n = 8/7$ - Largura de banda do canal múltiplo de 1.5 MHz: $n = 86/75$ - Largura de banda do canal múltiplo de 1.25 MHz: $n = 144/125$ - Largura de banda do canal múltiplo de 2.75 MHz: $n = 316/275$ - Largura de banda do canal múltiplo de 2 MHz: $n = 57/50$ - Para qualquer outra largura de banda do canal: $n = 8/7$
$F_s$	Frequência de amostragem = $\text{floor}(n * R / 8000) * 8000$ - 8000 é um valor de referência para garantir um número inteiro de subportadoras durante o tempo de um símbolo OFDM.
$\Delta_f$	Espaçamento entre as subportadoras = $\frac{F_s}{N_{FFT}}$
$T_b$	Tempo útil do símbolo = $\frac{1}{\Delta_f}$
$G$	O valor de $G$ pode variar entre 1/4, 1/8, 1/16, 1/32. A proporção indica o tempo destinado ao CP em relação a porção do símbolo destinada a transmissão de dados úteis ( $T_b$ ). Por exemplo, quando a BS especifica a utilização de CP igual a 1/4, significa que 25% da duração do símbolo será destinada a transmissão de informações redundantes.
$T_g$	Duração do prefixo cíclico ou tempo de guarda = $G * T_b$
$T_s$	Duração do símbolo OFDM = $T_b + T_g$
$T_{amostragem}$	Tempo de amostragem = $\frac{T_b}{N_{FFT}}$
$R_s$	Taxa de símbolo OFDM = $\frac{1}{T_s}$

A duração efetiva do CP é calculada considerando a proporção de informações redundantes empregadas  $G$  e o tempo utilizado para transmissão de informações úteis  $T_b$  e é definido por

$$T_g = G * T_b \Leftrightarrow G = \frac{T_g}{T_b} . \quad (19)$$

O intervalo do símbolo OFDM destinado a  $T_b$ , é obtido em relação ao espaçamento existente entre as subportadoras ortogonais  $\Delta_f$ , definido por

$$T_b = \frac{1}{\Delta_f}, \quad (20)$$

sendo que o valor de  $\Delta_f$  é definido em unidades de *Hertz* (Hz).

O valor do espaçamento entre subportadoras,  $\Delta_f$ , é obtido pela Equação (21), onde  $N_{FFT}$  representa o número total de subportadoras, incluindo as subportadoras de dados, pilotos e as destinadas aos intervalos de guarda.

$$\Delta_f = \frac{F_s}{N_{FFT}} \Leftrightarrow T_b = \frac{N_{FFT}}{F_s}. \quad (21)$$

A frequência de amostragem  $F_s$ , é definida em função da largura de banda do canal físico  $R$  e do fator de amostragem  $n$ , que é associado a uma largura de banda múltipla de cada valor (descrito na Tabela 7) e expresso por

$$F_s = \left\lfloor \frac{R * n}{8000} \right\rfloor * 8000. \quad (22)$$

Substituindo as ocorrências de  $F_s$  na Equação (22), através da utilização da Equação (21), a Equação (23) é gerada. Através dela, é possível calcular diretamente o tempo destinado a transmissão de informações úteis em cada símbolo OFDM.

$$T_b = \frac{N_{FFT}}{\left\lfloor \frac{R * n}{8000} \right\rfloor * 8000}. \quad (23)$$

Com base no exposto acima e os parâmetros descritos na Tabela 7, a duração de um símbolo OFDM  $T_s$  pode ser calculada através de

$$T_s = \left( \frac{1}{n * \frac{R}{N_{FFT}}} \right) * (1 + G). \quad (24)$$

Como exemplo, é mostrado o cálculo da duração de um símbolo OFDM com a utilização dos parâmetros definidos a seguir:

- $R = 20$  MHz.
- $N_{FFT} = 256$  portadoras.
- $n = 144/125$ , com  $R$  sendo 20 MHz, o fator de amostragem é múltiplo de 1.75 MHz.

- $G = 1/4$ , valor do intervalo de guarda utilizado pelo prefixo cíclico.

Assim, aplicando os valores na Equação (24), a duração de um símbolo OFDM é de  $13,89 \mu s$ , valor utilizado neste trabalho para a validação do modelo no Capítulo 5.

A partir da definição do símbolo OFDM é possível determinar a capacidade de transmissão da unidade de alocação mínima utilizada no subquadro UL, ou seja, o segmento físico UL  $N_{segUL}$ , definido por

$$N_{segUL} = \left\lfloor \frac{TQ}{T_s} \right\rfloor, \quad (25)$$

sendo que o tempo de duração de um quadro, determinado pela variável  $TQ$ , é calculado na *Camada 1*, conforme os parâmetros da rede e, a cada iteração do escalonador, seu valor poderá ser alterado, de acordo com o comportamento da rede.

Retornando ao exemplo acima, com a utilização da Equação (24), o valor calculado da duração de um símbolo OFDM é de  $13,89 \mu s$ . Se a duração do TQ for calculada em  $10 ms$ , com a utilização da Equação (25), ter-se-á 720 segmentos físicos para serem utilizados por este quadro.

Com o valor de  $N_{segUL}$ , a *Camada 2* da nova arquitetura define as áreas do subquadro UL utilizadas para requisição de banda e transmissão de rajadas de dados que são enviadas pela mensagem UL\_MAP. A quantidade de segmentos físicos destinados a cada uma dessas áreas é definida pela nova arquitetura, com a seguinte ordem de prioridade: (i) registro inicial; (ii) IE *polling*, (iii) oportunidades para requisição de banda por contenção (IE contenção); e (iv) oportunidades para transmissão de dados (IE dados), conforme visto na Figura 26. A *Camada 2* realiza a alocação apenas dos três primeiros grupos. A porção do quadro destinada as oportunidades de transmissão de dados (IE Dados), é gerenciada pela *Camada 3*.

Desta forma, a *Camada 2* fornece para a *Camada 3* a quantidade de segmentos físicos disponíveis para alocação de oportunidades de transmissão de dados ( $N_{segDados}$ ), que serão distribuídos entre as SSs para transmissão de dados. Esse parâmetro é calculado por

$$N_{segDados} = \left\lfloor \frac{TQ - (T_{DL} + T_{TTG} + T_{RTG} + \Delta)}{T_s} \right\rfloor, \quad (26)$$

sendo que, do valor de  $TQ$ , são subtraídos os valores que não são úteis para a transmissão dos dados no subquadro de UL. Assim feito, tem-se a quantidade de segmentos físicos disponíveis para a transmissão de dados.

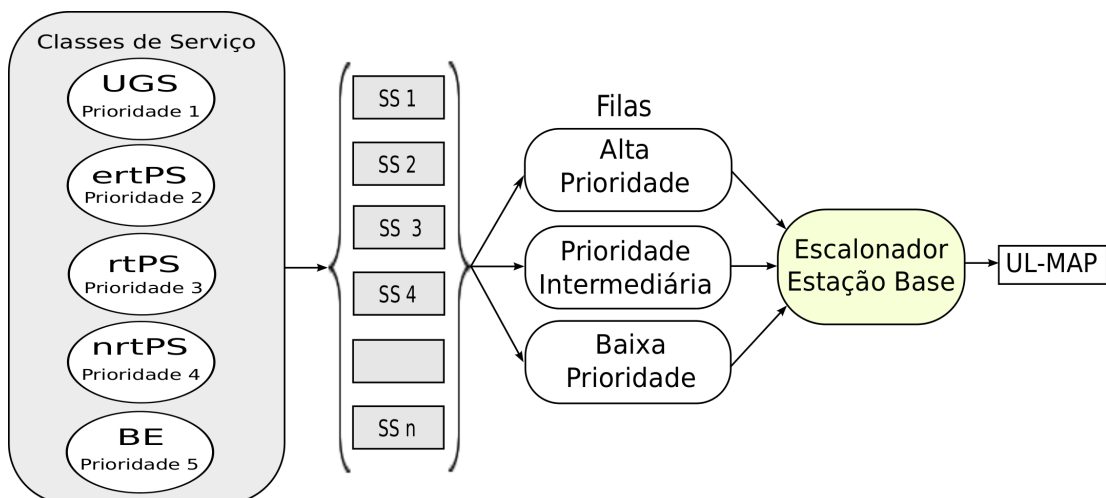
#### 4.4 CAMADA 3 - CONTROLE DE ADMISSÃO DE CONEXÕES (CAC)

A *Camada 3* da nova arquitetura foi projetada para realizar o controle de admissão das conexões, sendo responsável pelo gerenciamento das filas de requisições de banda, no sentido *uplink*, enviadas pelas SSs nos quadros anteriores, que estão armazenadas e aguardam atendimento pela BS. Uma conexão só é admitida quando a largura de banda calculada pela primeira restrição na Equação (10) e a largura de banda calculada pela segunda restrição na Equação (11) forem menor ou igual ao total da largura de banda  $R$  disponibilizada pelo sistema.

Desta forma, fecha-se o ciclo da mensagem no escalonador de forma hierárquica, sendo que a *Camada 1*, calcula a duração do  $TQ$  Ideal, a *Camada 2*, realiza a alocação dos recursos disponíveis para transmissões no sentido *uplink* e, com a sinalização dos recursos recebidos da *Camada 2*, a *Camada 3* gerencia as filas de requisições de banda.

Para isso, a solução apresentada para o escalonamento na *Camada 3*, é baseada em um controle de admissão com prioridades, onde, para cada classe de serviço, é atribuída uma prioridade. É importante destacar que existe um tratamento diferenciado para o tráfego em tempo real, que é priorizado devido ao fato de ser sensível ao atraso, com o armazenamento nas filas de maior prioridade para ser atendido antes do que o tráfego que não necessita de execução em tempo real, que é tolerante ao atraso.

Na nova arquitetura, os níveis de prioridade foram definidos conforme mostrado na Figura 31. Foram estabelecidas 5 prioridades, sendo a prioridade 1 para a classe de serviço UGS, a prioridade 2 para a classe de serviço ertPS, a prioridade 3 para a classe de serviço rtPS, a prioridade 4 para a classe de serviço nrtPS e prioridade 5 para a classe de serviço BE.



**Figura 31: Prioridades de atendimento de requisições**

Fonte: Autoria própria.

A classe de serviço ertPS tem prioridade maior do que a prioridade da classe de serviço rtPS, pelo fato de que esta classe combina a eficiência das classes de serviço UGS e rtPS, utilizando *codecs* EVRC (*Enhanced Variable Rate Codec*) (3GPP2, 2004) para VoIP, o que desperdiça menos recursos em relação ao tráfego na classe rtPS.

Após definidas as prioridades, o mecanismo de escalonamento da nova arquitetura utiliza três filas de escalonamento, alta prioridade, prioridade intermediária e baixa prioridade. A fila de alta prioridade é reservada para as requisições de largura de banda das classes de serviço UGS. A fila intermediária é reservada para armazenar as requisições de largura de banda das classes de serviços ertPS e rtPS. A fila de baixa prioridade é utilizada para as requisições de largura de banda da classe de serviço nrtPS e BE.

O principal destaque da nova arquitetura, além da inclusão de prioridades para as classes de serviço, é o cálculo do  $TQ$  Ideal, que é realizado pela *Camada 1*, e que possibilita um aumento do seu desempenho em relação a determinados parâmetros de QoS, bem como a maximização da utilização dos recursos do enlace *uplink*.

O cálculo do  $TQ$  Ideal será executado para as classes de serviço UGS, ertPS e rtPS. Esse cálculo será executado cada vez que novas SSs requisitarem acesso para a BS, se SSs deixarem a rede, ou se uma SS trocar seus parâmetros de QoS. Para isso, o Algoritmo 3, descrito na Seção 4.5 é executado, utilizando uma nova disciplina, denominada  $TQ/PSO$ , originalmente desenvolvida para esta nova arquitetura e que se utiliza das disciplinas PQ (*Priority Queuing*) e LR (*Latency-Rate*).

#### 4.5 PSEUDOCÓDIGOS DA NOVA ARQUITETURA

A nova arquitetura com otimização do  $TQ$  através da utilização de PSO em redes WiMAX, tem seu módulo principal descrito no Algoritmo 3. O mecanismo de escalonamento proposto é executado de forma hierárquica, conforme a descrição das 3 camadas nas Seções 4.2, 4.3 e 4.4.

A primeira tarefa que a nova arquitetura deve realizar é a definição do  $TQ$ . Para isso, no passo 1, o Algoritmo 3 executa uma chamada ao módulo  $TQIdeal$ , descrito no Algoritmo 4. Após o cálculo do  $TQ$  Ideal, nos passos 2 até 4 do módulo, as oportunidades de requisição de banda e as concessões de envio de dados de forma periódica, começam a ser distribuídas, de acordo com as prioridades que foram estabelecidas pela nova arquitetura, sempre iniciando pela prioridade 1 e que serão escalonadas no próximo quadro.

No passo 5, é realizada a verificação para saber se existe largura de banda suficiente

**Algoritmo 3: Algoritmo Principal**  
**Fonte: Autoria própria.**

- 
- 1: Calcular a duração do  $TQ$  Ideal através do módulo  $TQ$  Ideal.
  - 2: Inserir, na fila de alta prioridade, as requisições da classe de serviço UGS, que possuem concessões de dados periódicos e que devem ser escalonadas no próximo quadro.
  - 3: Inserir, na fila de prioridade intermediária, as requisições das classes de serviço ertPS e rtPS, para serem escalonadas no próximo quadro.
  - 4: Inserir, na fila de baixa prioridade, as requisições das classes de serviço nrtPS e BE, para serem escalonadas no próximo quadro.
  - 5: Verificar se a largura de banda da primeira restrição mais a largura de banda da segunda restrição é menor ou igual a largura de banda disponibilizada pelo sistema.
  - 6: Atender a todos as requisições na fila com alta prioridade em primeiro lugar.
- 

para o valor do  $TQ$  Ideal calculado, ou seja, se a largura de banda da primeira restrição e a largura de banda da segunda restrição são iguais ou menores do que a largura de banda  $R$  disponibilizada pelo sistema.

No passo 6, a nova arquitetura começa o atendimento de todas as requisições que estão na fila e foram alocadas, sempre pela prioridade mais alta, ou seja, a prioridade 1, da classe de serviço UGS.

O Algoritmo 4 descreve o módulo  $TQIdeal$  que tem a função de calcular o tempo ideal de duração de um quadro.

**Algoritmo 4: módulo  $TQIdeal$**   
**Fonte: Autoria própria.**

**Dados de Entrada:**

- $L'_{max,i}$  = Tamanho máximo de um pacote de dados com a sobrecarga  $\Delta$ .
- $R$  = Taxa Física do Canal.
- $D_{max,i}$  = Atraso requerido pelo usuário.
- $BT_i$  = Tamanho do Balde de Fichas.
- $BR'_i$  = Taxa do Balde de Fichas com a sobrecarga  $\Delta$ .
- $T_{TTG}$  = Intervalo de tempo entre o subquadro de *downlink* e *uplink*.
- $T_{RTG}$  = Intervalo de tempo entre o subquadro de *uplink* e *downlink*.
- $\Delta$  = Sobrecarga para transmissão dos pacotes.

**Algoritmo:**

- 1: Inicializar o valor de  $TQ$ .
- 2: Executar o módulo  $TQ\_PSO$  para otimizar o  $TQ$  com PSO.
- 3: Executar o módulo  $Segmentos\_SimboloOFDM$  para calcular a quantidade de segmentos físicos por símbolo OFDM.

**Dado de Retorno:**

Duração do  $TQ$  Ideal.

---

Esse módulo inicializa, no passo 1, o valor do  $TQ$ , que o padrão IEEE 802.16

estabelece como duração mínima do tempo de um quadro o valor de  $2.50\text{ ms}$ . Esta inicialização se faz necessária para o passo 2, que tem a finalidade de executar o módulo  $TQ\_PSO$ , descrito no Algoritmo 5.

Após a execução do módulo  $TQ\_PSO$ , o passo 3 executa o módulo  $Segmentos\_SimboloOFDM$ , descrito no Algoritmo 9, que calcula a quantidade de segmentos físicos por símbolo OFDM. O Algoritmo 4 retorna a duração do  $TQ$  Ideal para o Algoritmo *Principal*.

O Algoritmo 5, descreve o módulo  $TQ\_PSO$ , que tem a função de calcular o valor do  $TQ$  Ideal, dentro do espaço de busca do enxame de partículas. Os parâmetros de entrada, descritos no início do módulo, são necessários para que o módulo execute suas funções.

Três restrições que são apresentadas neste módulo. A restrição *A* significa que o  $TQ$  deve estar entre os limites do valor do  $TQ$  mínimo ( $TQ_{min}$ ) e do  $TQ$  máximo ( $TQ_{max}$ ), que, de acordo com o padrão IEEE 802.16, deve estar entre  $2,50\text{ ms}$  e  $20,00\text{ ms}$ . Também, na restrição *A*, é realizado o controle de banda, ou seja, o valor da somatória das taxas alocadas por cada servidor  $r'_i$  deve ser maior ou igual do que zero e menor ou igual ao valor da taxa física  $R$  do sistema. A restrição *B*, é a condição de restrição que se refere ao atraso. A restrição *C* é a condição de restrição que está relacionada ao balde de fichas.

A Tabela 8 descreve os valores dos parâmetros utilizados pelo algoritmo PSO desenvolvido para este trabalho.

**Tabela 8:** Parâmetros do algoritmo PSO para o cálculo do  $TQ$  Ideal

Fonte: Autoria própria.	
Parâmetros	Valores
$w$	$1 / (2 * \ln 2)$
$w_{min}$	0.4
$w_{max}$	0.9
$c_1$	$0.3 + \ln (2)$
$c_2$	$0.5 + \ln (2)$

O passo 1 do módulo  $TQ\_PSO$  calcula a duração de um símbolo OFDM, através da execução do módulo  $SimboloOFDM$ , descrito no Algoritmo 6. O valor do símbolo OFDM é utilizado para calcular a quantidade de segmentos físicos que um quadro pode comportar.

O passo 2 executa o algoritmo de otimização PSO e tem duas importantes características: (i) durante o procedimento de inicialização, todas as partículas são inicializadas repetidamente até satisfazer as três restrições para o cálculo do  $TQ$  Ideal com a utilização da metaheurística PSO; e, (ii) ao calcular os valores  $pbest_i$  e  $gbest$ , apenas as posições no espaço



### Algoritmo 5: módulo $TQ\_PSO$

Fonte: Autoria própria.

#### Parâmetros de Entrada:

$s$  = Tamanho do enxame.

$i$  = Partícula que representa uma solução candidata.

$v_i$  = Velocidade atual da partícula  $i$ .

$x_i$  = Posição atual da partícula  $i$ .

$c_1$  = Parâmetro de confiança → Coeficiente cognitivo.

$c_2$  = Parâmetro de confiança → Coeficiente social.

$w$  = Fator de inércia.

$r_1$  = Vetor de números aleatórios distribuídos no intervalo  $[0,1]$  → variável aleatória do coeficiente cognitivo.

$r_2$  = Vetor de números aleatórios distribuídos no intervalo  $[0,1]$  → variável aleatória do coeficiente social.

$pbest_i$  = Melhor posição alcançada pela partícula  $i$ , com base na experiência própria.

$gbest$  = Melhor posição encontrada pelo enxame, com base nas melhores posições de  $pbest$ .

$TQ_{min}$  = Duração mínima do  $TQ$ . Segundo o padrão IEEE 802.16  $TQ_{min} = 2,50$  ms.

$TQ_{max}$  = Duração máxima do  $TQ$ . Segundo o padrão IEEE 802.16  $TQ_{max} = 20,00$  ms.

#### Restrições do Algoritmo PSO:

Função Objetivo do Algoritmo PSO:  $Minimizar f(TQ) = \frac{\sum_{i=1}^n TU_i}{TQ}$

Sujeito, para cada sessão  $i$ , a:

$$A) \quad TQ_{min} \leq TQ \leq TQ_{max} \\ 0 \leq \sum r'_i \leq R$$

$$B) \quad \frac{(BT'_i - L'_{max,i}) * TQ}{r'_i * TQ - \Delta * R + L'_{max,i}} + T_{DL} + T_{DL} + \frac{L'_{max,i}}{R} + T_{TTG} + T_{RTG} \leq D_{max,i}$$

$$C) \quad BR_i + \frac{\Delta * R + L'_{max,i}}{TQ} \leq r'_i$$

#### Algoritmo:

- 1: Calcular a duração do símbolo OFDM através do módulo *SimboloOFDM*
- 2: **para cada** partícula  $i = 1$  até  $s$  **faça** ( $s$  = tamanho do enxame)
- 3:     Calcular o valor da primeira restrição através do módulo *PrimeiraRestricao*
- 4:     Calcular o valor da segunda restrição através do módulo *SegundaRestricao*
- 5:     Iniciar  $x_i$  com uma solução aleatória para o problema
- 6:     Iniciar  $v_i$  com uma velocidade aleatória  $< V_{max}$  ( $V_{max}$  = valor máximo da velocidade)
- 7:      $pbest_i \leftarrow gbest \leftarrow x_i$
- 8: **fim para cada**
- 9: **enquanto** não atingir a condição de parada **faça** (condição de parada = tamanho do enxame)
- 10:    **para cada** partícula  $i = 1$  até  $s$  **faça**
- 11:     **se**  $x_i$  for melhor que  $pbest_i$  **E** a partícula está no espaço viável **então**
- 12:        $pbest_i = x_i$
- 13:     **fim se**
- 14:     **se**  $pbest_i$  for melhor que  $gbest$  **então**
- 15:        $gbest = pbest_i$
- 16:     **fim se**
- 17:     Atualizar a velocidade ( $v$ ) da partícula  $i$ , através da equação:  

$$v_{(i+1)} = w * v_i + c_1 * r_1(pbest_i - x_i) + c_2 * r_2(gbest - x_i)$$
- 18:     Atualizar a posição ( $x$ ) da partícula  $i$ , através da equação:  

$$x_{i+1} = x_i + v_i$$
- 19:    **fim para cada**
- 20: **fim para cada**

#### Retorno:

Duração do  $TQ$  Ideal.

Duração de um símbolo OFDM ( $T_s$ ).

viável são computadas, ou seja, todas as soluções precisam estar localizadas no espaço de busca viável, conforme descrito no passo 11.

Nos passos 3 e 4, antes de iniciar a busca no enxame, as restrições  $B$  e  $C$  são inicializadas através dos módulos *PrimeiraRestricao*, descrito no Algoritmo 7, e *SegundaRestricao*, descrito no Algoritmo 8.

A partir do passo 5, o módulo *TQ\_PSO*, do Algoritmo 5, segue as etapas descritas em Fukuyama (2007) e Eberhart et al. (1996):

1. **Inicialização:** Para cada partícula  $i$ , inicializar aleatoriamente os vetores de posição  $x_i$  e velocidade  $v_i$  no espaço de busca  $D$ -dimensional;
2. **Avaliação:** Avaliar a função objetivo de cada partícula  $i$ ;
3. **Atualização do  $pbest_i$ :** Comparar a avaliação da função de aptidão da partícula  $i$ ,  $x_i$  com o  $pbest_i$  da partícula. Se o valor corrente é melhor que  $pbest_i$ , então o valor de  $pbest_i$  passa a ser igual ao valor da função de aptidão da partícula  $i$ , e a localização do  $pbest_i$  passa ser igual a localização atual no espaço  $D$ -dimensional;
4. **Atualização do  $gbest$ :** Comparar o valor obtido em  $pbest_i$  com o melhor valor global  $gbest$ . Se o valor de  $pbest_i$  for melhor, atualizar  $gbest$  com o novo valor;
5. **Atualização da Velocidade:** Atualizar a velocidade da partícula  $i$  de acordo com a equação definida no passo 17;
6. **Atualização da Posição:** Atualizar a posição da partícula  $i$  de acordo com a equação definida no passo 18;
7. **Critério de Parada:** Repetir os passos 9-19 até que um critério de parada seja encontrado.

Ao final da execução do módulo *TQ/PSO* do Algoritmo 5, os valores das variáveis  $TQ$  Ideal e da duração de um símbolo OFDM  $T_s$ , são retornadas para o módulo *TQIdeal* do Algoritmo 4.

O módulo *SimboloOFDM*, descrito no Algoritmo 6, tem a função de executar o cálculo da duração de um símbolo OFDM. Através dos parâmetros de entrada, o módulo executa, no passo 1, o cálculo da duração do símbolo OFDM, que será utilizado para alocar os segmentos físicos em cada quadro. O módulo retorna a duração do símbolo OFDM  $T_s$  para o módulo *TQ\_PSO* do Algoritmo 5.

**Algoritmo 6: módulo *SimboloOFDM***  
**Fonte: Autoria própria.**

Dados de Entrada:

- $R$  = Taxa Física do Canal.  
 $N_{FFT}$  = Número de subportadoras utilizadas para dados.  
 $n$  = Fator de amostragem.  
 $G$  = Valor do intervalo de guarda.

Algoritmo:

- 1: Calcular o valor da duração do símbolo OFDM através da fórmula:

$$T_s = \left( \frac{1}{n * \frac{R}{N_{FFT}}} \right) * (1 + G)$$

Retorno:

- Duração de um símbolo OFDM ( $T_s$ ).

O módulo *PrimeiraRestricao*, descrito no Algoritmo 7, executa a restrição referente ao atraso, através da equação no passo 1 e descrita na Equação 10 da Seção 4.2.1 deste capítulo. O retorno, para o módulo *TQ\_PSO* do Algoritmo 5, é o valor total da taxa alocada pelo servidor ( $r'_i$ ).

**Algoritmo 7: módulo *PrimeiraRestricao***  
**Fonte: Autoria própria.**

- 1: Calcular, para cada sessão  $i$ , o valor da taxa alocada pelo servidor ( $r'_i$ ) através da equação:

$$r'_i = \frac{\frac{(BT'_i - L'_{max,i}) * TQ}{L'_{max,i}} + \Delta * R + L'_{max,i}}{\frac{D_{max,i} - (T_{DL} + T_{UL}) - \frac{L'_{max,i}}{R} - (T_{TTG} + T_{RTG})}{TQ}}$$

Retorno:

- Primeira restrição: Total da taxa alocada pelo servidor ( $r'_i$ ).

O módulo *SegundaRestricao*, descrito no Algoritmo 8, executa a restrição referente ao balde de fichas, através da equação descrita no passo 1. O retorno, para o módulo *TQ\_PSO* do Algoritmo 5, é o valor da condição do balde de fichas ( $TBR$ ).

**Algoritmo 8: módulo *SegundaRestricao***  
**Fonte: Autoria própria.**

- 1: Calcular, para cada sessão  $i$ , o valor da Condição do Balde de Fichas ( $TBR$ ) através da equação:

$$TBR = BR_i + (((\Delta * R) + L'_{max,i}) / TQ)$$

Retorno:

- Segunda restrição: Condição do Balde de Fichas ( $TBR$ ).

No Algoritmo 9, a quantidade de segmentos físicos que serão alocadas em cada quadro é calculado pelo módulo *Segmentos\_SimboloOFDM*, descrito abaixo.

**Algoritmo 9: módulo Segmentos Simbolo OFDM**  
**Fonte: Autoria própria.**

Algoritmo:

1: Calcular a quantidade de segmentos físicos por quadro através da fórmula:

$$N_{segUL} = \left\lfloor \frac{TQ}{T_s} \right\rfloor$$

2: Calcular a quantidade de segmentos físicos disponíveis para transmissão de dados através da fórmula:

$$N_{segDados} = \left\lfloor \frac{TQ - (T_{DL} + T_{TTG} + T_{RTG} + \Delta)}{T_s} \right\rfloor$$

Dados de Retorno:

Quantidade de segmentos físicos por quadro ( $N_{segUL}$ ).

Quantidade de segmentos físicos disponíveis para transmissão de dados ( $N_{segDados}$ ).

No passo 2, a quantidade de segmentos físicos que estarão disponíveis para transmissão de dados é calculada. Este módulo retorna, para o módulo  $TQ_{Ideal}$  do Algoritmo 4, a quantidade de segmentos físicos para alocação em cada quadro ( $N_{segUL}$ ) e a quantidade de segmentos físicos disponíveis para transmissão de dados ( $N_{segDados}$ ).

#### 4.6 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou uma nova arquitetura que otimiza o tempo de um quadro utilizando enxame de partículas em redes WiMAX fixas, com o objetivo de garantir a QoS para os diferentes tipos de serviços oferecidos por uma rede, tratando as aplicações de acordo com diferentes tipos de prioridade.

A nova arquitetura é utilizada no sentido de tráfego *uplink* para a BS em redes IEEE 802.16, e é baseado na modulação OFDM, utilizando a topologia PMP (*Point to Multipoint*) e a duplexação por divisão de tempo (*Time Division Duplexing* - TDD).

Na implementação de algoritmos de escalonamento, o principal desafio é o comportamento das filas de conexões localizadas nas SSs. A minimização do tempo entre o envio e o recebimento de um pacote é importante para garantir os requisitos de atraso, variação do atraso e vazão. Com isso, o mecanismo de escalonamento se torna um importante componente para obtenção de QoS em redes IEEE 802.16. A norma define apenas a utilização de um algoritmo padrão.

Assim, este capítulo descreveu uma nova arquitetura, que tem como objetivo prover QoS para as cinco classes de serviços definidas pelo padrão IEEE 802.16, com garantia do limite de atraso para àquelas sensíveis ao atraso. Descreveu também, que uma importante modificação é introduzida, através do cálculo da duração do tempo do quadro, em cada iteração

do escalonador quando houver a necessidade, utilizando a abordagem de otimização PSO. Essa nova abordagem vai permitir a utilização de forma eficiente a largura de banda disponível, possibilitando que mais usuários possam ser alocados na rede, aumentando o nível de utilização do sistema.

A Seção 4.5 apresentou o Algoritmo *Principal*, com os módulos que fazem sua composição e que foram implementados. Os resultados desta implementação são apresentados e discutidos no Capítulo 5.

## 5 ANÁLISE DE DESEMPENHO DA NOVA ARQUITETURA

Este capítulo faz a análise de desempenho da nova arquitetura. A análise é dividida em duas avaliações, com dois cenários e com diferentes cargas de tráfego. Na primeira avaliação, os resultados obtidos são comparados com os resultados obtidos com uma avaliação utilizando o método de busca exaustiva (força bruta), com o objetivo de validar o cálculo da duração do tempo do quadro  $TQ$  Ideal utilizando a metaheurística PSO, desenvolvido neste trabalho.

Com o valor calculado do  $TQ$  Ideal, uma segunda avaliação é realizada, desta vez comparando a nova arquitetura com o trabalho em (JAKIMOSKI; JANEVSKI, 2013). Nesta avaliação, é verificado se o cálculo do  $TQ$  Ideal gera alguma vantagem nos quesitos de atraso médio, variação média do atraso, taxa útil média e número de usuários admitidos, com o mesmo cenário, com a mesma carga de tráfego, ou seja, com os mesmos parâmetros, tendo como alteração a duração do tempo do quadro, que no trabalho relacionado é fixado em  $5\text{ ms}$  e, neste trabalho, é calculado com o módulo  $TQ\_PSO$  no Algoritmo 5, do Capítulo 4.

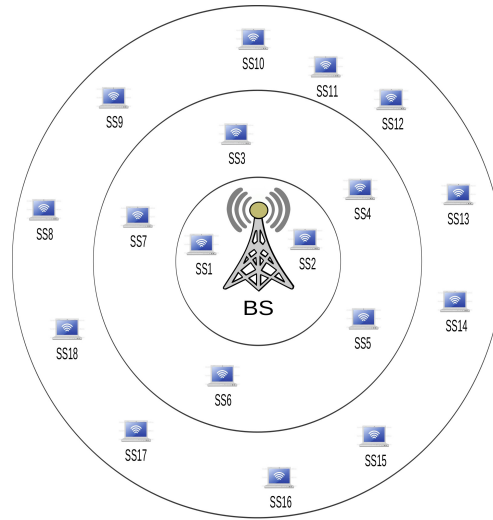
### 5.1 CENÁRIOS E PARÂMETROS DO AMBIENTE DE AVALIAÇÃO E DE SIMULAÇÃO

Nesta seção, as principais características da modelagem, envolvendo a avaliação e a simulação são apresentadas, através da descrição dos cenários e dos parâmetros utilizados. Na Seção 5.1.1, é realizada uma comparação com os resultados obtidos com uma avaliação utilizando o método de busca exaustiva, para validar o desempenho do cálculo da duração do  $TQ$  Ideal. Na Seção 5.1.2, é realizada uma avaliação do desempenho baseado em simulações com o *software* NS-2 (NS-2, 2015) e comparado com (JAKIMOSKI; JANEVSKI, 2013).

#### 5.1.1 CENÁRIO 1 - CÁLCULO DO TQ IDEAL

Para a realização dos testes de desempenho da nova arquitetura, foi utilizada uma rede IEEE 802.16 composta por uma BS que se comunica com dezoito SSs, com somente um tráfego por SS e o destino de todos os fluxos é a BS, conforme mostrado na Figura 32. O ambiente

de avaliação utilizado, foi desenvolvido na linguagem de programação C (KERNIGHAN; RITCHIE, 1989), tanto no modelo de busca exaustiva, como na nova arquitetura.



**Figura 32: Cenário de avaliação de uma rede para o cálculo do  $TQ$  Ideal**  
**Fonte: adaptado de (FERREIRA, 2011).**

Nessa topologia, utilizada para avaliação do atraso máximo requerido, existem diferentes tipos de tráfegos, sendo que, conforme a sumarização apresentada na Tabela 9, das dezoito SSs, seis enviam, para a BS, tráfego de áudio CBR, seis SSs enviam tráfego de vídeo VBR e seis SSs enviam tráfego de vídeo MPEG4.

**Tabela 9: Descrição de diferentes tráfegos**  
**Fonte: Autoria própria.**

SS	Aplicação	Período de Chegada (ms)	Tamanho do Pacote (max) (bytes)	Taxa de Envio (kbps) (média)
1 → 6	Áudio	4.7	160	64
7 → 12	Vídeo VBR	26	1024	≈ 200
13 → 18	Vídeo MPEG4	2	800	3200

A Tabela 10, apresenta os parâmetros das camadas PHY e MAC, utilizados na avaliação.

**Tabela 10: Parâmetros das camadas PHY e MAC**  
**Fonte: Autoria própria.**

<b>Largura de Banda</b>	20MHz
<b>Duração do Símbolo OFDM</b>	13.89 $\mu s$
<b>Atraso Requerido pelo Usuário</b>	5 / 10 / 15 e 20 ms
$\Delta$ (Initial Ranging e BW Request) → 9 símbolos OFDM	125.10 $\mu s$
<b>TTG + RTG</b> → 1 símbolo OFDM	13.89 $\mu s$
<b>Sub-quadro de UL (cabeçalho)</b> → 10% Símbolo OFDM	1.39 $\mu s$
<b>Taxa de Dados Máxima</b>	70Mbps

O desempenho da nova arquitetura é avaliado através do atraso máximo requerido pelo usuário e das estações alocadas. O resultado da alocação de estações no sistema, através de um  $TQ$  Ideal, limitado pelo atraso requerido, são descritos a seguir. Inicialmente, os parâmetros do balde de fichas são estimados de acordo com as características do tráfego de entrada conforme apresentados na Tabela 11.

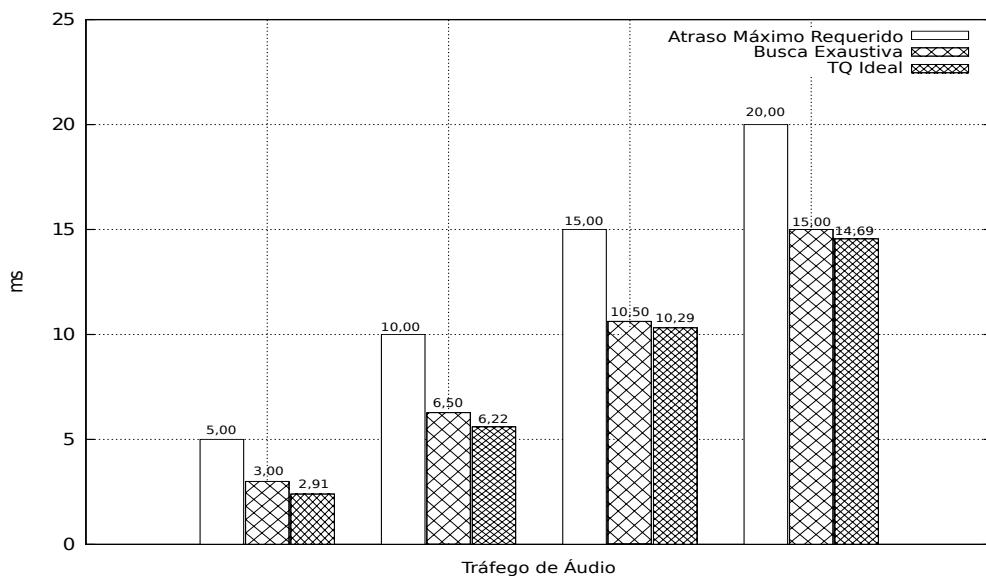
**Tabela 11:** Parâmetros do balde de fichas  
Fonte: Autoria própria.

	Áudio	Vídeo VBR	Vídeo MPEG4
Tamanho do Balde (bits)	3000	18000	10000
Taxa do Balde (kb/s)	64	500	4100

#### 5.1.1.1 RESULTADOS OBTIDOS COM A PRIMEIRA AVALIAÇÃO

Para calcular o valor da duração do  $TQ$  Ideal, são utilizados os parâmetros do balde de fichas, o atraso máximo requerido pelo usuário, a taxa física e o tamanho máximo do pacote.

A Figura 33 apresenta o valor que corresponde a duração de um  $TQ$  Ideal, para o tráfego de áudio, com diferentes valores de  $TQ$  para quatro valores de atraso máximo requerido pelo usuário. No método de busca exaustiva, representado no gráfico por *Busca Exaustiva*, quando o usuário solicita um atraso máximo de 5,00 ms, o valor ótimo do quadro é de 3,00 ms. Para o atraso máximo de 10,00 ms, tem-se um valor ótimo de quadro de 6,50 ms. Para um atraso máximo de 15,00 ms, o valor ótimo do quadro é de 10,50 ms e, para um atraso máximo de 20,00 ms, o valor ótimo do quadro é de 15,00 ms.

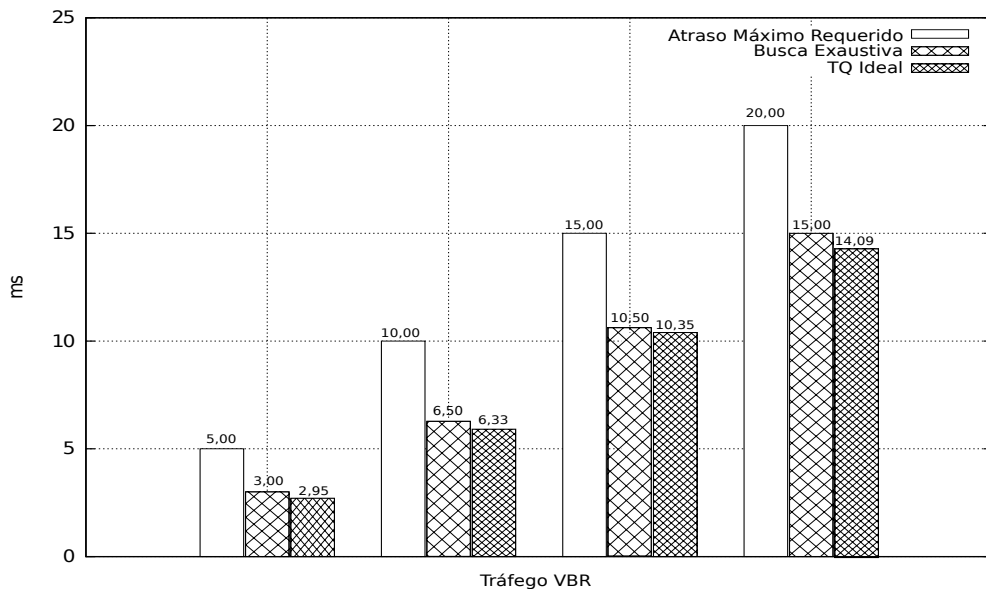


**Figura 33:** Duração do  $TQ$  Ideal para o tráfego de áudio  
Fonte: Autoria própria.



Com a utilização do PSO, é possível verificar que o cálculo do tempo do quadro ideal, representado no gráfico da Figura 33 por *TQ Ideal*, tem uma pequena redução no seu valor, quando comparado com o método da busca exaustiva, o que valida o resultado do cálculo do *TQ Ideal* com a utilização da metaheurística PSO, ou seja, para um atraso de 5,00 ms, o valor de *TQ* é de 2,91 ms. Para um atraso de 10,00 ms, tem-se um valor de *TQ* de 6,22 ms. Para um atraso de 15,00 ms, o valor de *TQ* é de 10,29 ms, e, para um atraso de 20,00 ms, o valor do *TQ* é de 14,69 ms.

No gráfico da Figura 34, para o tráfego de vídeo VBR, o método da busca exaustiva calcula, para um atraso máximo solicitado pelo usuário de 5,00 ms, o valor ótimo do quadro é de 3,00 ms. Para o atraso máximo de 10,00 ms, tem-se um valor ótimo de quadro de 6,50 ms. Para um atraso máximo de 15,00 ms, o valor ótimo do quadro é de 10,50 ms e, para um atraso máximo de 20,00 ms, o valor ótimo do quadro é de 15,00 ms.



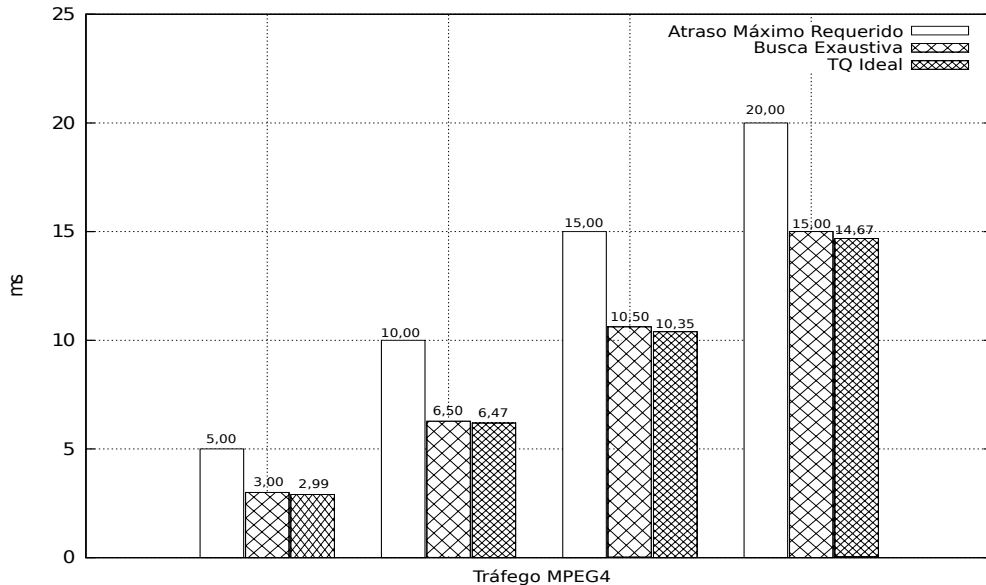
**Figura 34: Duração do *TQ Ideal* para o tráfego VBR**

**Fonte: Autoria própria.**

Para o tráfego de vídeo VBR, o gráfico da Figura 34 mostra novamente que, com a utilização do PSO, os valores do cálculo do tempo de um quadro sofrem uma pequena redução. Para um atraso de 5,00 ms, o valor de *TQ* é de 2,95 ms. Para um atraso de 10,00 ms, tem-se um valor de *TQ* de 6,33 ms. Para um atraso de 15,00 ms, o valor de *TQ* é de 10,35 ms, e, para um atraso de 20,00 ms, o valor do *TQ* é de 14,09 ms.

Já para o tráfego de vídeo MPEG4, o cálculo com o método da busca exaustiva, apresentados no gráfico da Figura 35, para um atraso máximo solicitado pelo usuário de 5,00 ms, o valor ótimo do quadro é de 3,00 ms. Para o atraso máximo de 10,00 ms, tem-se um valor

ótimo de quadro de 6,50 *ms*. Para um atraso máximo de 15,00 *ms*, o valor ótimo do quadro é de 10,50 *ms* e, para um atraso máximo de 20,00 *ms*, o valor ótimo do quadro é de 15,00 *ms*.



**Figura 35: Duração do  $TQ$  Ideal para o tráfego MPEG4**

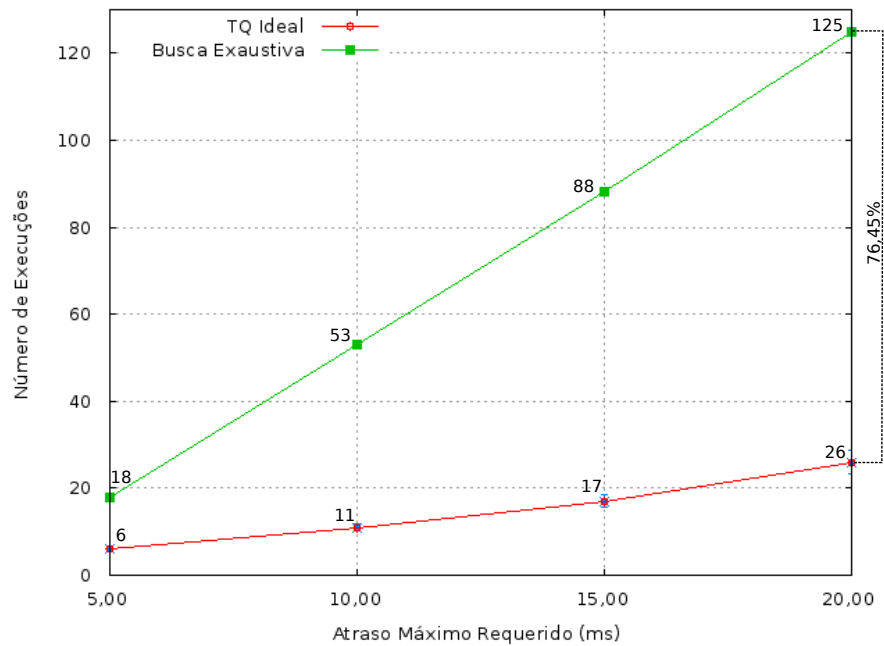
Fonte: Autoria própria.

Os valores calculados com a utilização do PSO, para o tráfego de vídeo MPEG4, apresentados no gráfico da Figura 35, também têm a característica dos dois gráficos apresentados anteriormente, ou seja, para um atraso de 5,00 *ms*, o valor de  $TQ$  é de 2,99 *ms*. Para um atraso de 10,00 *ms*, tem-se um valor de  $TQ$  de 6,47 *ms*. Para um atraso de 15,00 *ms*, o valor de  $TQ$  é de 10,35 *ms*, e, para um atraso de 20,00 *ms*, o valor do  $TQ$  é de 14,67 *ms*.

Com os resultados apresentados acima, percebe-se que os valores do tempo do quadro, calculados com o método da busca exaustiva, e os valores do  $TQ$  Ideal, calculados com a utilização da metaheurística PSO, possuem uma pequena diferença do tempo de quadro, o que valida a utilização do PSO para o cálculo do  $TQ$  Ideal.

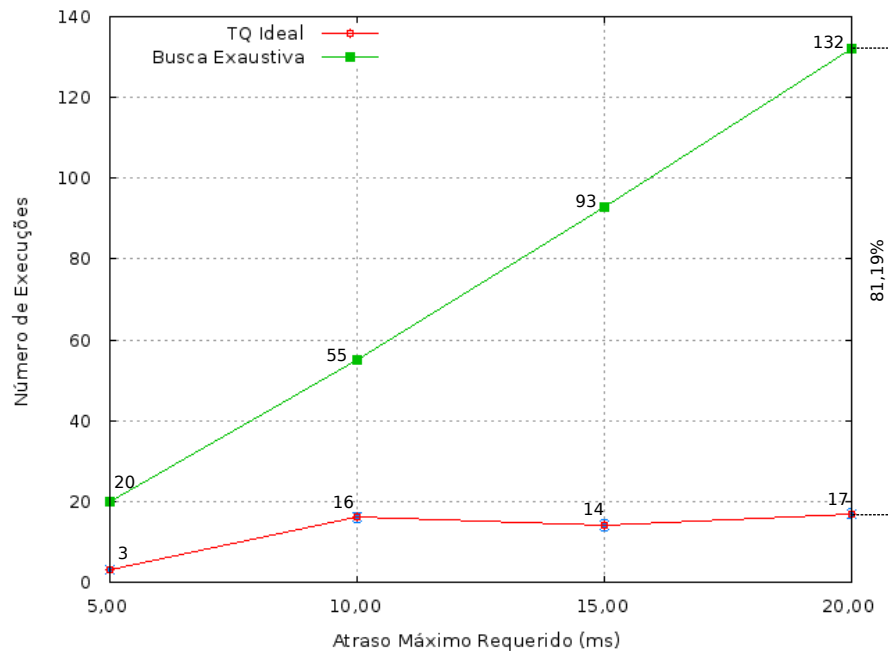
Além disso, dois outros fatores influenciam positivamente para a utilização do PSO nesta nova arquitetura, em relação ao cálculo do tempo do quadro com a busca exaustiva: (i) o número de execuções para calcular o valor do  $TQ$  Ideal, é menor; e (ii) o tempo de execução do algoritmo também é menor, gerando um ganho em tempo de execução para efetuar o cálculo da duração do  $TQ$  Ideal. Os resultados são mostrados a seguir.

Iniciando pelo número de execuções para o tráfego de áudio, o gráfico da Figura 36 mostra que, com a utilização do PSO, é possível alcançar um melhor desempenho para a nova arquitetura, com uma redução no número de execuções em média de 76,45% em relação ao método da busca exaustiva.



**Figura 36: Número de execuções para encontrar o  $TQ$  Ideal para o tráfego de áudio**  
**Fonte: Autoria própria.**

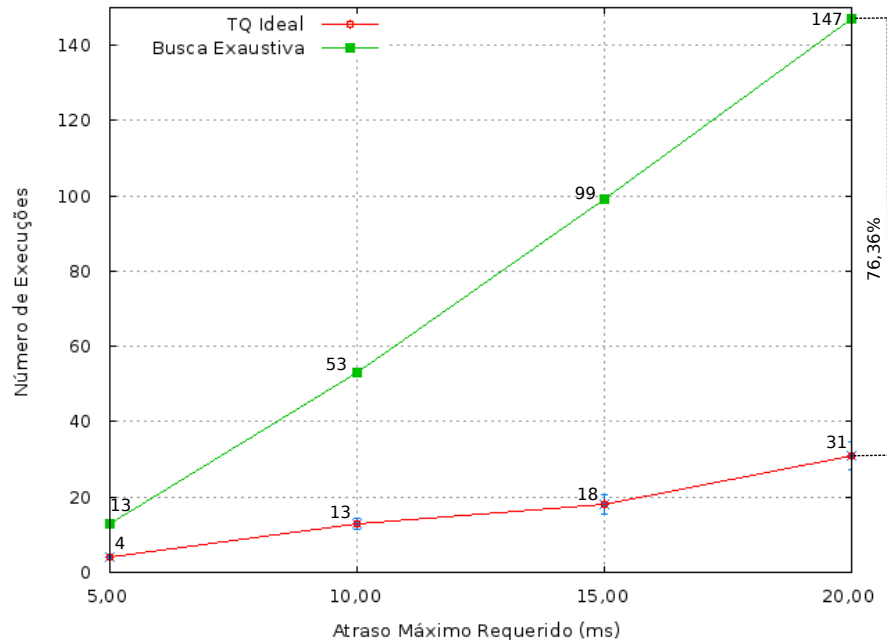
Para o tráfego de vídeo VBR, o gráfico da Figura 37 mostra que, com a utilização da ferramenta de otimização PSO, é possível reduzir o número de execuções em média de 81,19%, quando comparado com o método da busca exaustiva.



**Figura 37: Número de execuções para encontrar o  $TQ$  Ideal para o tráfego de vídeo VBR**  
**Fonte: Autoria própria.**

Com a utilização da ferramenta de otimização PSO, para o tráfego de vídeo MPEG4, é

possível alcançar um melhor desempenho com a nova arquitetura, com uma redução no número de execuções em média de 76,36% em relação ao método da busca exaustiva, conforme mostra o gráfico da Figura 38.



**Figura 38:** Número de execuções para encontrar o  $TQ$  Ideal para o tráfego de vídeo MPEG4  
**Fonte:** Autoria própria.

Desta forma, pode-se concluir que, com a utilização de uma ferramenta de otimização, o número de execuções para calcular o valor do  $TQ$  Ideal é diminuído, possibilitando que a rede tenha um melhor rendimento, alocando todo o espaço disponível do quadro na transmissão dos pacotes, gerando uma economia no tempo de execução do algoritmo, como pode ser verificado a seguir.

Com o tempo de execução do algoritmo PSO para o cálculo do  $TQ$  Ideal, também ocorre uma redução para todos os tipos de tráfego utilizados na avaliação.

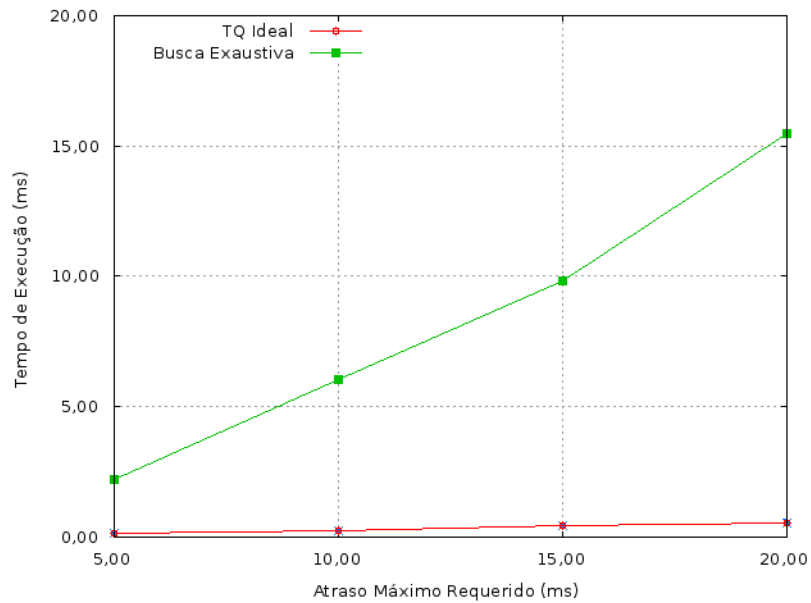
Na Tabela 12 todos os valores utilizados para a mostrar o desempenho do cálculo do  $TQ$  Ideal, em relação ao cálculo do tempo do quadro com o método da busca exaustiva.

**Tabela 12:** Tempos para calcular o quadro com PSO e com o método de busca exaustiva no tráfego de áudio

**Fonte:** Autoria própria.

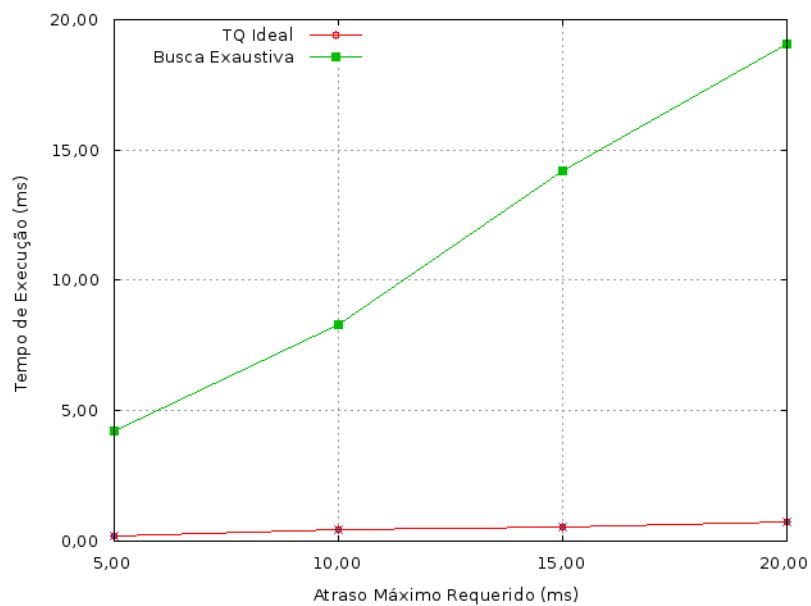
Atraso máximo requerido $D_{max,i}$ (ms)	$TQ$ Ideal (ms)	Busca Exaustiva (ms)
5,00	0,16	2,20
10,00	0,25	6,04
15,00	0,44	9,83
20,00	0,56	15,49

O gráfico da Figura 39 mostra o tempo de execução para calcular o quadro com o PSO e com o método da busca exaustiva para o tráfego de áudio.



**Figura 39:** Tempo de execução para encontrar o  $TQ$  Ideal para o tráfego de áudio  
**Fonte:** Autoria própria.

Para o tráfego de vídeo VBR, também ocorre a redução no tempo de execução. O gráfico da Figura 40 mostra o tempo de execução em relação ao tempo para calcular o quadro no método da busca exaustiva.



**Figura 40:** Tempo de execução para encontrar o  $TQ$  Ideal para o tráfego de vídeo VBR  
**Fonte:** Autoria própria.

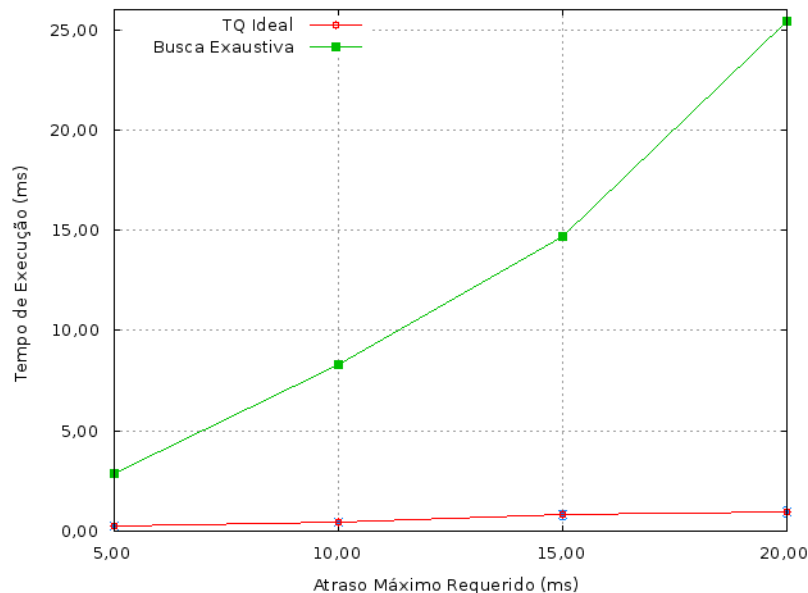
A Tabela 13 resume os valores utilizados para a confecção do gráfico da Figura 40.

**Tabela 13:** Tempos para calcular o quadro com PSO e com o método de busca exaustiva no tráfego de vídeo VBR

**Fonte: Autoria própria.**

Atraso máximo requerido $D_{max,i}$ (ms)	$TQ$ Ideal (ms)	Busca Exaustiva (ms)
5,00	0,19	4,23
10,00	0,44	8,33
15,00	0,55	14,20
20,00	0,75	19,06

Para o tráfego de vídeo MPEG4, também ocorre a redução no tempo de execução para o cálculo do tempo do quadro, utilizando o PSO em relação ao método da busca exaustiva. O gráfico da Figura 41 mostra esta relação.



**Figura 41:** Tempo de execução para encontrar o  $TQ$  Ideal para o tráfego de vídeo MPEG4

**Fonte: Autoria própria.**

A Tabela 14 resume os valores utilizados para construção do gráfico da Figura 41.

**Tabela 14:** Tempos para calcular o quadro com PSO e com o método de busca exaustiva no tráfego de vídeo MPEG4

**Fonte: Autoria própria.**

Atraso máximo requerido $D_{max,i}$ (ms)	$TQ$ Ideal (ms)	Busca Exaustiva (ms)
5,00	0,24	2,85
10,00	0,44	8,31
15,00	0,80	14,66
20,00	0,96	25,45

Baseando-se nos resultados acima, pode-se concluir que com a utilização da metaheurística PSO para o cálculo do  $TQ$  Ideal, ocorre um ganho de desempenho em relação ao número de execuções e ao tempo de execução, quando medido com o cálculo do tempo

do quadro utilizando o método da busca exaustiva, o que indica uma redução do esforço computacional.

### 5.1.2 CENÁRIO 2 - UTILIZAÇÃO DO $TQ$ IDEAL PARA VALIDAÇÃO DA NOVA ARQUITETURA

O cenário 2 apresenta os resultados da simulação com a utilização do  $TQ$  Ideal já calculado. O ambiente é modelado e avaliado através da ferramenta de simulação de redes de computadores NS-2, na versão NS-2.33, em conjunto com o módulo que implementa a camada MAC do padrão IEEE 802.16, desenvolvido em (BORIN; FONSECA, 2008). Os resultados obtidos são comparados com o trabalho desenvolvido em (JAKIMOSKI; JANEVSKI, 2013). As simulações realizadas têm como objetivo verificar as propriedades da nova arquitetura e analisar seu comportamento em uma rede com canal em condições ideais. A Tabela 15 apresenta as características do modelo de rede utilizado.

**Tabela 15:** Principais parâmetros de configuração da rede

Fonte: Autoria própria.

Parâmetro	Valor
Arquitetura da Rede	PMP
Duplexação	TDD
Camada Física	OFDM
Modulação	BPSK
Duração do Tempo do Quadro	calculado em cada iteração
Capacidade do Canal	40 Mbps
Tempo do Símbolo OFDM	13,89 $\mu$
Prefixo Cíclico	0,25
Fator de Amostragem	144/125
Fluxos de Dados	UGS, ertPS, rtPS, nrtPS e BE
Taxa UL/DL	1:1

O cenário de simulação utilizado consiste de uma BS e até 70 SSs, sendo 10 SSs com tráfego UGS, 10 SSs com tráfego rtPS, 10 SSs com tráfego nrtPS, 10 SSs com tráfego BE, e as SSs com tráfego ertPS variam de 5-30 durante a simulação, a fim de testar as performances sob diferentes cargas de tráfego ertPS. A BS está localizada no centro de uma área de 250 x 250 m, conectada a um nó fixo por meio de um enlace de 100 Mbps e retardo de 2 ms, assim como em (JAKIMOSKI; JANEVSKI, 2013), com as SSs distribuídas no entorno da BS, de maneira aleatória. Cada SS tem um fluxo *uplink* e um fluxo *downlink*, que são mapeados para o mesmo tipo de serviço.

Neste cenário, cada SS possui apenas uma aplicação. As aplicações são representadas pelas fontes de tráfego de voz, voz com supressão de silêncio, vídeo, FTP e *Web* que são servidas, respectivamente pelos serviços UGS, ertPS, rtPS, nrtPS e BE.

O tráfego de voz, utilizado pela classe de serviço UGS, é modelado por uma fonte *on/off* com distribuição exponencial e médias iguais a 1,2 e 1,8 segundos, respectivamente, com pacotes de 66 *bytes* gerados a cada 20 *ms* durante o período *on* (BRADY, 1969). Para simular a classe de serviço ertPS um simulador EVRC (*Enhanced Variable Rate Codec*) com um modelo de voz com supressão de silêncio (3GPP2, 2004), com pacotes gerados a cada 20 *ms* empregando taxa 1 (171 bits/pacote), taxa 1/2 (80 bits/pacote), taxa 1/4 (40 bits/pacote) ou taxa 1/8 (16 bits/pacote). O tráfego de vídeo, utilizado pela classe de serviço rtPS, foi obtido a partir de *traces* de vídeos reais, com as características apresentadas na Tabela 16.

**Tabela 16:** Características dos *traces* de vídeos utilizados nas simulações  
Fonte: (SEELING et al., 2004)

<b>vídeo</b>	<b>taxa de pico (Mbps)</b>	<b>Variância do tamanho (Mbps)</b>	<b>Hurst dos frames</b>
Aladdin	2.242800	1.217128	0.876541
Friends vol4 DVD	1.175520	1.820629	0.818063
Gupta Lecture	2.218080	1.398482	0.777175
Baseball	1.859760	1.284597	0.684708
MTV	2.976240	1.448513	0.842930
Star Wars V	1.720800	1.492767	0.767742
Lady Tramp DVD	0.910080	1.763723	0.813108
Ice Age DVD	1.058400	1.532578	0.834963
Dreamcatcher DVD	0.932400	1.400108	0.700975

O tráfego de transferência de arquivos via FTP é gerado utilizando mensagens de comprimentos exponencialmente distribuídos com média de 512 *KB*. O tráfego *Web* é gerado a partir de uma distribuição híbrida *Lognormal/Pareto*, cujo corpo da distribuição corresponde à uma área de 0,88 e é modelado por uma distribuição *Lognormal* com média igual a 7247 *bytes*, enquanto que a cauda é modelada por uma distribuição *Pareto* com média de 10558 *bytes* (BARFORD et al., 1998). O intervalo entre concessões de dados periódicos para os serviços UGS e ertPS é de 20 *ms*, pois, de acordo com o padrão IEEE 802.16, a BS deve alocar concessões para esses serviços em intervalos iguais aos intervalos em que a aplicação gera os pacotes (IEEE, 2012). O intervalo entre *polling unicast* (concessões periódicas) para o envio de mensagens de requisição de banda é de 20 *ms* para o serviço rtPS e de 1 *s* para os fluxos nrtPS. O requisito de atraso máximo limitado para os fluxos da classe rtPS é de 100 *ms* e o requisito de banda mínima varia para cada conexão de acordo com a taxa média do vídeo transmitido, conforme mostrado na Tabela 16. Os fluxos nrtPS tem requisito de banda mínima de 200 *kbps* e os fluxos BE não tem nenhum requisito de QoS.

Para cada valor testado referente à taxa de chegada das conexões, foram executadas 15 replicações com sementes diferentes, para gerar o intervalo de confiança de 95%. Cada simulação teve duração de 3600 *s*.



### 5.1.2.1 RESULTADOS OBTIDOS COM A SEGUNDA AVALIAÇÃO

No trabalho desenvolvido em (JAKIMOSKI; JANEVSKI, 2013), a duração do tempo do quadro é fixa, ou seja, não tem variação durante a execução das simulações. Uma das propriedades da nova arquitetura é otimizar o cálculo da duração do  $TQ$  Ideal, com a metaheurística PSO, para melhorar o desempenho da rede. Assim, cada vez que o escalonador é utilizado, o módulo  $PSO\_TQ$  descrito no Algoritmo 5, no Capítulo 4, é executado para calcular a duração do  $TQ$  Ideal.

Desta forma, o objetivo deste experimento é investigar o impacto da nova arquitetura, com o valor do  $TQ$  Ideal calculado, com diferentes cargas de tráfego para a classe de serviço ertPS, enquanto que as outras quatro classes de serviço tem a mesma carga, variando o número de conexões ertPS de 5 a 30, com um incremento de 5 conexões. São simulados 6 cenários, utilizando o simulador de redes NS-2 (NS-2, 2015), com 10 conexões para cada uma das classes de serviço, e o número de conexões ertPS sendo de 5, 10, 15, 20, 25 e 30. Os resultados são apresentados em relação ao atraso médio, a taxa útil média e a variação média do atraso.

O cálculo para o  $TQ$  Ideal é realizado, com a utilização dos mesmos parâmetros dos experimentos de (JAKIMOSKI; JANEVSKI, 2013), especificados na Tabela 15, sendo que a duração do quadro é de 5 ms para todos os cenários daquele trabalho. O valor calculado da duração do  $TQ$  Ideal, relacionado na Tabela 17, tem uma variação conforme o número de conexões ertPS que está sendo escalonada.

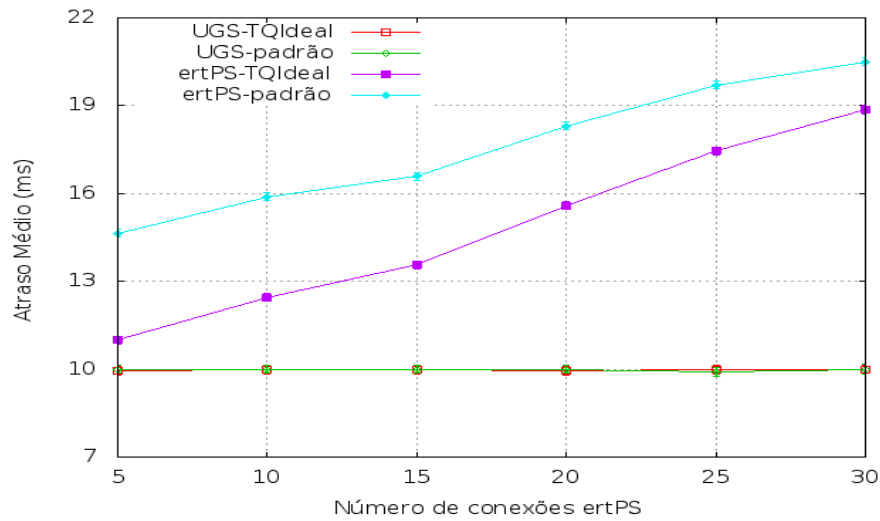
**Tabela 17:** Valor do  $TQ$  Ideal em relação ao tempo do quadro de (JAKIMOSKI; JANEVSKI, 2013)

Fonte: Autoria própria.

Número de conexões ertPS	Tempo do Quadro (ms) (JAKIMOSKI; JANEVSKI, 2013)	$TQ$ Ideal (ms)	% Redução com o uso do $TQ$ Ideal
5	5,00	3,75	25,00%
10	5,00	3,92	21,60%
15	5,00	4,09	18,20%
20	5,00	4,26	14,80%
25	5,00	4,43	11,40%
30	5,00	4,60	8,00%

A Figura 42 apresenta o atraso médio de uma quantidade de conexões ertPS que variam de 5 a 30 e a média de atraso de 10 conexões UGS, em todos os 6 casos com as diferentes cargas de tráfego ertPS. UGS-padrão e ertPS-padrão, na Figura 42, é o resultado obtido com o mecanismo de escalonamento do trabalho de (JAKIMOSKI; JANEVSKI, 2013). UGS-TQIdeal e ertPS-TQIdeal é o resultado obtido com a nova arquitetura.

Utilizando a nova arquitetura com o cálculo do  $TQ$  Ideal, os resultados, com base nas conexões ertPS, são melhores. Para cada um dos 6 cenários, um valor de  $TQ$  Ideal foi



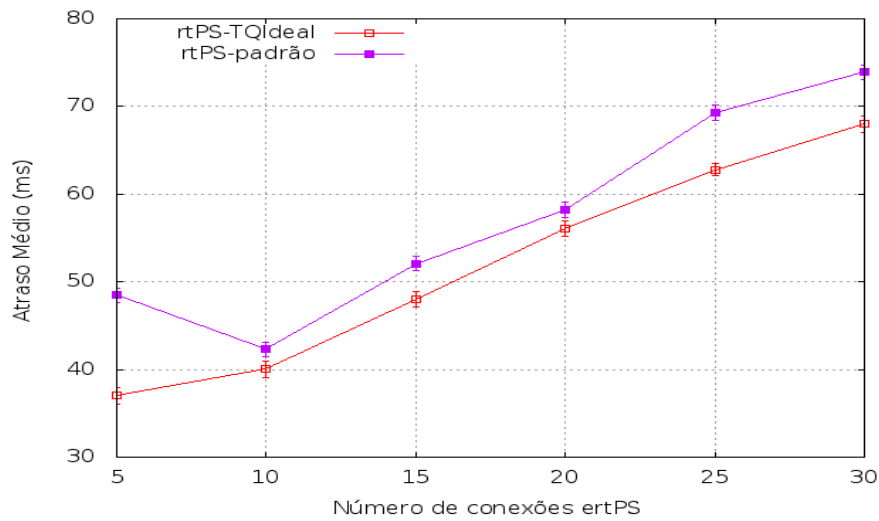
**Figura 42: Atraso médio de 10 conexões UGS com diferentes cargas de tráfego ertPS**

**Fonte: Autoria própria.**

calculado, conforme apresentado na Tabela 17. Assim, na Figura 42 percebe-se que, conforme a carga de tráfego das conexões ertPS for aumentado e o valor do  $TQ$  Ideal for aumentando, as curvas vão ficando mais próximas, porém, existe uma otimização quando da utilização do  $TQ$  Ideal calculado, porque todo o quadro foi utilizado para transportar as conexões que estavam na fila de escalonamento. Por exemplo, com a nova arquitetura, quando a classe de serviço ertPS está com 5 conexões, o  $TQ$  Ideal calculado é de  $3,75\ ms$ , sendo  $25\ %$  menor do que valor do tempo do quadro utilizado no trabalho de (JAKIMOSKI; JANEVSKI, 2013), que é de  $5,00\ ms$ . O atraso médio das conexões UGS não é afetado após a implementação da nova arquitetura, pois o aumento da carga de tráfego ertPS não tem qualquer influência sobre o atraso médio das conexões UGS em ambos os casos, pelo fato de o tráfego de dados da classe de serviço UGS estar na fila com prioridade alta, juntamente com o tráfego de dados da classe de serviço ertPS. O intervalo de confiança é bastante reduzido, o que mostra que a variação do atraso dos pacotes transmitidos por conexões associadas a esses dois tipos de serviço foi pequena e, portanto, a nova arquitetura foi capaz de alocar oportunidades de transmissão periodicamente como especificado no padrão IEEE 802.16.

A Figura 43, apresenta o atraso médio de 10 conexões rtPS, com as conexões ertPS variando de 5 a 30 conexões. Pelo fato de a classe de serviço rtPS estar com prioridade intermediária na fila de escalonamento e com prioridade 3 para as suas conexões, o atraso médio é influenciado pelo aumento da carga de tráfego de dados da classe de serviço ertPS, que está na mesma fila de prioridade intermediária para o escalonamento, mas tem prioridade 2 para as suas conexões. Com o cálculo da duração do  $TQ$  Ideal, a nova arquitetura mantém um aumento do atraso médio em crescimento uniforme para as conexões rtPS, mas sem extrapolar

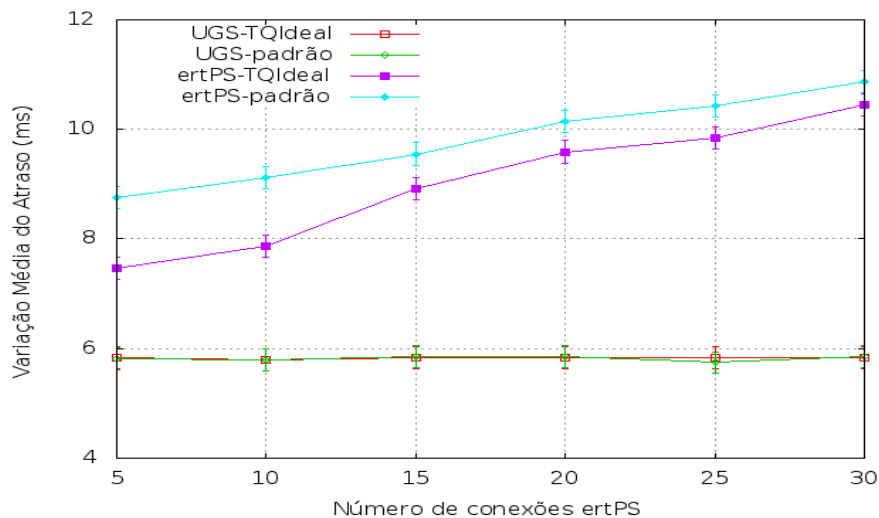
o requisito de atraso máximo para esta classe de serviço, que é de 100 *ms*.



**Figura 43:** Atraso médio de 10 conexões rtPS com diferentes cargas de tráfego ertPS

Fonte: Autoria própria.

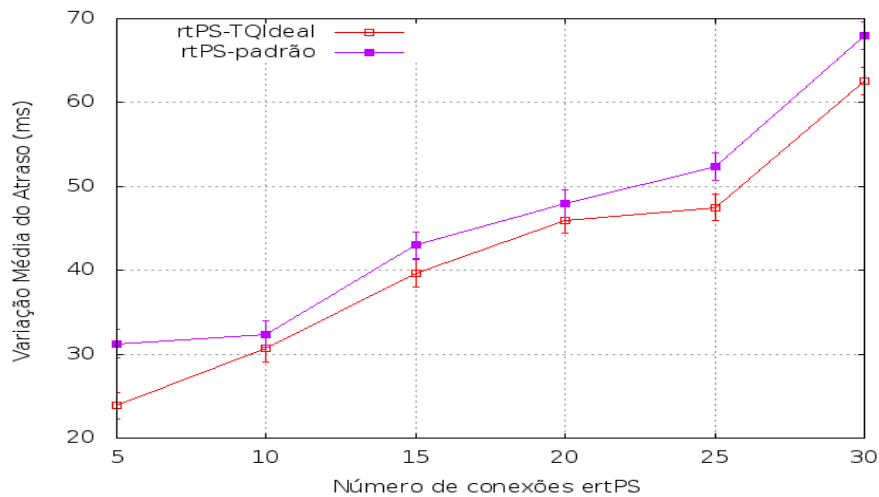
A Figura 44, apresenta o resultado da variação média do atraso para o tráfego de dados das classes de serviço UGS e ertPS. Similar aos resultados obtidos com o atraso médio, os resultados para a variação média do atraso da classe de serviço ertPS são melhores com a utilização da nova arquitetura, quando comparados com o trabalho de (JAKIMOSKI; JANEVSKI, 2013). O resultado da variação média do atraso para a classe de serviço UGS, em ambos os casos, não é afetado pelo aumento da carga de tráfego das conexões ertPS, tendo seus valores próximos de 6 *ms* em todos os cenários simulados, em ambos os casos.



**Figura 44:** Variação média do atraso de 10 conexões UGS e variação média do atraso de conexões ertPS de 5 para 30 estações clientes (SSs)

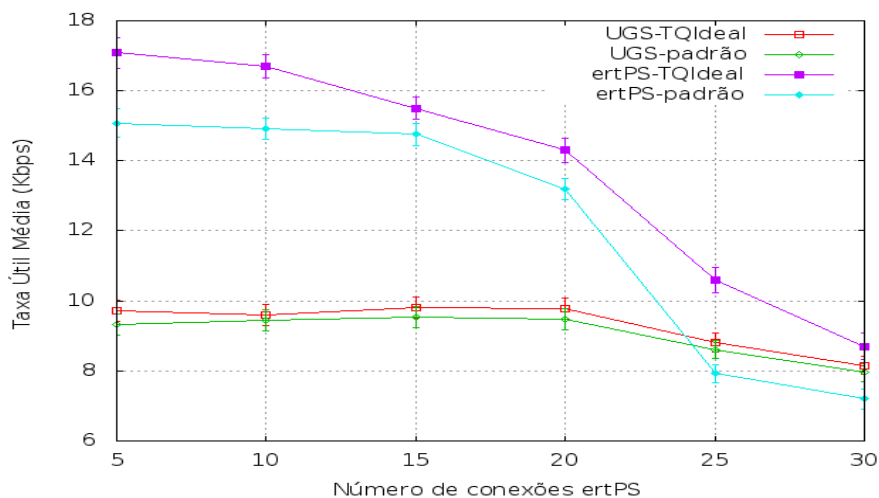
Fonte: Autoria própria.

A Figura 45, apresenta o resultado da variação média do atraso para o tráfego de 10 conexões rtPS, com as conexões ertPS variando de 5 para 30 conexões. Pode-se observar que a nova arquitetura tem uma variação média do atraso menor do que o trabalho de (JAKIMOSKI; JANEVSKI, 2013), pelo fato de que os valores utilizados para o tempo do quadro serem otimizados na nova arquitetura, gerando um ganho de desempenho na variação média do atraso.



**Figura 45: Variação média do atraso de 10 conexões rtPS com diferentes cargas de tráfego ertPS**  
**Fonte: Autoria própria.**

A Figura 46, apresenta o resultado da taxa útil média de 10 conexões UGS quando o número de conexões ertPS varia de 5 para 30. Também, no mesmo gráfico, o resultado da taxa útil média das conexões ertPS. O número de conexões das classes de serviço rtPS, nrtPS e BE são de 10 para todos os casos simulados.

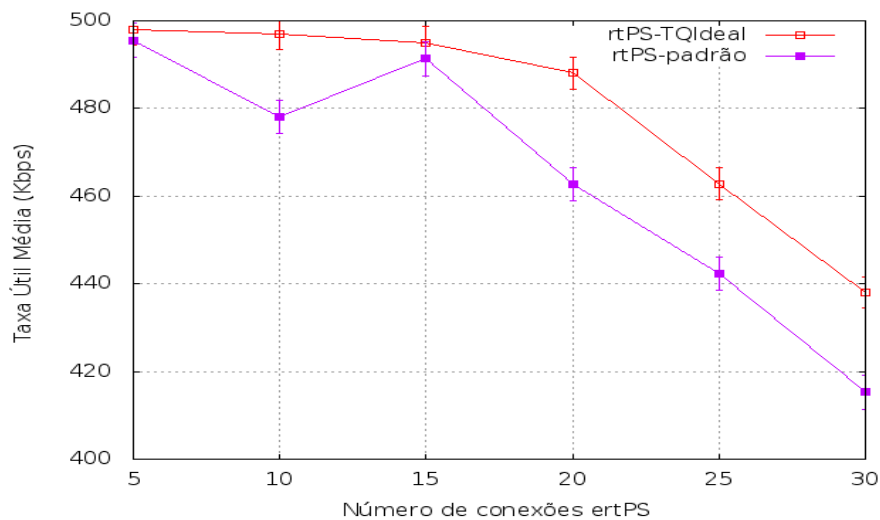


**Figura 46: Taxa útil média de 10 conexões UGS e taxa útil média de conexões ertPS variando de 5 para 30 estações clientes (SSs)**

**Fonte: Autoria própria.**

O que pode ser observado na Figura 46 é que os resultados com a nova arquitetura tem um melhor desempenho no quesito de taxa útil para as classes UGS e ertPS para todas as seis diferentes cargas de tráfego. Isso se deve ao fato de que a nova arquitetura tem um melhor desempenho quando o número de conexões ertPS aumenta no sistema. Este caso também se repete quando comparados com as conexões rtPS, nrtPS e BE, mostrados nas Figuras 47 e 48.

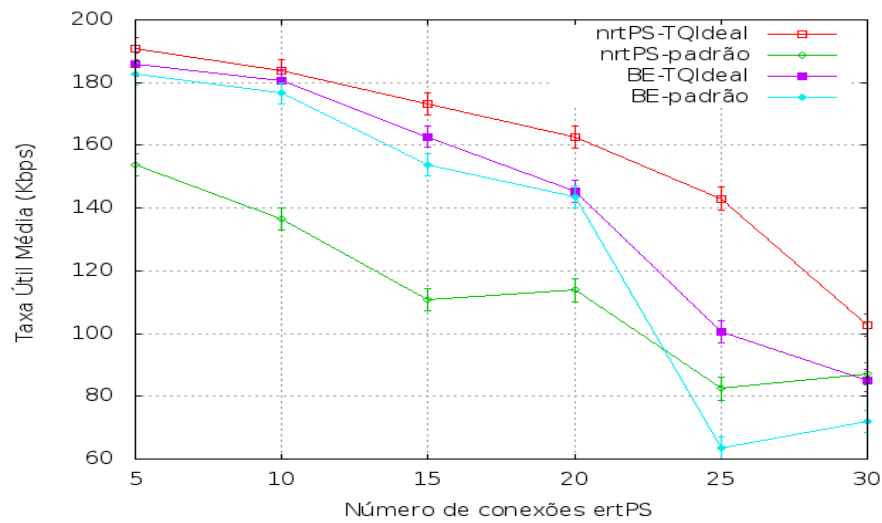
Além de comparar os resultados da taxa útil média das 10 conexões rtPS da nova arquitetura com o esquema de escalonamento do trabalho de (JAKIMOSKI; JANEVSKI, 2013), na Figura 47, verifica-se que a taxa mínima de tráfego e taxa máxima de tráfego requeridas, estão de acordo com o que é especificado no padrão para redes WiMAX (IEEE, 2012). Em todos os cenários simulados para obtenção da taxa útil média, a taxa mínima requerida de 300 *kbps* para as conexões rtPS e a taxa máxima requerida de 500 *kbps*, estão dentro da faixa, independentemente da carga de tráfego das conexões ertPS.



**Figura 47: Taxa útil média de 10 conexões rtPS com diferentes cargas de tráfego ertPS**

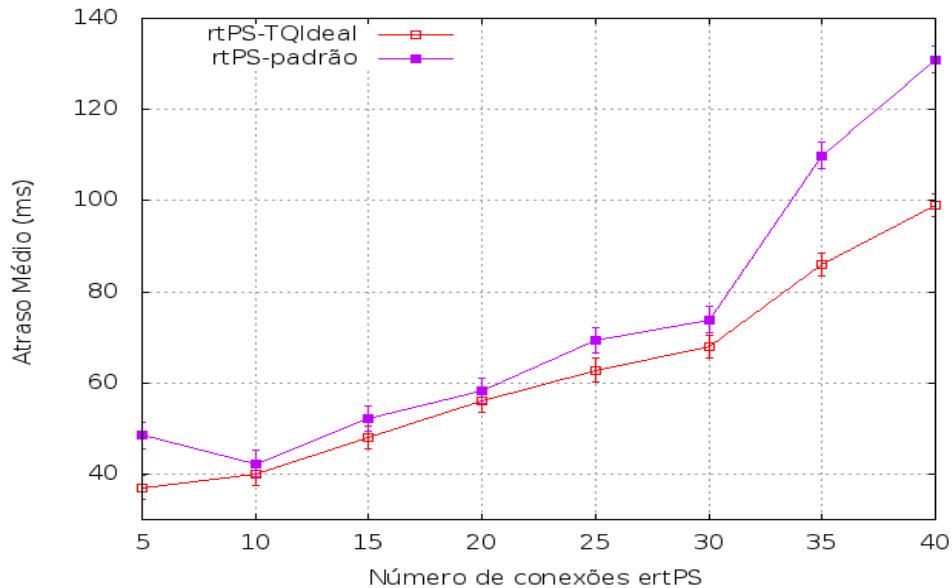
Fonte: Autoria própria.

As observações finais que podem ser feitas das Figuras 47 e 48 são as mesmas discutidas na Figura 46, por apresentarem os resultados médios da taxa útil de 10 conexões rtPS, 10 conexões nrtPS e 10 conexões BE, com diferentes cargas para as conexões ertPS. Comparando-se o impacto do aumento do número de conexões ertPS sobre as outras quatro classes de serviço, pode-se concluir que o aumento do número de conexões ertPS tem um menor impacto sobre a classe de serviço UGS. Esta situação pode ser explicada pelo fato de que a classe de serviço UGS também estar na fila de atendimento de alta prioridade, assim como as conexões da classe de serviço ertPS e, pacotes a partir destas classes de serviço, ter mais prioridade do que os pacotes das classes de serviço rtPS, nrtPS e BE.



**Figura 48:** Taxa útil média de 10 conexões nrtPS e 10 conexões BE com diferentes cargas de tráfego ertPS  
**Fonte:** Autoria própria.

Por fim, a Figura 49 mostra uma evolução do número de estações clientes com conexões da classe de serviço ertPS, até que o  $TQ$  Ideal seja de  $5,00\ ms$ , tendo, desta forma, o mesmo tempo do quadro do trabalho de (JAKIMOSKI; JANEVSKI, 2013).



**Figura 49:** Aumento do número de conexões ertPS  
**Fonte:** Autoria própria.

Os resultados mostram que com o valor do  $TQ$  Ideal em  $5,00\ ms$ , a nova arquitetura suporta 40 conexões ertPS com garantia do atraso máximo para as conexões rtPS, que é de  $100\ ms$ . Portanto, este experimento tem um cenário com 10 conexões rtPS com as cargas das conexões ertPS variando de 5 até 40. As classes de serviço UGS, nrtPS e BE tem 10 conexões

cada. Pode-se observar que com o aumento do número de conexões ertPS, além das 30 estações dos cenários anteriores, o trabalho de (JAKIMOSKI; JANEVSKI, 2013), já com 35 conexões ertPS, não garante o atraso máximo para as conexões rtPS.

A Tabela 18 mostra um resumo do gráfico apresentado na Figura 49, destacando os valores calculados para o  $TQ$  Ideal, que com 35 conexões ertPS é de 4,77 ms, tendo uma redução de 4,60 % e, para as 40 conexões ertPS, o tempo do quadro é de 5,00 ms para ambos. Também destaca-se, na Tabela 18 a garantia do atraso máximo, que a nova arquitetura garante com o aumento das conexões ertPS para 35, com o valor de 85,89 ms e com 40 conexões ertPS, com o valor de 98,19 ms. Já no trabalho de (JAKIMOSKI; JANEVSKI, 2013), com 35 conexões ertPS, o valor do atraso médio é de 109,89 ms, e para 40 conexões ertPS, o atraso médio é de 130,91 ms, não garantido, em ambos os casos, o atraso máximo para as conexões rtPS que é de 100 ms. Na tabela, a coluna com **Tempo do Quadro Padrão**, se refere ao tempo do quadro do trabalho de (JAKIMOSKI; JANEVSKI, 2013), assim como **Atraso Médio TQ Padrão** se refere ao atraso médio daquele trabalho.

**Tabela 18:** Resumo do gráfico apresentado na Figura 49, com o valor do atraso médio

Fonte: Autoria própria.

Número de conexões ertPS	Tempo do Quadro Padrão (ms)	Atraso Médio (ms) TQ Padrão	$TQ$ Ideal (ms)	Atraso Médio (ms) do $TQ$ Ideal
5	5,00	48,49	3,75	37,01
10	5,00	42,38	3,92	40,08
15	5,00	52,16	4,09	48,13
20	5,00	58,20	4,26	56,04
25	5,00	69,33	4,43	62,80
30	5,00	73,95	4,60	67,95
35	5,00	109,89	4,77	85,89
40	5,00	130,91	5,00	98,19

## 5.2 CONCLUSÃO DO CAPÍTULO

Este capítulo apresentou uma análise do desempenho da nova arquitetura, através da modelagem, avaliação e simulação, utilizando dois cenários. No cenário 1, o cálculo da duração do  $TQ$  Ideal é realizado e seu desempenho é comparado com o cálculo do tempo do quadro através do método da busca exaustiva. O resultado obtido com o cálculo da duração do  $TQ$  Ideal, utilizando PSO, demonstrou ser bastante eficiente. Em relação a duração do tempo do quadro não houve redução, porém em relação ao número de execuções para encontrar o  $TQ$  Ideal e ao tempo de execução, houve uma redução bastante significativa, conforme mostrado no Capítulo 5, em relação ao cálculo do tempo do quadro utilizando o método da busca exaustiva.

No cenário 2, o ambiente foi modelado e simulado através da ferramenta de simulação

de redes de computadores NS-2, na versão NS-2.33. As simulações realizadas tem como objetivo verificar as propriedades da nova arquitetura e analisar seu comportamento em uma rede com canal em condições ideais, ou seja, sem perda ou alteração dos pacotes. Os resultados obtidos nesta simulação foram comparados com o trabalho de (JAKIMOSKI; JANEVSKI, 2013), utilizando os mesmos parâmetros, tendo somente a duração do tempo do quadro como diferencial. A nova arquitetura utilizou o cálculo do  $TQ$  Ideal e encontrou, conforme a carga utilizada, diferentes valores para o  $TQ$  Ideal, conforme mostrado na Tabela 17. No trabalho de (JAKIMOSKI; JANEVSKI, 2013) a duração do tempo do quadro é fixa em 5,00 *ms* durante toda a simulação.

Assim, com a utilização do cálculo do  $TQ$  Ideal e com o estabelecimento de prioridades par ao escalonamento, os resultados obtidos provaram que a nova arquitetura tem melhores resultados para o atraso médio, taxa útil média e variação média do atraso, em comparação com o mecanismo de escalonamento *uplink* do trabalho de (JAKIMOSKI; JANEVSKI, 2013).

Também foi avaliado se o aumento do número de conexões ertPS provocaria um impacto sobre as outras quatro classes de serviço. Nos dois trabalhos comparados, ficou claro que este aumento da carga de tráfego ertPS, quando comparado com a classe de serviço UGS, é insignificante em ambos os casos, especialmente para atraso médio e variação média do atraso, porque a classe de serviço UGS está na fila de alta prioridade. O impacto sobre as outras classes de serviço é perceptível em alguns casos no trabalho de (JAKIMOSKI; JANEVSKI, 2013), mas na nova arquitetura os resultados são melhores para o atraso médio, taxa útil média e variação média do atraso, especialmente quando o número de conexões ertPS é maior.

Na última etapa de simulações do cenário 2, foi simulado o atraso médio de 10 conexões rtPS em relação a um aumento da carga de conexões ertPS para os dois modelos, chegando a 40 conexões estabelecidas. A nova arquitetura conseguiu garantir o atraso máximo de 100 *ms*, o que não ocorreu com o trabalho de (JAKIMOSKI; JANEVSKI, 2013).

Os resultados alcançados demonstram que com o cálculo de um  $TQ$  Ideal, o sistema pode se beneficiar, evitando que os quadros trafeguem com espaços vazios, alocando mais usuários em cada quadro e evitando o desperdício de banda. Os resultados da análise de desempenho demonstraram que a nova arquitetura fornece garantias de atraso máximo limitado para as aplicações de tempo real e largura de banda mínima para as aplicações de tempo real e não tempo real, sendo compatível com o padrão IEEE 802.16.



## 6 CONCLUSÕES

O padrão IEEE 802.16 é uma solução para as redes de acesso de banda larga sem fio, que também é conhecido comercialmente como WiMAX. Uma de suas grandes vantagens é a especificação, na camada MAC, de mecanismos de provisão de QoS. Porém, o padrão deixa em aberto algumas questões, para que os fabricantes de dispositivos possam diferenciar seus produtos além de permitir que a comunidade científica realize pesquisas nesta área.

Na norma definida pelo padrão IEEE 802.16 (IEEE, 2012), a camada física (PHY) e a camada de controle de acesso ao meio (MAC), estão definidas como sendo orientadas à conexão, e que a comunicação entre os elementos da rede, realizado pela estação base (BS) e estação cliente (SS), é gerenciada através de fluxos de serviços, que são mapeados e atendidos de acordo com os requisitos de QoS das classes de serviço UGS, ertPS, rtPS, nrtPS e BE. Assim, o padrão é capaz de prover QoS para classes heterogêneas com diferentes requisitos de QoS. O tráfego de tempo real é mapeado para as classes de serviço UGS, ertPS ou rtPS, e o tráfego não tempo real para as classes de serviço nrtPS ou BE. O tráfego de tempo real possui requisitos de QoS específicos e mais críticos, como o atraso máximo, e por isso precisa de tratamento diferenciado, com priorização de tráfego.

Face ao exposto, este trabalho de pesquisa tem seu foco no desenvolvimento de uma nova arquitetura para o tráfego no sentido *uplink* para a BS em redes IEEE 802.16, tendo como principal função a provisão de QoS para as cinco classes de serviço especificadas pelo padrão IEEE 802.16, com a garantia de recursos para as classes com tráfego em tempo real e com garantia de recursos mínimos necessários para as classes com tráfego que não necessitem ser em tempo real. A nova arquitetura é composta por três etapas de processamento, executadas de forma hierárquica, e denominadas de *Camada 1*, *Camada 2* e *Camada 3*.

A *Camada 1* é o componente principal da nova arquitetura, e tem como objetivo calcular um tempo ideal para o quadro através das duas equações de restrição (10) e (11) que são descritas na Seção 4.2. Esta abordagem utiliza a metaheurística de otimização por enxame de partículas (PSO). Com o cálculo da duração do *TQ* Ideal, o sistema é otimizado

pelo fato de que mais usuários podem transmitir seus dados dentro de um mesmo quadro. A *Camada 2* trata do problema da alocação dos recursos disponíveis para transmissões no sentido *uplink*. Esses recursos devem ser divididos considerando a capacidade de transmissão de cada segmento no subquadro de *uplink* e, então, alocados para oportunidades de requisição de banda ou de transmissão de dados pelas SSs. A *Camada 3* é responsável pelo gerenciamento das filas de requisições de banda no sentido *uplink* enviadas pelas SSs nos quadros anteriores, que estão armazenadas e aguardam atendimento pela BS.

A apresentação e a análise dos resultados obtidos a partir da nova arquitetura foram efetuadas por meio de modelagem, avaliação e simulação, com a utilização de dois cenários. No cenário 1, o cálculo da duração do *TQ* Ideal é realizado e seu desempenho é comparado com o cálculo do tempo do quadro utilizando o método da busca exaustiva.

O resultado obtido com o cálculo da duração do *TQ* Ideal, utilizando PSO, demonstrou ser bastante eficiente. Em relação a duração do tempo do quadro não houve redução, porém em relação ao número de execuções para encontrar o *TQ* Ideal e ao tempo de execução, houve uma redução bastante significativa, conforme mostrado no Capítulo 5, em relação ao cálculo do tempo do quadro com busca exaustiva.

No cenário 2, o ambiente foi modelado e avaliado através da ferramenta de simulação de redes de computadores NS-2, na versão NS-2.33, em conjunto com o módulo que implementa a camada MAC do padrão IEEE 802.16, desenvolvido em (BORIN; FONSECA, 2008). As simulações realizadas tem como objetivo verificar as propriedades da nova arquitetura e analisar seu comportamento em uma rede com canal em condições ideais, ou seja, sem perda ou alteração dos pacotes. Os resultados obtidos nesta simulação foram comparados com o trabalho de (JAKIMOSKI; JANEVSKI, 2013), utilizando os mesmos parâmetros, tendo somente a duração do tempo do quadro como diferencial. A nova arquitetura utilizou o cálculo do *TQ* Ideal e encontrou, conforme a carga utilizada, diferentes valores para o *TQ* Ideal, conforme mostrado na Tabela 17. No trabalho de (JAKIMOSKI; JANEVSKI, 2013) a duração do tempo do quadro é fixa em 5,00 *ms* durante toda a simulação.

Os resultados obtidos demonstraram que a nova arquitetura, com a utilização do cálculo do *TQ* Ideal e com a utilização do mecanismo de escalonamento baseado em prioridades, tem melhores resultados para o atraso médio, taxa útil média e variação média do atraso, em comparação com o mecanismo de escalonamento *uplink* do trabalho de (JAKIMOSKI; JANEVSKI, 2013). Também foi analisado o impacto do aumento do número de conexões ertPS sobre as outras quatro classes de serviço. Nos dois trabalhos, ficou claro que o impacto do aumento da carga de tráfego ertPS na classe de serviço UGS é insignificante

em ambos os casos comparados, especialmente para atraso médio e variação média do atraso, porque a classe de serviço UGS está na fila de alta prioridade. O impacto sobre as outras classes de serviço é perceptível em alguns casos no trabalho de (JAKIMOSKI; JANEVSKI, 2013), mas na nova arquitetura os resultados são melhores para o atraso médio, taxa útil média e variação média do atraso, especialmente quando o número de conexões ertPS é maior.

Por fim, foi simulado o atraso médio de 10 conexões rtPS em relação a um aumento da carga de conexões ertPS para os dois modelos, chegando a 40 conexões estabelecidas. A nova arquitetura conseguiu garantir o atraso máximo de 100 *ms*, o que não ocorreu com o trabalho de (JAKIMOSKI; JANEVSKI, 2013).

Desta forma, os resultados alcançados demonstram que com o cálculo de um *TQ* Ideal, o sistema pode se beneficiar, evitando que os quadros trafeguem com espaços vazios, alocando mais usuários em cada quadro e evitando o desperdício de banda. Os resultados da análise de desempenho demonstraram que a nova arquitetura fornece garantias de atraso máximo limitado para as aplicações de tempo real e largura de banda mínima para as aplicações de tempo real e não tempo real, sendo compatível com o padrão IEEE 802.16.

## 6.1 TRABALHOS FUTUROS

Alguns tópicos para o desenvolvimento de trabalhos futuros:

- Incorporar o módulo que realiza o cálculo do *TQ* Ideal, desta nova arquitetura, ao módulo WiMAX, desenvolvido em (BORIN; FONSECA, 2008), no simulador NS-2 (NS-2, 2015);
- Estender o modelo analítico para a interface física OFDMA para utilizar com redes WiMAX móveis;
- Testar o modelo da nova arquitetura com outras metaheurísticas.

## REFERÊNCIAS

3GPP2. *Enhanced Variable Rate Codec, Speech Service Option 3 for Wideband Spread Spectrum Digital Systems*. Lucent Technologies, 2004. 144 p.

AHMED, Z.; HAMMA, S. Efficient and fair scheduling of rtPS traffic in IEEE 802.16 point-to-multipoint networks. In: IFIP-IEEE. **Wireless and Mobile Networking Conference - WMNC'2011 4th Joint IFIP**. Toulouse, 2011. v. 1, p. 1–7.

ALRASHIDI, M.; EL-HAWARY, M. A survey of particle swarm optimization applications in electric power systems. **Evolutionary Computation, IEEE Transactions on**, v. 13, n. 4, p. 913–918, 2009.

ALSAHAG, A. et al. Fair bandwidth assignment in hierarchical scheduling for mobile WiMAX system. In: IEICE-IEEE. **Asia-Pacific Conference on Communications - APCC'2011**. Sabah, 2011. p. 743–747.

ANATEL. **Portal da Anatel**. <http://www.anatel.gov.br/Portal/exibirPortalInternet.do>, Acessado em 17 de março 2015.

ANDREWS, J. G.; GHOSH, A.; MUHAMED, R. **Fundamentals of WiMAX: Understanding broadband wireless networking**. Nova Jersey: Prentice Hall, 2007.

ARHAIF, M. S. Comparative study of scheduling algorithms in WiMAX. **International Journal of Scientific & Engineering Research**, v. 2, n. 2, p. 87–93, 2011.

AWAL, M. A.; BOUKHATEM, L. **WiMAX and end-to-end QoS support**. Book chapter in: Radio Resources Management in WiMAX: From theoretical capacity to system simulations, 2009. 105-136 p.

BACIOCCOLA, A. et al. IEEE 802.16: History, status and future trends. **Computer Communications**, v. 33, n. 2, p. 113–123, 2010.

BARFORD, P. et al. **Changes in Web Client Access Patterns: Characteristics and Caching Implications**. Boston, 1998. 30 p.

BELGHITH, A.; NUAYMI, L. **Scheduling techniques for WiMAX**. Book chapter in: Current Technology Developments of WiMax Systems, 2009. 61-84 p.

BERGH, F. van den. **An analysis of particle swarm optimizers**. 300 p. Tese (Doutorado) — Faculty of Natural and Agricultural Science, University of Pretoria, 2002.

BORIN, J.; FONSECA, N. da. Um módulo para simulação de redes WiMAX no simulador NS-2. In: SBC. **Congresso da Sociedade Brasileira de Computação**. Belém, 2008. v. 1, p. 180–199.

- BORIN, J.; FONSECA, N. da. Uplink scheduler and admission control for the IEEE 802.16 standard. In: IEEE. **Global Telecommunications Conference, 2009. GLOBECOM'09. IEEE**. Honolulu, 2009. p. 1–6.
- BORIN, J. F. **Mecanismos para provisão de qualidade de serviços em redes IEEE 802.16**. 104 p. Tese (Doutorado) — Instituto de Computação, Universidade Estadual de Campinas, 2010.
- BRADY, P. T. A model for generating on-off speech patterns in two-way conversations. **Bell System Technical Journal**, v. 48, p. 2445–2472, 1969.
- CHENG, S. T.; HSIEH, M. T.; CHEN, B. Fairness-based scheduling algorithm for time division duplex mode IEEE 802.16 broadband wireless access systems. **Communications, IET**, v. 4, n. 9, p. 1065–1072, 2010.
- CICCONETTI, C. et al. Performance evaluation of the IEEE 802.16 MAC for QoS support. **IEEE Transactions on Mobile Computing**, v. 6, n. 1, p. 26–38, 2007.
- COATH, G.; HALGAMUGE, S. A comparison of constraint-handling methods for the application of particle swarm optimization to constrained nonlinear optimization problems. **The Congress on Evolutionary Computation, 2003. CEC '03**, v. 4, p. 2419–2425, 2003.
- DORIGO, M.; BIRATTARI, M. Swarm intelligence. **Scholarpedia**, v. 2, n. 9, p. 1462, 2007.
- DOSCIATTI, E. R. **Um novo escalonador com controle de admissão de conexão para o padrão IEEE 802.16 com garantia de limite de atraso**. Dissertação (Mestrado) — UTFPR, 2010.
- DOSCIATTI, E. R.; GODOY, W.; FORONDA, A. An efficient approach of scheduling with call admission control to fixed WiMAX networks. **Latin America Transactions, IEEE (Revista IEEE America Latina)**, v. 10, n. 1, p. 1256–1264, 2012.
- EBERHART, R.; SIMPSON, P.; DOBBINS, R. **Computational intelligence PC tools**. San Diego: Academic Press Professional, 1996.
- ENGELBRECHT, A. P. **Computational intelligence: An introduction**. 2. ed. West Sussex: John Wiley and Sons, 2002.
- FERRARI, D.; VERMA, D. A scheme for real-time channel establishment in wide-area networks. **IEEE Journal on Selected Areas in Communications**, v. 8, n. 3, p. 368–379, 1990.
- FERREIRA, F. A. **Uma proposta de escalonamento baseado na disciplina Priority Queuing (PQ) para redes IEEE 802.16**. 118 p. Dissertação (Mestrado) — Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, 2011.
- FUKUYAMA, Y. **Fundamentals of particle swarm optimization techniques**. Book chapter in: *Modern Heuristic Optimization Techniques*, 2007. 71-87 p.
- GAING, Z.-L. Particle swarm optimization to solving the economic dispatch considering the generator constraints. **IEEE Transactions on Power Systems**, v. 18, n. 3, p. 1187–1195, 2003.
- GLOVER, F. Tabu search - part I. **ORSA Journal on Computing**, v. 1, n. 3, p. 190–206, 1989.

- GLOVER, F. Tabu search - part II. **ORSA Journal on Computing**, v. 2, n. 1, p. 4–32, 1990.
- GOMES, L. A. **Método baseado em lógica nebulosa para alocação de geradores distribuídos sob a óptica do perfil de tensão**. 112 p. Dissertação (Mestrado) — Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, 2009.
- GONCALVES, A. R. **Otimização em ambientes dinâmicos com variáveis contínuas empregando algoritmos de estimação de distribuição**. 132 p. Dissertação (Mestrado) — Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, 2011.
- HOYMANN, C. Analysis and performance evaluation of the OFDM-based metropolitan area network IEEE 802.16. **Computer Networks**, v. 49, n. 3, p. 341–363, 2005.
- HU, X.; EBERHART, R. Solving constrained nonlinear optimization problems with particle swarm optimization. In: CERNET. **6th World Multiconference on Systemics, Cybernetics and Informatics -SCI'02**. Orlando, 2002. p. 203–206.
- IEEE. IEEE Standard for Air Interface for Broadband Wireless Access Systems. **IEEE Std 802.16-2012 (Revision of IEEE Std 802.16-2009)**, p. 1–2542, 2012.
- JAKIMOSKI, K.; JANEVSKI, T. Priority based uplink scheduling scheme for WiMAX service classes. **International Journal Information Technology and Computer Science, IJITCS**, v. 08, n. 1, p. 66–74, 2013.
- KENNEDY, J.; EBERHART, R. C. Particle swarm optimization. **IEEE Proceedings of the IEEE International Joint Conference on Neural Networks**, v. 4, p. 1942–1948, 1995.
- KENNEDY, J.; EBERHART, R. C.; SHI, Y. **Swarm intelligence**. San Francisco: Morgan Kaufmann, 2001.
- KERNIGHAN, B. W.; RITCHIE, D. M. **C - A linguagem de programação padrão ANSI**. São Paulo: Campus-Elsevier, 1989.
- KOLIAS, C.; KAMBOURAKIS, G.; GRITZALIS, S. Attacks and countermeasures on 802.16: Analysis and assessment. **Communications Surveys Tutorials, IEEE**, v. 15, n. 1, p. 487–514, 2013.
- KUROSE, J. F.; ROSS, K. W. **Redes de computadores e a Internet: Uma abordagem top-down**. 5. ed. São Paulo: Addison Wesley, 2010.
- LAARHOVEN, P. V.; AARTS, E. **Simulated annealing: Theory and applications**. Eindhoven: Kluwer Academic Publishers, 1987.
- LEE, K.; PARK, J.-B. Application of particle swarm optimization to economic dispatch problem: Advantages and disadvantages. In: IEEE PES. **Power Systems Conference and Exposition, 2006. PSCE '06**. Atlanta, 2006. p. 188–192.
- LIU, Y.; PASSINO, K. M. **Swarm intelligence: Literature overview**. 2000. Literature Overview.
- LOVBERG, M. **Improving particle swarm optimization by hybridization of stochastic search heuristics and self organized critically**. 80 p. Dissertação (Mestrado) — Department of Computer Science, University of Aarhus, 2002.

MA, M.; LU, J. **QoS provision mechanisms in WiMAX**. Book chapter in: *Current Technology Developments of WiMAX Systems*, 2009. 85-114 p.

MASRI, M.; ABDELLATIF, S.; JUANOLE, G. An uplink bandwidth management framework for IEEE 802.16 with QoS guarantees. In: *IFIP. 8th International IFIP-TC 6 Networking Conference, NETWORKING'09*. Aachen, 2009. v. 5550, p. 651–663.

MICHALEWICZ, Z.; FOGEL, D. B. **How to solve it: Modern heuristics**. 2. ed. Berlin: Springer Berlin Heidelberg, 2013.

MILLONAS, M. M. **Swarms, phase transitions and collective intelligence**. Book chapter in: Addison-Wesley, 1994. 1-43 p.

MORAES, H. B. **Proposta de um mecanismo de policiamento de tráfego baseado em token bucket para redes IEEE 802.16**. 124 p. Dissertação (Mestrado) — Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, 2013.

MSADAA, I.; CAMARA, D.; FILALI, F. Scheduling and CAC in IEEE 802.16 fixed BWNs: A comprehensive survey and taxonomy. **Communications Surveys Tutorials, IEEE**, v. 12, n. 4, p. 459–487, 2010.

NS-2. **Network Simulator 2**. <http://www.isi.edu/nsnam/ns>, Acessado em 17 de março 2015.

OLIVEIRA, A. N. de. **Um estudo de escalonamento uplink considerando a camada física OFDM do padrão IEEE 802.16**. 118 p. Dissertação (Mestrado) — Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, 2011.

OLIVEIRA, S.; VILLELA, L.; R., M. Estudo da difusão tecnológica no setor de telecomunicações - tecnologia WiMAX. **Revista de Gestão e Operações Produtivas**, v. 3, n. 3, p. 1–12, 2012.

PAPAPANAGIOTOU, I. et al. A survey on next generation mobile WiMAX networks: Objectives, features and technical challenges. **Communications Surveys Tutorials, IEEE**, v. 11, n. 4, p. 3–18, 2009.

PARSOPOULOS, K. E.; VRAHATIS, M. N. **Particle swarm optimization and intelligence: Advances and applications**. Hershey: IGI Global, 2010.

PRASAD, R.; VELEZ, F. J. **WiMAX networks: Techno-economic vision and challenges**. Heidelberg: Springer, 2010.

RATNAWEERA, A.; HALGAMUGE, S.; WATSON, H. Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. **Evolutionary Computation, IEEE Transactions on**, v. 8, n. 3, p. 240–255, 2004.

RINI, D. P.; SHAMSUDDIN, S. M.; YUHANIZ, S. S. Particle swarm optimization: Technique, system and challenges. **International Journal of Computer Applications**, v. 14, p. 19–27, 2011.

ROSA, E. C.; GUARDIERO, P. R. CAC and uplink scheduling algorithms in WiMAX networks. **Revista Telecomunicações**, v. 13, n. 2, p. 32–39, 2011.

- RUANGCHAIJATUPON, N.; WANG, L.; JI, Y. A study on the performance of scheduling schemes for broadband wireless access networks. In: **ISCIT. International Symposium on Communications and Information Technologies, 2006. ISCIT '06.** Bangkok, 2006. p. 1008–1012.
- SAYENKO, A.; ALANEN, O.; HÄMÄLÄINEN, T. Scheduling solution for the IEEE 802.16 base station. **Computer Networks**, v. 52, n. 1, p. 96–115, 2008.
- SCHOOFS, L.; NAUDTS, B. Swarm intelligence on the binary constraint satisfaction problem. **Proceedings of the 2002 Congress on Evolutionary Computation, 2002. CEC'02**, v. 2, p. 1444–1449, 2002.
- SEELING, P.; REISSLEIN, M.; KULAPALA, B. Network performance evaluation using frame size and quality traces of single-layer and two-layer video: A tutorial. **Communications Surveys Tutorials, IEEE**, v. 6, n. 3, p. 58–78, 2004.
- SETTEMBRE, M. et al. Performance analysis of an efficient packet-based IEEE 802.16 MAC supporting adaptive modulation and coding. In: **IEEE. 7th International Symposium on Computer Networks, 2006 - ISCN'06.** Istanbul, 2006. p. 11–16.
- SHA, D.; HSU, C.-Y. A hybrid particle swarm optimization for job shop scheduling problem. **Computers & Industrial Engineering**, v. 51, n. 4, p. 791–808, 2006.
- SILVA, S. P. da; BIONDI, A. **Caminhos para a universalização da internet banda larga: Experiências internacionais e desafios brasileiros.** São Paulo: Intervezes - Coletivo Brasil de Comunicação Social, 2012.
- SPALL, J. C. **Introduction to stochastic search and optimization.** Nova Jersey: John Wiley & Sons Ltd., 2003.
- STILIADIS, D.; VARMA, A. Latency-rate servers: A general model for analysis of traffic scheduling algorithms. **IEEE/ACM Transactions on Networking**, v. 6, n. 5, p. 611–624, 1998.
- TAGHIPOOR, M.; MJAFARI, S.; HOSSEINI, V. **Quality of service and resource allocation in WiMAX.** Rijeka: InTech, 2012.
- TANG, S.; MULLER, P.; SHARIF, H. **WiMAX security and quality of service: An end-to-end perspective.** Nova Jersey: John Wiley & Sons, 2010.
- TELES, M. L.; GOMES, H. M. Comparação de algoritmos genéticos e programação quadrática sequencial para otimização de problemas em engenharia. **Teoria e Prática na Engenharia Civil**, v. 15, n. 1, p. 29–39, 2010.
- TSAI, T.-C.; JIANG, C.-H.; WANG, C.-Y. CAC and packet scheduling using token bucket for IEEE 802.16 networks. **Journal of Communications**, v. 1, n. 2, p. 30–37, 2006.
- VALLE, Y. del et al. Particle swarm optimization: Basic concepts, variants and applications in power systems. **IEEE Transactions on Evolutionary Computation**, v. 12, n. 2, p. 171–195, 2008.
- VENTER, G.; SOBIESZCZANSKI-SOBIESKI, J. Particle swarm optimization. In: **ARC. 43rd AIAA/ASME/ASCE/AHA/ASC Structures, Structural Dynamics and Materials.** Denver, 2002.



VINAY, K. et al. Performance evaluation of end-to-end delay by hybrid scheduling algorithm for QoS in IEEE 802.16 network. In: IFIP-IEEE. **International Conference on Wireless and Optical Communications Networks**. Bangarole, 2006. p. 5 pp. –5.

WANG, C.; LIN, H.-C.; LO, H.-K. Adaptive admission control algorithm in IEEE 802.16e broadband wireless access networks. In: IEEE. **6th World Congress on Services - SERVICES-1**. Miami, 2010. p. 273–279.

WIMAX, F. **WiMAX Forum**. <http://www.wimaxforum.org/>, Acessado em 17 de março 2015.

WONGTHAVARAWAT, K.; GANZ, A. Packet scheduling for QoS support in IEEE 802.16 broadband wireless access systems. **International Journal of Communication Systems**, v. 16, n. 1, p. 81–96, 2003.

ZHANG, Y. **WiMAX network planning and optimization**. London: CRC Press, 2009.

ZUBEN, F. J. V.; ATTUX, R. R. F. **Inteligência de enxame**. 2008. 80 p. DCA/FEEC/Unicamp e DECOM/FEEC/Unicamp.

## APÊNDICE A - OBTENÇÃO DA EQUAÇÃO (9)

Substituindo (3) em (5) tem-se

$$D_i \leq \frac{BT'_i - L'_{max,i}}{r_i} + T_{DL} + T_{UL} + \frac{L'_{max,i}}{R} + T_{TTG} + T_{RTG}, \quad (27)$$

onde  $L'_{max,i}$  é  $L_{max,i}$  com a sobrecarga de (7) e  $BT'_i$  é o tamanho do balde de fichas com a mesma sobrecarga.

Substituindo (8) em (27) tem-se

$$D_i \leq \frac{BT'_i - L'_{max,i}}{r'_i - \frac{\Delta * R}{TQ_i}} + T_{DL} + T_{UL} + \frac{L'_{max,i}}{R} + T_{TTG} + T_{RTG}, \quad (28)$$

onde  $r'_i$  é o total da taxa alocada pelo servidor.

Então, considerando a transmissão de pelo menos um pacote  $\frac{L'_{max,i}}{TQ_i}$  tem-se

$$D_i \leq \frac{BT'_i - L'_{max,i}}{r'_i - \frac{\Delta * R}{TQ_i} + \frac{L'_{max,i}}{TQ_i}} + T_{DL} + T_{UL} + \frac{L'_{max,i}}{R} + T_{TTG} + T_{RTG}. \quad (29)$$

Assim, (9) é

$$D_i \leq \frac{(BT'_i - L'_{max,i}) * TQ_i}{r'_i * TQ_i - \Delta * R + L'_{max,i}} + T_{DL} + T_{UL} + \frac{L'_{max,i}}{R} + T_{TTG} + T_{RTG}. \quad (30)$$