

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

**Método de Avaliação de Qualidade de Serviço em Sistemas
Multimídia Utilizando um Sistema Genético Nebuloso
Baseado em Regras**

Fabiana Cristina Bertoni

São Carlos
Agosto 2005

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

B547ma

Bertoni, Fabiana Cristina.

Método de avaliação de qualidade de serviço em sistemas multimídia utilizando um sistema genético nebuloso baseado em regras / Fabiana Cristina Bertoni. -- São Carlos : UFSCar, 2006.

103 p.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2005.

1. Sistemas multimídia. 2. Qualidade de serviço. 3. Sistema genético nebuloso. 4. Algoritmos genéticos. I. Título.

CDD: 006.7 (20^a)

Universidade Federal de São Carlos

Centro de Ciências Exatas e de Tecnologia

Programa de Pós-Graduação em Ciência da Computação

“Método de Avaliação de Qualidade de Serviço em Sistemas Multimídia utilizando um Sistema Genético nebuloso baseado em Regras”

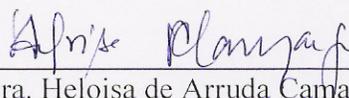
FABIANA CRISTINA BERTONI

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

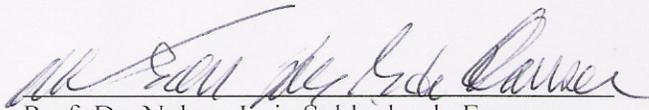
Membros da Banca:



Prof. Dr. Sérgio Donizetti Zorzo
(Orientador – DC/UFSCar)



Profa. Dra. Heloisa de Arruda Camargo
(DC/UFSCar)



Prof. Dr. Nelson Luis Saldanha da Fonseca
(IC/UNICAMP)

São Carlos
Agosto/2005

Adquire pois a sabedoria. Emprega tudo o que possuis na aquisição de conhecimento e entendimento. Exalta-a, e ela te exaltará, e, abraçando-a tu, ela te honrará. Dará à tua cabeça uma diadema de graça e uma coroa de glória te entregará.

Provérbios 4, 7-9

Agradecimentos

Agradeço a DEUS, no nome de JESUS, por permitir que eu vencesse todas as dificuldades e conquistasse mais esta vitória, e, acima de tudo, pela família e vida maravilhosas que me deste.

Aos meus pais e avó materna que estiveram sempre presentes em todos os momentos de minha vida, fossem eles bons ou ruins, apoiando-me e incentivando-me a alcançar todos os objetivos aos quais eu me propunha.

Agradeço de maneira especial a meu namorado Matheus que muitas vezes se privou de seu tempo e paciência e me ajudou a compor este trabalho. Que tinha sempre um abraço apertado, muito carinho e palavras doces para me dizer.

À dona Hilda, que me acolheu em sua casa como uma mãe; e às meninas da pensão (Roberta, Simone, Vanessa, Claudia, Ester), que se tornaram minha segunda família, como irmãs que nunca tive.

Um agradecimento ao meu orientador, Sérgio Donizetti Zorzo, que como a própria designação diz, mostrou-me de forma honrosa como proceder na composição deste trabalho, dividindo dificuldades e comemorando vitórias.

A Prof^a. Dr^a. Heloísa de Arruda Camargo, que me auxiliou no desenvolvimento deste trabalho, sempre disposta a esclarecer minhas dúvidas e me ajudar a superar problemas.

A todos os meus amigos do laboratório de Sistemas Distribuídos e Redes (Bel, Jéssica, Renatinho, Warta, Bocudo, Crazy, Bruno, Eugênio, Loko e Shimoda), que me motivaram e me apoiaram sempre que necessitei da ajuda de cada um deles.

A “galera do AP” (Plotze, Lucas, Tiagão e Jean), com os quais compartilhei um delicioso e modesto Big Mac...

Aos amigos das demais áreas de pesquisa deste programa de mestrado, pelo companheirismo a mim dedicado: Banco de Dados (Cláudia, Luciene, Dinho, Escovar, Genta, Evandro e Karina), Engenharia de Software (Val, Vini, André, Ronaldo, Lucrédio, Biligui, Rafael, Clô e Milena), Inteligência Artificial (Pablo e Ricardo) e Processamento de Imagem (Luciana, Taciana, Mairum, Moacir e Rar).

Sumário

LISTA DE FIGURAS	IV
LISTA DE TABELAS	V
RESUMO	VI
ABSTRACT	VII
CAPÍTULO 1 INTRODUÇÃO	1
1.1 MOTIVAÇÃO	1
1.2 JUSTIFICATIVA	1
1.3 OBJETIVO	2
1.4 ORGANIZAÇÃO DO TRABALHO	3
CAPÍTULO 2 SISTEMAS MULTIMÍDIA	4
2.1 COMPONENTES OU SUBSISTEMAS DE SISTEMAS DE COMUNICAÇÃO MULTIMÍDIA	5
2.1.1 APLICAÇÕES CONVERSACIONAIS	5
2.1.2 APLICAÇÕES DE APRESENTAÇÃO	5
2.2 REQUISITOS DAS APLICAÇÕES MULTIMÍDIA	6
2.2.1 ALTAS TAXAS DE TRANSMISSÃO	6
2.2.2 RESTRIÇÕES TEMPORAIS	7
2.2.3 GARANTIA NOS SERVIÇOS OFERECIDOS	8
2.2.4 COMUNICAÇÃO DE GRUPO	9
2.3 CONSIDERAÇÕES FINAIS	9
CAPÍTULO 3 QUALIDADE DE SERVIÇO EM SISTEMAS MULTIMÍDIA	10
3.1 MODELO CONCEITUAL DE QUALIDADE DE SERVIÇO	11
3.1.1 CAMADA DO USUÁRIO	12
3.1.2 CAMADA DE APLICAÇÃO	12
3.1.3 CAMADA DE SISTEMA	13
3.1.4 CAMADA DE REDE	13
3.2 ELEMENTOS NECESSÁRIOS PARA PROVER GARANTIA DE QUALIDADE DE SERVIÇO	14

3.2.1	ESPECIFICAÇÃO DE QoS	15
3.2.2	GERENCIAMENTO DE QoS	15
3.2.3	NEGOCIAÇÃO E RENEGOCIAÇÃO DE QoS	17
3.2.4	SISTEMA DE RESERVA DE RECURSOS	17
3.2.5	CONTROLE DE TRÁFEGO	18
3.3	PARÂMETROS DE QoS	18
3.3.1	PARÂMETROS DE QoS NO NÍVEL DE USUÁRIO	19
3.3.2	PARÂMETROS INTERNOS DE QoS	20
3.4	CONSIDERAÇÕES DE DESEMPENHO PARA DIFERENTES APLICAÇÕES DE ÁUDIO E VÍDEO	23
3.5	CONSIDERAÇÕES FINAIS	25

CAPÍTULO 4 LÓGICA NEBULOSA E SISTEMAS NEBULOSOS **26**

4.1	CONJUNTOS NEBULOSOS E SUA REPRESENTAÇÃO	27
4.1.1	OPERAÇÕES BÁSICAS COM CONJUNTOS NEBULOSOS	29
4.1.2	VARIÁVEIS LINGÜÍSTICAS	30
4.2	SISTEMA NEBULOSO	32
4.2.1	FUZIFICAÇÃO	33
4.2.2	BASE DE CONHECIMENTO NEBULOSO	33
4.2.3	MOTORES DE INFERÊNCIA	34
4.2.4	DESFUZIFICAÇÃO	37
4.3	CONSIDERAÇÕES FINAIS	38

CAPÍTULO 5 ALGORITMOS GENÉTICOS **40**

5.1	MÉTODOS DE SELEÇÃO	43
5.1.1	SELEÇÃO PELA ROLETA	43
5.1.2	SELEÇÃO ELITISTA OU ELITISMO	44
5.2	OPERADORES GENÉTICOS	45
5.2.1	OPERADORES GENÉTICOS PARA CODIFICAÇÃO BINÁRIA	46
5.3	CONSIDERAÇÕES FINAIS	48

CAPÍTULO 6 MÉTODO DE AVALIAÇÃO DE QUALIDADE DE SERVIÇO EM SISTEMAS MULTIMÍDIA UTILIZANDO UM SISTEMA GENÉTICO NEBULOSO BASEADO EM REGRAS **50**

6.1	TRANSFORMAÇÃO DOS PARÂMETROS DE QoS	51
6.2	APLICAÇÃO DO SISTEMA GENÉTICO NEBULOSO BASEADO EM REGRAS NA AVALIAÇÃO DE QoS	55
6.2.1	MÓDULO DE FUZIFICAÇÃO	55
	Aplicação de Videoconferência	57
	Aplicação de Vídeo sob Demanda	58
6.2.2	BASE DE REGRAS	59
	Seleção de Regras Nebulosas por Algoritmos Genéticos	63
6.2.3	MÓDULO DE DESFUZIFICAÇÃO	67
6.3	CONSIDERAÇÕES FINAIS	69

CAPÍTULO 7 EXPERIMENTOS E RESULTADOS **70**

7.1	ESTUDO DE CASO 1: VIDEOCONFERÊNCIA	70
7.2	ESTUDO DE CASO 2: VÍDEO SOB DEMANDA	71
7.3	RESULTADOS DOS ESTUDOS DE CASO	72
7.4	ANÁLISE DA VARIAÇÃO DA QOS	74
7.5	VALIDAÇÃO DO MÉTODO PROPOSTO	78
7.5.1	MÉTRICA NUMÉRICA PROPOSTA POR DRESSLER	78
7.5.2	MÉTRICA NUMÉRICA PROPOSTA POR LUNARDI E DOTTI	80
7.6	ANÁLISE DOS RESULTADOS	82
7.7	CONSIDERAÇÕES FINAIS	83

CAPÍTULO 8 CONCLUSÕES, TRABALHOS RELACIONADOS E TRABALHOS FUTUROS **84**

8.1	CONCLUSÕES	84
8.2	TRABALHOS RELACIONADOS	85
8.3	TRABALHOS FUTUROS	86

REFERÊNCIAS BIBLIOGRÁFICAS **88**

APÊNDICE A: ANÁLISE DAS BASES DE REGRAS DO SISTEMA NEBULOSO **95**

APÊNDICE B: MODELAGEM DO SISTEMA GENÉTICO NEBULOSO BASEADO EM REGRAS **100**

B.1	DIAGRAMA DE CLASSES	101
B.2	DIAGRAMAS DE CASOS DE USO	102

Lista de Figuras

FIGURA 3-1: UM MODELO CONCEITUAL DE QoS	11
FIGURA 3-2: ATRASOS INTRODUZIDOS	22
FIGURA 4-1: REPRESENTAÇÃO TRIANGULAR	28
FIGURA 4-2: REPRESENTAÇÃO TRAPEZOIDAL	29
FIGURA 4-3: REPRESENTAÇÃO DE VARIÁVEL LINGÜÍSTICA	31
FIGURA 4-4: SISTEMA NEBULOSO	32
FIGURA 4-5: MODELO DE INFERÊNCIA DE MAMDANI	36
FIGURA 4-6: MÉTODO DO CENTRO DE MASSA	37
FIGURA 4-7: MÉTODO DO CENTRO DE MÁXIMA	38
FIGURA 5-1: : EXEMPLO DE UMA ROLETA DE SELEÇÃO	44
FIGURA 5-2: CRUZAMENTO DE UM-PONTO	46
FIGURA 5-3: CRUZAMENTO DE TRÊS-PONTOS	47
FIGURA 5-4: CRUZAMENTO UNIFORME	47
FIGURA 5-5: MUTAÇÃO SIMPLES	48
FIGURA 6-1. EXEMPLO DE CÁLCULO DE PERTINÊNCIA	59
FIGURA 6-2: INTERFACE GRÁFICA XNIST NET	61
FIGURA 6-3: IMAGENS NAS TRÊS CATEGORIAS DEFINIDAS PELO SISTEMA	62
FIGURA 6-4: EXEMPLO DO CÁLCULO DE DESFUZIFICAÇÃO	67
FIGURA 7-1: APLICAÇÃO DE VIDEOCONFERÊNCIA	70
FIGURA 7-2: APLICAÇÃO DE VÍDEO SOB DEMANDA	72
FIGURA 7-3: VARIAÇÃO DA QoS – ANÁLISE 1	75
FIGURA 7-4: VARIAÇÃO DA QoS – ANÁLISE 2	75
FIGURA 7-5: VARIAÇÃO DA QoS – ANÁLISE 3	75
FIGURA 7-6: VARIAÇÃO DA QoS – ANÁLISE 4	75
FIGURA 7-7: VARIAÇÃO DA QoS – ANÁLISE 5	75
FIGURA 7-8: VARIAÇÃO DA QoS – ANÁLISE 6	75
FIGURA 7-9: VARIAÇÃO DA QoS – ANÁLISE 7	76
FIGURA 7-10: VARIAÇÃO DA QoS – ANÁLISE 8	76
FIGURA 7-11: VARIAÇÃO DA QoS – ANÁLISE 9	76

Lista de Tabelas

TABELA 2-1.: MÉDIA DAS TAXAS DE DADOS DE ALGUNS SERVIÇOS	7
TABELA 3-1: OBJETIVOS DE DESEMPENHO PARA APLICAÇÕES DE ÁUDIO E VÍDEO	25
TABELA 6-1: VALORES ESPERADOS DE ALGUNS PARÂMETROS DE QoS	52
TABELA 6-2: VALORES DOS PARAMETROS DE QoS	52
TABELA 6-3: VALORES ACEITÁVEIS PARA OS PARÂMETROS DE QoS	53
TABELA 6-4: FAIXAS ACEITÁVEIS PARA OS PARÂMETROS DE QoS	53
TABELA 6-5: VALORES DOS PARÂMETROS DE QoS CONSIDERADOS ACEITÁVEIS	53
TABELA 6-6: MAPEAMENTO DOS PARÂMETROS DE QoS - USUÁRIO/REDE	54
TABELA 6-7: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE ENTRADA VAZÃO	57
TABELA 6-8: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE ENTRADA DELAY	57
TABELA 6-9: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE ENTRADA JITTER	57
TABELA 6-10: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE ENTRADA PERDA DE PACOTES	57
TABELA 6-11: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE ENTRADA VAZÃO	58
TABELA 6-12: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE ENTRADA DELAY	58
TABELA 6-13: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE ENTRADA JITTER	58
TABELA 6-14: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE ENTRADA PERDA DE PACOTES	58
TABELA 6-15: PARÂMETROS DAS FUNÇÕES DE PERTINÊNCIA PARA A VARIÁVEL DE SAÍDA QoS	67
TABELA 7-1: VIDEOCONFERÊNCIA - RESULTADOS DE MEDIDAS DOS PARÂMETROS	71
TABELA 7-2: VÍDEO SOB DEMANDA - RESULTADOS DE MEDIDAS DOS PARÂMETROS	72
TABELA 7-3. VIDEOCONFERÊNCIA - RESULTADOS DE AVALIAÇÃO DE QoS UTILIZANDO UM SGNBR	73
TABELA 7-4: VÍDEO SOB DEMANDA - RESULTADOS DE AVALIAÇÃO DE QoS UTILIZANDO UM SGNBR	74
TABELA 7-5. VIDEOCONFERÊNCIA - ANÁLISE DE QoS SEGUNDO DRESSLER	79
TABELA 7-6: VÍDEO SOB DEMANDA - ANÁLISE DE QoS SEGUNDO DRESSLER	80
TABELA 7-7: VIDEOCONFERÊNCIA - RESULTADOS DE AVALIAÇÃO DE QoS UTILIZANDO MÉTRICAS NUMÉRICAS	81
TABELA 7-8: VÍDEO SOB DEMANDA - RESULTADOS DE AVALIAÇÃO DE QoS UTILIZANDO MÉTRICAS NUMÉRICAS	81

Resumo

A Qualidade de Serviço (QoS) disponível na rede é o principal fator para a garantia do sucesso em uma transmissão multimídia [Dressler 2003a]. No contexto atual dos sistemas multimídia, mostra-se imprescindível sistemas que dêem suporte ao controle de QoS, de forma que as aplicações tenham a qualidade desejada na efetivação do serviço.

O usuário humano de uma aplicação multimídia é o ponto inicial de todas as considerações a respeito de QoS, tendo esta que se preocupar com os efeitos perceptíveis a esse usuário. Sabe-se que ele normalmente não tem condições de especificar parâmetros de baixo nível, como largura de banda, nem mesmo de entender valores numéricos relativos a medições dos mesmos. Assim sendo, é necessário que o sistema forneça um nível de abstração para que o usuário compreenda qual a qualidade de serviço que a rede pode oferecer em um determinado instante. Cabe ao usuário, então, decidir se colocará sua aplicação em execução ou não.

Este trabalho apresenta um novo método para avaliação de QoS em uma conexão multimídia. Esse método utiliza valores discretos dos parâmetros de QoS estabelecendo uma relação entre eles, pelo uso de um Sistema Genético Nebuloso Baseado em Regras, de forma a obter um único valor resultante que expresse o nível de QoS da conexão. Este valor final é classificado dentro de um padrão de especificação de QoS pré-definido, que inclui três conjuntos: Ótimo, Aceitável e Não Tolerado; tendo como resultante deste processo uma definição da Qualidade de Serviço em uma linguagem em nível de usuário.

Abstract

The Quality of Service (QoS) available in the network is the most important parameter for the guarantee of the success in a multimedia transmission [Dressler 2003a]. In the current context of multimedia systems, handle systems to QoS control are essential to the applications have the desired quality in the service accomplishment.

The human user of a multimedia application is the initial point of all the considerations regarding QoS, having this that worry about the perceivable effect to this user. Usually, users are not aware on how to specify low-level parameters, such bandwidth, not even to understand numerical values related to measurements of these. Thus, it is necessary that the system provides an abstraction level, so that the user can understand which is the quality of service which the network can offer in a certain instant. Then, the user can decide if he will place the application in execution or not.

This work presents a new method to express the Quality of Service (QoS) in a multimedia connection. This method uses discrete values of QoS parameters, establishing a relationship among them through the use of a Fuzzy Genetic Rule-Based System (FGRBS) to obtain a single value that expresses the QoS of the connection. This final value is classified within a predefined QoS specification standard consisting of three sets: Optimal, Acceptable and Unacceptable, resulting in a definition of the Quality of Service in a user-level language.

Capítulo 1

Introdução

1.1 Motivação

Observa-se uma grande demanda por uma classe de aplicações conhecidas como Sistemas Multimídia [Mbise e Woods 2003]. Esta classe de sistemas distingue-se das outras pela necessidade de transmissão de dados multimídia, tais como áudio e vídeo, gerando grandes fluxos de dados e acarretando um grande aumento no consumo de recursos nas redes de computadores e sistemas fim [Teixeira e Barbar 2003]. Nestas condições, os sistemas multimídia, dentre outras características, provêm serviços voltados à execução de aplicações que necessitam de garantias de recursos para que seu desempenho não seja prejudicado.

As aplicações multimídia requerem uma variedade de níveis de qualidade das redes de comunicação, para que possam ser apresentadas ao usuário de forma satisfatória. Para que isso ocorra, existe a alternativa de tomar conhecimento da qualidade de serviço (QoS) oferecida pela rede para aplicações deste tipo, fazendo uma Provisão de QoS [Kamienski e Sadok 2000], que é o objeto de estudo deste trabalho.

1.2 Justificativa

A Qualidade de Serviço é uma necessidade cada vez mais recorrente para aplicações multimídia, em função do aumento do uso das mesmas na Internet [Kamienski e Sadok 2000],

sendo necessário um mecanismo auxiliar que avalie esta rede, objetivando qualidade no serviço oferecido.

Os mecanismos existentes não devolvem uma resposta ao usuário em uma linguagem de fácil entendimento e nem estão preparados para o tratamento de “verdades parciais”, como por exemplo, a classificação de algum serviço como “*está bom, mas não muito bom*”. A utilização do Sistema Genético Nebuloso Baseado em Regras possibilitará esses dois fatores.

1.3 Objetivo

O principal foco de pesquisa deste trabalho é o desenvolvimento de um novo método para avaliar a Qualidade de Serviço em um Sistema Multimídia, utilizando metodologias de Inteligência Computacional para estabelecer pesos para cada parâmetro de QoS e posteriormente unificar estes valores para explicitar o nível de qualidade de uma conexão. Esse método utiliza valores discretos dos parâmetros de QoS estabelecendo uma relação entre eles, pelo uso de um Sistema Genético Nebuloso Baseado em Regras, de forma a obter um único valor resultante que expresse o nível de QoS da conexão. Este valor final é classificado dentro de um padrão de especificação de QoS pré-definido, que inclui três conjuntos: Ótimo, Aceitável e Não Tolerado; tendo como resultante deste processo uma definição da Qualidade de Serviço em uma linguagem acessível aos diversos níveis de usuário e um valor numérico representativo dessa QoS.

Para garantir a validade do método desenvolvido, o valor unificado resultante foi comparado com valores obtidos através de métricas numéricas convencionais e com o resultado observado na análise feita com o uso de um emulador.

Obtendo as definições de QoS para uma classe de problemas, é possível estendê-las a outras aplicações. Dois casos com características distintas foram analisados: uma aplicação de vídeo sob demanda e uma videoconferência. Os resultados destes processos poderão ser utilizados diretamente por usuários como forma de saber qual a qualidade que a rede pode oferecer em um determinado instante, ou por mecanismos que renegociem a QoS contratada, podendo ser mapeados para um framework que faça o gerenciamento dinâmico da qualidade de serviço, não havendo a intersecção do usuário no processo.

1.4 Organização do Trabalho

Esta dissertação de mestrado está organizada da seguinte forma:

O capítulo 2 apresenta definições relacionadas aos Sistemas Multimídia, suas características, requisitos, tipos de aplicações existentes e seus componentes.

No capítulo 3, são mostrados os conceitos ligados à Qualidade de Serviço em Sistemas Multimídia, seu modelo em camadas e seus parâmetros. Este capítulo mostra também um Framework geral de QoS, tratando de sua especificação, gerenciamento, negociação e renegociação, sistema de reserva de recursos, controle de tráfego, e como pode ser provida em diferentes níveis, além de considerações de desempenho para as diversas aplicações.

Nos capítulos 4 e 5, são apresentadas as metodologias utilizadas para o desenvolvimento do trabalho, sendo elas: Sistemas Nebulosos e Algoritmos Genéticos.

O capítulo 6 relata cada fase do desenvolvimento do método proposto, detalhando todas as técnicas e parâmetros utilizados na implementação e validação do mesmo.

Os Experimentos e Resultados são mostrados no capítulo 7, os quais tratam da análise de duas aplicações multimídia: uma aplicação de videoconferência e outra de vídeo sob demanda, utilizando o Sistema Genético Nebuloso Baseado em Regras desenvolvido e outras métricas numéricas convencionais, estabelecendo um comparativo. Uma análise contínua da variação da QoS também é apresentada.

Por fim, o capítulo 8 traz as conclusões e alguns comentários sobre trabalhos relacionados e futuros.

Capítulo 2

Sistemas Multimídia

Sistemas multimídia processam informações que combinam um conjunto de dados de vários tipos de mídia, tais como texto, gráficos, imagens, áudio, vídeo ou informações de controle [Oliveira 2001].

Estes tipos de mídia são classificados em duas categorias: Mídias discretas, onde estão envolvidas apenas dimensões espaciais (textos, hipertextos e imagens) e Mídias contínuas, que envolvem tanto dimensões espaciais como temporais (áudio e vídeos). O armazenamento e distribuição das mídias discretas não acarretam dificuldades devido ao pequeno volume de dados que essas mídias apresentam. Para as mídias contínuas, devemos considerar que o sistema pretende implementar transmissão em tempo real, isso é, os dados de áudio e vídeo serão transmitidos "por demanda". O áudio, apesar de ser uma mídia contínua, apresenta baixas taxas de transmissão, principalmente quando empregado uma técnica de compressão. Já os vídeos digitalizados, sejam imagens em movimento ou animações gráficas, apresentam grandes volumes de dados e, mesmo utilizando técnicas de compressão, a dificuldade em sua transmissão requer certos cuidados [Silveira 2000].

Um sistema de mídia contínua possui uma natureza isocronológica, ou seja, deve ser apresentado regularmente para que não haja danos em sua qualidade, não sendo possível perceber falhas e atrasos, ou quando presentes, que sejam aceitáveis. Esta natureza, no desenvolvimento de aplicações multimídia de grande escala, impõe ao sistema restrições em todos os seus componentes, sejam eles de rede ou de sistemas fim [Teixeira e Barbar 2003]. Por exemplo, uma transmissão de voz em um telefone digital deve ocorrer com uma amostragem de 8 bits a cada 125 milisegundos. Se os próximos 8 bits não forem transmitidos em um tempo menor ou igual a 125 milisegundos, o atraso será humanamente perceptível.

2.1 Componentes ou Subsistemas de Sistemas de Comunicação Multimídia

Aplicações multimídia requerem a transmissão de dados multimídia entre diferentes localizações. Tipicamente, algumas estações geram o dado (em tempo real ou acessando um banco de dados) e outras estações são consumidoras ou apresentadoras dos dados [Silveira 2000]. O termo usado para este tipo de aplicação é conhecido como fim-a-fim e refere-se ao sistema como um todo.

Os componentes de um sistema multimídia dependem do tipo de aplicação que está sendo considerada. Existem duas classes de aplicações multimídia: Aplicações Conversacionais e Aplicações de Apresentação, conforme descrito em [ITU 2001].

2.1.1 Aplicações Conversacionais

Sinais de áudio e vídeo são gerados em tempo real por microfones ou câmeras de vídeo. Estes sinais são convertidos para o formato digital e codificados sendo posteriormente transmitidos através da rede por uma pilha de protocolos de transporte. No destino, os dados são recebidos e decodificados, quando então são convertidos em analógicos para serem apresentados em algum dispositivo [Lu 1996].

Essas aplicações são mais sensíveis ao atraso. As informações são geradas em tempo real e deveriam ser consumidas em tempo real. Como isso geralmente não acontece, dados podem ser perdidos, uma vez que o dado produzido possui um tempo de vida para circular na rede. Exemplos deste tipo de aplicação são o Videofone e a Videoconferência.

2.1.2 Aplicações de Apresentação

Aplicações de Apresentação utilizam informação multimídia armazenada em um ou mais dispositivos de armazenamento de alta capacidade (servidores) e usuários podem acessá-las em tempo real [Bochmann 2002]. A aplicação pode ser sob Demanda, envolvendo o

armazenamento de dados multimídia em servidores que podem ser acessados por um cliente, ou Broadcast, envolvendo apresentação para grupos de pessoas, onde um servidor transmite e os grupos recebem os dados simultaneamente. Em ambas, a informação é composta e codificada off-line, oriundas de dispositivos de armazenamento (servidores). Há dificuldade em determinar tempo de acesso e taxa de transferência, uma vez que o sistema deve fornecer respostas a diversas requisições, com a fonte das informações podendo estar em servidores diferentes.

Baseado nas características descritas acima, é perceptível que o desenvolvimento de Aplicações de Apresentação depende de avançados sistemas de comunicação (redes de alta velocidade), armazenamento (bancos de dados orientados à objeto) e processamento (sistemas operacionais de tempo real).

2.2 Requisitos das Aplicações Multimídia

Aplicações multimídia são caracterizadas por manipular mídias contínuas, suportar diferentes tipos de mídia e seus relacionamentos temporais e prover serviços voltados à execução de aplicações que necessitam de garantias de recursos para que seu desempenho não seja prejudicado. Estas características impõem novos requisitos aos sistemas de comunicação e aos sistemas fim, tais como altas taxas de transmissão, restrições temporais, garantia nos serviços oferecidos e comunicação de grupo.

Estes requisitos são detalhados nos itens a seguir.

2.2.1 Altas Taxas de Transmissão

Devido aos diversos tipos de dados associados em uma transmissão de mídia contínua, especialmente vídeo, os sistemas deveriam suportar altas taxas de transmissão para executar aplicações multimídia. Para tentar amenizar o problema, as taxas de dados são reduzidas utilizando técnicas de compressão, mas isso introduz latência adicional e diminui a flexibilidade.

Além da alta largura de banda, outros requisitos são necessários, tais como redes de alta velocidade, poderosas estações de trabalho e interfaces adequadas de acesso aos computadores da rede para o processamento das informações [Canan e Raabe 2004].

A Tabela 2-1 [Cotroneo et al. 2003] mostra as médias das taxas de dados de alguns serviços, considerando que os valores são dependentes da tecnologia e sua evolução.

Tabela 2-1.: Média das taxas de dados de alguns serviços

Dado Multimídia	Serviço	Taxa de Dados
Áudio	Voz	4-64 Kb/s
Áudio	Mensagem de Voz	4-32 Kb/s
Áudio	Streaming de Áudio	16-128 Kb/s
Vídeo	Videoconferência	4-384 Kb/s
Vídeo	Video sob Demanda	4-384 Kb/s

2.2.2 Restrições Temporais

Mídias contínuas tem uma inerente dimensão temporal, resultando em um conjunto de requisitos de sincronização. Estes requisitos se devem a três categorias: Sincronização de Streams (fluxos), Sincronização de Eventos e Sincronização de Grupos.

Sincronização de Streams

Trata da manutenção das relações temporais existentes entre duas ou mais streams, como por exemplo, uma apresentação que envolve áudio e vídeo. Existem dois tipos de sincronização de streams: Sincronização Intra-Stream e Sincronização Inter-Stream. A Sincronização Intra-Stream envolve o atraso máximo e a máxima variação do atraso aceitáveis. Qualquer dado que exceda este atraso é descartado. O atraso inserido é dependente do tipo de mídia/stream e do estado dos componentes usados na transmissão. A Sincronização Inter-Stream trata das relações temporais que podem existir entre diferentes streams.

Sincronização de Eventos

Objetiva habilitar aplicações multimídia para reagir de maneira apropriada em resposta a um evento de notificação em uma progressiva apresentação de mídia. Isso introduz severas restrições de tempo para a execução de uma ação.

Sincronização de Grupos

Preocupa-se com o seguinte princípio: “O que você vê é o que eu vejo”, isto é, todos os participantes de uma sessão devem ter a mesma visão de janelas compartilhadas em um mesmo momento. Isto se torna muito importante para aplicações multimídia que se preocupam com interações entre um número de usuários, os quais podem estar recebendo a mesma informação ao mesmo tempo.

2.2.3 Garantia nos Serviços Oferecidos

A garantia nos serviços oferecidos deve existir para dar suporte às aplicações multimídia, ou seja, redes e sistemas fim devem ser capazes de oferecer algum tipo de reserva de recursos para suportar as taxas de dados e restrições temporais [Silveira 2000]. Por outro lado, essas mesmas redes e sistemas fim podem não oferecer uma qualidade aceitável para os serviços solicitados, podendo ser o dado perdido ou estar atrasado, devido à carência de recursos.

Apesar desta possibilidade, aplicações multimídia devem receber garantia nos serviços oferecidos, e isto com especificação do grau de garantia para as diferentes classes de qualidade, as quais são associadas com os tipos de serviços disponíveis.

2.2.4 Comunicação de Grupo

Muitas aplicações envolvem vários usuários compartilhando a mesma informação. Tais aplicações possuem restrições quanto ao tempo no qual a informação deve ser recebida por todos os membros do grupo. Para suportar estes requisitos, são necessários serviços de comunicação multiponto de alto desempenho. Assim, a provisão dos serviços de multicast e broadcast com específicas taxas de dados e latência é de grande importância. Serviços multicast e broadcast reduzem o tempo e os recursos requeridos na entrega do mesmo dado para vários receptores [Almeida et al. 2004].

2.3 Considerações Finais

Este capítulo descreveu uma visão geral sobre conceitos que formam a base dos Sistemas Multimídia, os componentes relativos a cada tipo de aplicação e os requisitos necessários para a execução destas aplicações foram apresentados. Depois de compreendido o funcionamento e as necessidades deste tipo de sistema, torna-se necessário tomar conhecimento da qualidade de serviço que pode ser oferecida pela rede para o mesmo, sendo possível escolher um modelo adequado que garanta seus requisitos de QoS.

No próximo capítulo serão descritos os conceitos fundamentais de Qualidade de Serviço, bem como seu importante papel em sistemas que envolvem o processamento destes dados multimídia.

Capítulo 3

Qualidade de Serviço em Sistemas Multimídia

O termo Qualidade de Serviço, relacionado a redes de computadores, tem sido usado na literatura com diferentes definições.

Em [Vogel et al. 1995], QoS é definida como um conjunto de características quantitativas e qualitativas de um sistema multimídia distribuído, necessárias para satisfazer os requisitos de funcionalidade de uma aplicação.

Na visão da ISO, segundo [Kamienski e Sadok 2000], QoS é definida como o efeito coletivo do desempenho de um serviço, o qual determina o grau de satisfação de um usuário do serviço.

Segundo [Teixeira e Barbar 2003], QoS pode ser definida como um requisito das aplicações multimídia, onde se exige que determinados parâmetros relativos a estas aplicações estejam dentro de limites bem definidos (valor máximo, valor mínimo). Estes parâmetros especificam a quantidade de recursos a serem alocados para uma determinada aplicação, podendo assim garantir qualidade para os usuários.

Neste trabalho, o termo Qualidade de Serviço é considerado como um conjunto de técnicas e parâmetros de um sistema multimídia distribuído, os quais influenciam a apresentação dos dados multimídia ao usuário e a satisfação do mesmo com a aplicação. Este conjunto de técnicas e parâmetros inclui atributos orientados ao desempenho, tais como vazão, atraso na transmissão de pacotes, variação do atraso e perda de pacotes.

As seções seguintes tratam dos seguintes aspectos: a seção 3.1 descreve o Modelo Conceitual de Qualidade de serviço e suas camadas; Na seção 3.2 são retratados os Elementos Necessários para se Garantir Qualidade de Serviço; A seção 3.3 descreve os Parâmetros de Qualidade de Serviço, tanto no nível do usuário quanto no nível mais interno; e por fim, a

seção 3.4 traz as Considerações de Desempenho para Diferentes Tipos de Aplicações Multimídia.

3.1 Modelo Conceitual de Qualidade de Serviço

O modelo conceitual de QoS é dividido em diversas camadas, que podem ser observadas na Figura 3-1, definida em [Nahrstedt e Steinmetz 1995] e posteriormente reafirmada em [Teixeira e Barbar 2003]:

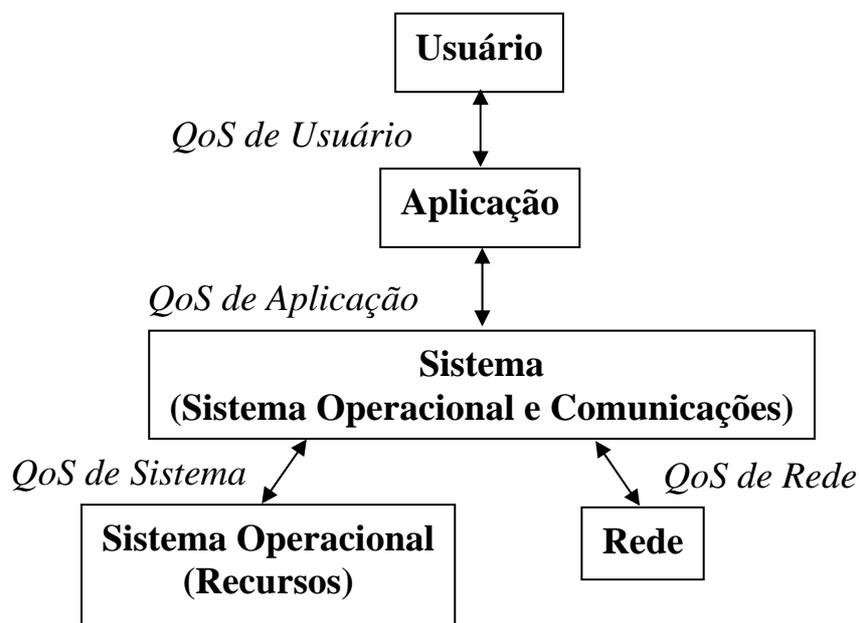


Figura 3-1: Um modelo conceitual de QoS

Este modelo de QoS foi escolhido por compreender as visões de qualidade de serviço de usuário, de aplicação, de sistema e de rede, sendo que cada um aborda aspectos de qualidade em seu escopo, executando a função de tradução entre as representações da QoS destes diferentes níveis. Deste modo, parâmetros de QoS especificados por um usuário (nível do usuário) para um uma aplicação de vídeo, por exemplo, incluem resolução de frame (altura e largura do frame), números de bits/pixel e frames por segundo. Para o sistema operacional, estes parâmetros são traduzidos em recursos a serem alocados, como buffer de armazenamento e carga de processamento. Já para a camada de rede, estes parâmetros são traduzidos como porcentagem de largura de banda a ser reservada e taxa de transmissão.

Esses mecanismos são adaptados no sistema final, para obter QoS para as aplicações multimídia, visando sempre a satisfação do usuário.

3.1.1 Camada do Usuário

Os parâmetros de qualidade relacionados diretamente aos usuários devem ser necessariamente menos técnicos e mais subjetivos, normalmente mensurados de forma qualitativa. Assim, a essência dessa camada é ocultar, tanto quanto possível, os parâmetros de QoS internos do sistema, freqüentemente sem significado para o usuário. Para isso, deve-se apresentar exemplos que permitam realizar escolhas como imagens de tamanho e resolução diferentes, ou áudio com qualidade de CD ou telefone. Em um sistema, cada escolha impacta nas exigências de QoS da rede. Exemplo disso é a seleção de uma imagem de alta resolução que pode resultar em um aumento de banda necessária e atraso de transferência.

3.1.2 Camada de Aplicação

Nessa camada são descritos os requisitos para a aplicação em termos de qualidade da mídia e relacionamento entre mídias. Escolhas realizadas pelo usuário final, na camada de QoS do usuário, são mapeadas em um conjunto de parâmetros de aplicação, os quais devem atender aos seus anseios. Os parâmetros, nesta camada, são associados a quadros de vídeo (tamanho do quadro, taxa de quadros por segundo) ou amostras (taxa de amostragem, bits por amostragem). Relacionamentos entre as mídias também podem ser especificados.

Uma especificação de QoS pode ter como ponto de partida escolhas feitas pelo usuário. Entretanto, ele nem sempre pode escolher valores arbitrários. Por exemplo, o número de quadros por segundo de um vídeo e o tamanho do quadro podem ser inversamente proporcionais, limitando assim o espaço utilizado. Alguns requisitos não podem ser diretamente mensuráveis pelo usuário, embora ele possa perceber a diferença existente no resultado final. Nesse caso, ficaria a cargo da aplicação, e não do usuário, estabelecer limites baseando-se em perfis previamente estabelecidos.

3.1.3 Camada de Sistema

Parâmetros desta camada representam características de cada nó visitado, as quais podem ser especificadas em termos de critérios quantitativos e qualitativos. Critérios quantitativos são aqueles que podem ser avaliados em termos de medidas concretas, como tempo de processamento de uma tarefa e tamanho da unidade de dados. Critérios qualitativos especificam os serviços esperados, como sincronização, mecanismos de escalonamento e recuperação de erros.

Os parâmetros nesse nível podem ter um grande impacto na qualidade percebida pelo usuário. Por exemplo, a velocidade do barramento, memória disponível e tempo de CPU (*Central Process Unit*) podem limitar a taxa de quadros da apresentação de um vídeo, tornando sua exibição mais lenta.

3.1.4 Camada de Rede

Abordam-se os requisitos dos serviços de rede, associados a propriedades de pacotes ou bits, tais como atraso de pacotes e taxa de bits. Neste nível, parâmetros fim-a-fim são especificados.

Para aplicações multimídia, a rede deve satisfazer os requisitos de transmissão de dados multimídia. O tráfego de mídia discreta (transferência de arquivos ou recuperação de imagens) requer serviços livres de erros, mas é tolerante a atrasos. O tráfego de mídia contínua requer tempo real e transmissão em alta velocidade. Essa mídia é sensível a atrasos e suas variações, mas é tolerante a perdas ocasionais de pacotes.

3.2 Elementos Necessários para prover Garantia de Qualidade de Serviço

Um modelo simplificado de QoS para um sistema de comunicação multimídia é aquele em que a aplicação determina a QoS que precisa, solicitando-a ao sistema. Se o sistema possui recursos suficientes para atender a solicitação, aceita a aplicação e aloca os recursos necessários para servi-la, satisfazendo seus requisitos. Caso contrário, ele pode rejeitar a aplicação ou oferecê-la a executar em um nível inferior de QoS, que pode ser satisfeita. Se a aplicação aceitar o conjunto de parâmetros oferecidos pelo sistema, a comunicação é iniciada. Se não aceitar, ela é rejeitada e pode voltar a tentar, na espera de que alguns recursos possam ter sido liberados por outra aplicação.

Garantir QoS requer a cooperação de todos os níveis da rede, ou seja, de todos os elementos que a compõem fim-a-fim. Qualquer garantia de qualidade de serviço será no máximo tão boa quanto o elo mais fraco na “cadeia” entre origem e destino [Melo 2001], e para dar suporte a essa garantia, são necessários os seguintes elementos:

- Mecanismo de especificação de QoS de uma aplicação;
- Gerenciamento de QoS;
- Políticas de negociação, para maximizar o número de aplicações a serem servidas;
- Alocação e escalonamento de recursos, para garantir os requisitos de QoS das aplicações aceitas;
- Política de controle de tráfego, para garantir a quantidade de dados gerados por aplicação, de acordo com o negociado na especificação.

Além destes elementos, é necessário oferecer outros serviços complementares:

- Mecanismo de Renegociação, pois uma aplicação pode querer mudar sua especificação inicial;
- Monitoramento da sessão de comunicação, para garantir a qualidade de serviço caso haja algum problema, estabelecendo as medidas a serem tomadas;
- Uso de técnicas como a escalabilidade de mídia e a degradação de qualidade, para oferecimento e garantia de serviços satisfatórios;

- Cada subsistema deve oferecer seus próprios mecanismos para prover QoS em todos os níveis.

Os itens de maior relevância citados acima serão melhor detalhados a seguir.

3.2.1 Especificação de QoS

Para especificar e garantir QoS, deve ser utilizada uma sessão orientada à conexão, onde parâmetros de QoS seriam negociados entre todos os subsistemas envolvidos.

Outro ponto importante é que o resultado da QoS é a qualidade percebida pelo usuário. Então, ele é que deve especificar o nível de qualidade desejado e o que é aceitável ou não.

3.2.2 Gerenciamento de QoS

Em sistemas multimídia, QoS define características que influenciam a qualidade perceptível de uma aplicação. É uma especificação qualitativa e quantitativa de um requisito da aplicação, a qual um sistema multimídia deveria satisfazer, com o objetivo de conseguir a qualidade desejada para a aplicação. Baseado nesta definição, há dois aspectos de QoS: as aplicações especificam requisitos de qualidade de serviço e os sistemas provêm garantia de qualidade de serviço.

Para prover a garantia de qualidade especificada por uma aplicação, o sistema deve possuir um Controle de Admissão e prover Gerenciamento de Falhas.

O Controle de Admissão serve para garantir que se uma nova aplicação for admitida, isso não afetará a qualidade de serviço das outras que estão em execução. Para isso, é necessário definir políticas de gerenciamento de QoS. A definição de políticas de gerenciamento inclui: identificação dos perfis ou classes de fluxo de tráfego existentes e sua forma de agregação, associação dos tipos de serviços aos perfis e alocação da largura de banda para cada perfil [Melo 2001]. Considerando uma única classe de usuários, uma política geral de gerenciamento de QoS pode ser colocada em prática [Bochmann et al. 2001]. Todos

os usuários ativos deveriam desfrutar do mesmo tempo de resposta. Se a média do tempo de resposta alcança um valor R_{Max} , o sistema não deveria aceitar qualquer nova sessão de usuário.

No caso de haver duas classes de usuários, A e B, onde A pertence a uma classe de nível mais alto que B, poderia haver duas maneiras possíveis de priorizar A, como por exemplo:

Política 1:

- a. Se o tempo de resposta da classe A alcança R_{MaxA} , então não serão aceitas novas sessões de usuários da classe A;
- b. Se o tempo de resposta da classe B alcança R_{MaxB} , então não serão aceitas novas sessões de usuários da classe B;
- c. $R_{MaxA} < R_{MaxB}$.

Política 2:

- a. Se o tempo de resposta da classe A alcança R_{MaxA} , então não serão aceitas novas sessões de usuários da classe A nem de usuários da classe B;
- b. Se o tempo de resposta da classe B alcança R_{MaxB} , então não serão aceitas novas sessões de usuários da classe B;
- c. $R_{MaxA} < R_{MaxB}$.

O Gerenciamento de Falhas se preocupa em detectar porque uma requisição de QoS não está sendo satisfeita. Técnicas para diagnosticar falhas definem hipóteses das prováveis causas de degradação, e posteriormente são realizados testes para validar essas hipóteses.

Para que a arquitetura em questão ofereça suporte a detecção e diagnóstico de falhas, é necessário que possua monitores de recursos para a rede e para os nós que executam os componentes da aplicação. Seriam necessários também monitores que medem o comportamento da aplicação. Além destes monitores, a arquitetura deve possuir aplicações de gerenciamento que podem analisar a informação monitorada e determinar se a requisição de QoS foi violada, e caso seja, diagnosticar a causa [Bochmann et al. 2001].

3.2.3 Negociação e Renegociação de QoS

A negociação é feita por cada subsistema, um a um, até todos garantirem alguma qualidade de serviço. Parâmetros de QoS são transmitidos de uma camada para outra e cada subsistema indica se suporta ou não cada serviço. Se positivo, é reservado o recurso àquela sessão. Somente quando todos os subsistemas concordarem com os parâmetros de qualidade, é que a sessão é realmente estabelecida.

Uma renegociação pode ser necessária, caso haja alterações durante a transmissão, feitas pelo usuário. Estas alterações incluem: mudança de qualidade por parte do usuário de um determinado canal ou de toda uma transmissão, fechamento de um canal de áudio ou vídeo e abertura de novo canal para determinada mídia.

Um exemplo destes tipos de mudança ocorre quando o usuário define uma sessão de um canal de vídeo com baixa qualidade, pressupondo que ela é suficiente para sua aplicação, mas ao executar, percebe que a qualidade estava muito baixa e decide aumentar.

3.2.4 Sistema de Reserva de Recursos

Uma reserva de recursos começa com a solicitação do mesmo e termina após um intervalo de tempo. Para resolver os problemas quanto à reserva, tal como rejeição de conexão em função de recurso não disponível ou escalonamento na utilização, dando prioridade a aplicações mais importantes, este tipo de serviço foi separado em duas sessões, preocupadas em como dividir os recursos entre todas as classes: Sessões de Reserva Avançada e Sessões de Reserva Convencional. Esta divisão define formas de como especificar mais precisamente o tempo de duração das reservas e de que maneira o usuário será informado sobre o sucesso de seus pedidos de reserva, caso haja alguma falha no sistema.

Sobre a duração das reservas, existem duas soluções. Na primeira delas, o total de recursos é dividido em duas partes, sendo que uma é alocada a sessões convencionais e a outra a sessões mais avançadas. Na segunda, todas as sessões compartilham o total dos recursos.

Como o usuário será informado sobre o sucesso de seus pedidos de reserva é uma questão um pouco mais difícil de resolver. Seria necessário comunicar ao usuário um problema ocorrido para que ele possa tomar as medidas apropriadas, como por exemplo, re-escalonar suas sessões. O grande problema é que o sistema para este fim pode estar inativo.

3.2.5 Controle de Tráfego

Os protocolos de comunicação utilizados por um sistema multimídia devem incluir mecanismos de controle de tráfego, baseados nas taxas estabelecidas para QoS. Esse controle serve tanto para que os pacotes sejam enviados em um determinado padrão, como para verificar se os pacotes que estão sendo recebidos seguem esse padrão, ou seja, para garantir que a quantidade de dados gerados por aplicação esteja de acordo com o negociado na especificação.

Tais mecanismos podem adotar uma política de conservação de trabalho ou de não-conservação de trabalho [Nahrstedt e Steinmetz 1995]. No primeiro caso, os pacotes são enviados na taxa mais alta possível sem prejudicar as demais conexões, de modo que o nó mantenha-se sempre ocupado. Na outra alternativa, o tempo em que cada pacote deve ser enviado é determinado. Se em um dado momento, nenhum pacote for elegível para transmissão, nada será transmitido, mesmo que a CPU esteja inativa.

3.3 Parâmetros de QoS

Parâmetros de qualidade de serviço são usados para caracterizar o ambiente de gerenciamento de QoS ao longo do caminho de um fluxo de pacotes que requer controle de qualidade fim-a-fim [Shenker e Wroclawski 1997]. A definição de cada parâmetro usado para caracterizar um caminho através da rede possui dois tipos de semântica: local e composta. O valor local caracteriza um único elemento da rede. Valores compostos refletem a composição corrente de valores locais ao longo de um caminho, segundo uma regra de composição.

Cada definição de parâmetro especifica sua regra de composição. Essa regra diz como combinar um valor composto (da porção já percorrida do caminho) com o valor local. A

partir de então, gera-se um novo valor a ser passado ao próximo elemento da rede no caminho [Shenker e Wroclawski 1997]. Uma regra de composição pode ser especificada tanto para *downstream* (em direção ao receptor) como para *upstream* (em direção ao emissor). O valor do parâmetro relacionado ao elemento local da rede deve possuir uma única definição, porém mais de uma regra pode ser definida para a composição de um parâmetro ao longo do caminho, dependendo da necessidade e da natureza do parâmetro.

Os parâmetros de caracterização possuem essencialmente o sentido de *per-next-hop*, ou seja, de buscar sempre o próximo nó do caminho. Dessa forma, elementos de rede com várias possibilidades de encaminhamento, roteadores de borda de grandes sub-redes, por exemplo, podem assumir diferentes valores de composição de um parâmetro, dependendo do endereço do próximo elemento no caminho.

Com o objetivo de especificar o gerenciamento de Qualidade de Serviço, devido à existência de diferentes tipos de parâmetros, é necessário distinguir entre os parâmetros de alto nível de qualidade de serviço, relevantes no nível de usuário e os parâmetros internos de qualidade, pertencentes aos diferentes recursos do sistema, os quais estão sendo usados por uma determinada aplicação. Além de uma descrição dos parâmetros de QoS, é também apresentada uma proposta de quantização destes parâmetros em um intervalo de valores homogêneos.

3.3.1 Parâmetros de QoS no Nível de Usuário

Os parâmetros de qualidade importantes da perspectiva do usuário são [Bochmann et al. 2001]: Tempo de Resposta, Disponibilidade, Capacidade de Atendimento, Atendimento. O Tempo de Resposta é o tempo entre o momento em que um pedido é enviado até o tempo em que a resposta chega ao usuário; a Disponibilidade trata da medida do tempo livre efetivo do sistema; a Capacidade de Atendimento define a porcentagem de tempo que o servidor está disponível e pode aceitar a requisição do usuário; e o Atendimento é a medida do número de pedidos do usuário que são atendidos pelo sistema.

3.3.2 Parâmetros Internos de QoS

Os parâmetros internos de qualidade estão relacionados a diferentes componentes [Mbise e Woods 2003], que podem ser classificados como Parâmetros de Sistema, Parâmetros de Rede e Parâmetros Relacionados ao Dispositivo. Os Parâmetros de Sistema dependem apenas das condições locais da máquina, não dependendo da rede. Os parâmetros deste tipo são: Ociosidade de memória, Ociosidade de processamento, e Disponibilidade de disco. Os valores dos Parâmetros de Rede, residentes em cada nó ativo, são obtidos através da análise das características e da condição da conexão de rede entre o nó em questão e seus vizinhos imediatos. Dentre estes parâmetros estão: Alcançabilidade e Confiabilidade, Vazão e Largura de Banda, Atraso ou Delay, Variação do Atraso ou Jitter e Perda de Pacotes. Os Parâmetros Relacionados ao Dispositivo tratam das limitações dos dispositivos utilizados, tais como Carga de Bateria, Capacidades de Processamento e Memória.

De acordo com [Dressler 2003c];[Dobson 2004], os parâmetros com maior relevância e que atuam de forma decisiva em aplicações multimídia são os Parâmetros de Rede. Tais parâmetros são descritos no item a seguir.

Parâmetros de QoS de Rede

Alcançabilidade e Confiabilidade

Alcançabilidade ou Conectividade entre dois sistemas significa que é possível transmitir dados entre os dois. As medidas típicas de confiabilidade são baseadas nesta métrica simples.

Confiabilidade significa alcançabilidade em um certo período de tempo.

Vazão e Largura de Banda

Largura de banda é uma medida de capacidade de transmissão de dados, normalmente expressa em kilobits por segundo (Kbps) ou megabits por segundo (Mbps). A largura de banda indica a capacidade máxima de transmissão de uma conexão. Entretanto, na medida em que a taxa de transmissão utilizada se aproxima da largura de banda máxima, fatores negativos, como atraso na transmissão das informações podem causar deterioração na

qualidade [Melo 2001]. A largura de banda de uma rede pode ser vista como um tubo que transfere dados. Quanto maior o diâmetro do tubo, mais dados podem ser enviados através dele simultaneamente.

A vazão é o montante de tráfego de dados movidos de um nó da rede para outro em um determinado período de tempo [Melo 2001]. A vazão também é expressa em Kbps ou Mbps.

Estes dois termos estão diretamente relacionados, pois quanto maior for a Largura de Banda, maior poderá ser a Vazão de dados através do canal de comunicação.

Atraso ou Delay

Descreve o intervalo de tempo entre a transmissão de um pacote e sua recepção com sucesso no destino. Um dos mais importantes valores em uma comunicação multimídia de tempo real é o atraso da mensagem para percorrer o caminho de um nó a outro (one-way delay - OWD); isso por diversas razões, a saber [Monteiro et al. 2002]: todas as transmissões de áudio e vídeo possuem um sinal unidirecional de um nó para outro; o caminho de ida geralmente é diferente do caminho de volta; a largura de banda pode ser assimétrica; ao medir um sentido por vez, pode-se saber a diferença no desempenho de ambos; ainda que o caminho de ida seja igual ao caminho de volta, o atraso pode ser diferente em função do estado das filas dos roteadores; em algumas aplicações, o desempenho depende apenas de um único sentido.

Considere o envio de um pacote, como ilustrado na Figura 3-2, descrita por [Dressler 2003c]. Um pequeno atraso é introduzido pela aplicação, pelo sistema operacional e pelo hardware de rede, fazendo com que a solicitação de envio do nó A que ocorreu em t_{A0} , só seja efetivamente enviada em t_{A1} .

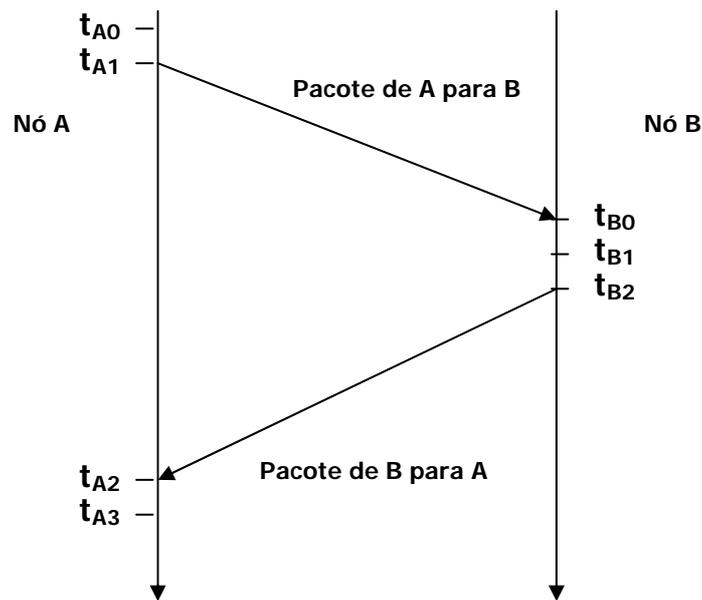


Figura 3-2: Atrasos introduzidos

O nó B recebe o pacote em t_{B0} e o timestamp, “etiqueta” digital que permite detectar a época da criação de determinados dados, é processado pela aplicação somente em t_{B1} , sendo t_{B1} igual ao tempo no qual o pacote foi recebido. Usando o timestamp t_{A0} , é possível calcular o OWD, atraso da mensagem para percorrer o caminho de um nó a outro, da seguinte forma:

$$\Delta t_{\text{OWD}} = t_{B1} - t_{A0}$$

Em algumas aplicações que exigem um tempo de resposta mais rápido, tais como videoconferência, o RTT (Round Trip Time), tempo requerido para um pacote ir da fonte ao destino e retornar à fonte, é um importante parâmetro a ser calculado.

Neste caso, uma solicitação de envio feita pelo nó B no tempo t_{B1} será efetivada em t_{B2} . Em t_{A2} , o pacote é recebido no nó A e o último timestamp é processado em t_{A3} .

Usando todos estes timestamps, é possível calcular o RTT:

$$\Delta t_{\text{RTT}} = \Delta t_{\text{OWD}} + \Delta t_{\text{OWD}} = [(t_{A3} - t_{B1}) + (t_{B1} - t_{A0})] = t_{A3} - t_{A0}$$

Como a medida do RTT depende somente de t_{A3} e t_{A0} , apenas o clock de um dos nós está envolvido no processo. Assim, a sincronização entre ambos os clocks não é necessária.

Varição do Atraso ou Jitter

A variação do intervalo de tempo entre chegadas de pacotes no receptor é conhecida como variação do atraso ou jitter e pode ser medida de duas maneiras [Dressler 2003b]. Na primeira delas, os atrasos são medidos em um determinado período de tempo e posteriormente calcula-se a média entre eles. Na segunda, analisa-se a variação do tempo entre chegadas dos pacotes. Usando uma constante de fluxo de pacotes com uma distância bem definida entre eles, o receptor pode medir essa distância facilmente. O jitter pode, então, ser calculado como a máxima variação do tempo entre chegadas em um período de tempo.

Perda de Pacotes

A taxa de perda de pacotes é calculada no lado do receptor como a razão entre a quantidade de pacotes perdidos e a quantidade de pacotes transmitidos, em cada intervalo de tempo considerado [Monteiro et al. 2002]. A perda de pacotes se dá em função do descarte que estes sofrem nas filas dos roteadores quando estas se tornam cheias e não permitem mais a admissão e o armazenamento de pacotes para posterior envio. Protocolos como o UDP (User Datagram Protocol), normalmente utilizado para transmissão de áudio e vídeo, não retransmite pacotes e, portanto estas perdas não são recuperadas.

Para medir a taxa de perda de pacotes, é necessário que cada pacote possua um número de seqüência.

3.4 Considerações de Desempenho para Diferentes Aplicações de Áudio e Vídeo

Diferentes tipos de aplicações multimídia de áudio e vídeo possuem requisitos distintos de desempenho para que executem de forma satisfatória. A fim de se implementar um serviço de rede adequado para estas aplicações foi necessário estudá-las e classificá-las de acordo com suas características principais, em Interativas e Não Interativas. Em [Monteiro et al. 2002], dentre outras classificações, encontra-se também uma classificação de mais alto nível destas aplicações, onde elas são agrupadas de acordo com o serviço que irão prestar:

Aplicações de Áudio

Os fatores que influenciam na qualidade de aplicações de áudio interativas são a fidelidade do sinal de voz, o eco, o atraso, o jitter fim-a-fim e a baixa perda. Uma aplicação de Voz é influenciada pelo OWD, que pode resultar em eco e impactar a dinâmica da conversação. Além disso, é intolerante ao jitter, e qualquer valor acima do tolerado causa danos bastante sensíveis ao ouvido humano. Tolerada determinada quantidade de perda de informação. No caso de Mensagens de Voz existem exigências para a perda de informação similares a uma aplicação de Voz, mas com uma maior tolerância ao atraso, desde que não haja conversação direta envolvida.

Aplicações não interativas ou com pouca interatividade são influenciadas por atraso, jitter e perda, mas o grau de influência desses fatores depende da implementação das aplicações. Um exemplo deste tipo de aplicação seria uma Streaming de Áudio de qualidade profissional. Uma Streaming de Áudio fornece melhor qualidade do que o telefone convencional, sendo maiores as exigências relacionadas à perda da informação em termos de perda de pacotes e possui poucas exigências com relação ao atraso.

Aplicações de Vídeo

Aplicações de vídeo interativas são influenciadas pelo atraso, jitter fim-a-fim e perda de pacotes, a depender da tecnologia de vídeo utilizada. Exemplos de aplicações com esta característica são a Videoconferência e o Videofone. Em se tratando do Videofone, implica em um sistema full-duplex, com vídeo e áudio sincronizados em um ambiente de conversação. Detém as mesmas exigências de atraso das aplicações de Voz, isto é, nenhum eco e efeito mínimo na dinâmica de conversação. É tolerante a uma determinada quantidade de perda da informação, a qual não é perceptível ao olho humano.

Nas aplicações não interativas, os requisitos variam a depender da aplicação, não existindo nenhum elemento de conversação envolvido. Um Streaming de Vídeo, por exemplo, de qualidade modesta, irá requerer banda em torno de 500 Kbps, poderá tolerar perdas e jitter, a depender da implementação. No caso de uma aplicação de Vídeo sob Demanda, há poucas exigências com relação ao atraso.

A Tabela 3-1 definida em [ITU 2001], mostra os valores objetivados para cada um dos parâmetros de algumas das aplicações de Áudio e Vídeo citadas acima, com maiores detalhes. Para cada dado multimídia, são apresentados a classe de aplicação, a simetria do fluxo da informação e os parâmetros de vazão, atraso, jitter e perda de pacotes.

Tabela 3-1: Objetivos de desempenho para aplicações de áudio e vídeo

Dado Multimídia	Aplicação	Grau de Simetria	Vazão	Delay	Jitter	Perda de Pacotes
Áudio	Voz	Duplo sentido	4-64 Kb/s	<150 mseg (preferido) <400 mseg (limite)	<1 mseg	< 3%
Áudio	Mensagem de Voz	Primariamente em um único sentido	4-32 Kb/s	<1 seg (para exibição) <2 seg (para gravação)	<1 mseg	< 3%
Áudio	Streaming de Áudio	Primariamente em um único sentido	16-128 Kb/s	<10 seg	<1 mseg	< 1%
Vídeo	Videofone	Duplo sentido	4-384 Kb/s	<150 mseg (preferido) <400 mseg (limite)	-	< 1%
Vídeo	Video sob Demanda	Único sentido	>64 Kb/s	<6 seg	-	< 1%

3.5 Considerações Finais

O objetivo deste capítulo foi proporcionar um entendimento básico sobre o que é Qualidade de Serviço e sua aplicação em Sistemas Multimídia.

Os dois capítulos seqüentes descrevem as metodologias de Inteligência Computacional utilizadas no desenvolvimento deste projeto de mestrado: Sistemas Nebulosos e Algoritmos Genéticos.

Capítulo 4

Lógica Nebulosa e Sistemas Nebulosos

Lógica Nebulosa pode ser definida como sendo uma metodologia capaz de capturar informações vagas, em geral descritas em uma linguagem natural, e convertê-las para um formato numérico. Quando se diz que o tempo está ensolarado, não quer dizer que não haja nuvens. Da mesma forma, em um tempo nublado, pode haver sol. Geralmente é aceito como ensolarado um céu com até 20% ou 30% de nuvens [Klir e Yuan 1995]. Assim, é preciso definir limites para estas proposições. Pode-se dizer que, quando o ser humano utiliza tais termos, ele infere um grau de verdade, ou seja, entre as certezas de ser ou não ser, existem infinitos graus de incerteza. Dentro de uma região tropical, por exemplo, pode-se aceitar nuvens em maior quantidade, visto que uma pouca quantidade de sol poderia ser suficiente para as atividades que dele dependem. Um meteorologista talvez estabelecesse limites mais rígidos e precisos. Os valores que definem o grau de verdade de uma proposição podem não ser conscientes, mas de alguma forma definidos pelo contexto social, lingüístico ou uma base de referência experimental. A extração (representação) destes termos e valores ocorre por meio do uso de Conjuntos Nebulosos e a definição e a modelagem dos mesmos constituem a base da teoria e do estudo dos Sistemas Nebulosos.

A teoria dos Conjuntos Nebulosos [Zadeh 1965] e os conceitos envolvendo Lógica Nebulosa [Zadeh 1973], podem ser usados para traduzir em termos matemáticos a informação imprecisa expressa por um conjunto de regras lingüísticas. Se um operador humano for capaz de articular sua estratégia de ação como um conjunto de regras da forma SE ENTÃO, um algoritmo passível de ser implementado em computador pode ser construído, conforme demonstrado nos anos 70 por [Mamdani 1974];[Mamdani e Assilian 1975];[Mamdani 1977].

O resultado é um sistema de inferência baseado em regras, no qual a teoria dos Conjuntos Nebulosos fornece o ferramental matemático para tratar tais regras lingüísticas.

4.1 Conjuntos Nebulosos e sua Representação

A partir da teoria apresentada por Zadeh [Zadeh 1965], um conjunto nebuloso tem sido definido como uma coleção de objetos com valores de pertinência entre 0 (exclusão completa) e 1 (inclusão completa). Os valores de pertinência expressam graus de compatibilidade dos objetos com as propriedades ou características distintas da coleção. Conjuntos nebulosos são formalmente caracterizados por uma função de pertinência que mapeia os elementos do domínio, espaço ou universo de discurso X no intervalo $[0,1]$. Isto é,

$$A: X \rightarrow [0,1]$$

Portanto, um conjunto nebuloso A em X pode ser representado como um conjunto de pares ordenados de um elemento genérico $x \in X$,

$$A = \{(x, A(x)) \mid x \in X\}$$

onde $A(x)$ representa o grau de x em A .

De acordo com um exemplo descrito em [Zimmermann 1990], considere um vendedor de casas que deseja classificar as casas oferecidas aos seus clientes. Um indicador de conforto é dado pelo número de quartos presentes em cada casa. Seja $X = \{1, 2, 3, 4, \dots, 10\}$ o conjunto disponível de tipos de casas descritas por $x =$ número de quartos. Então, o conjunto nebuloso “casa confortável para uma família com quatro pessoas” pode ser descrito como:

$$A = \{(1, 0.2), (2, 0.5), (3, 0.8), (4, 1), (5, 0.7), (6, 0.3)\}$$

O primeiro elemento denota o valor de x e o segundo elemento representa o grau de pertinência de x em A .

O conceito de função de pertinência desempenha um papel fundamental na teoria de representação dos conjuntos nebulosos. Os gráficos das funções podem ter diferentes formas e propriedades. Na prática é mais conveniente usar as funções de formato mais simples, tais como Função Triangular e Função Trapezoidal [Belchior 1995].

Função de Pertinência Triangular

A função Triangular é definida da seguinte forma:

$$A(x) = \begin{cases} 0, & \text{se } x \leq a \\ \frac{x-a}{m-a}, & \text{se } x \in [a, m] \\ \frac{b-x}{b-m}, & \text{se } x \in [m, b] \\ 0, & \text{se } x \geq b \end{cases}$$

onde m é um valor modal, e a e b denotam o limite inferior e superior, respectivamente, para valores não nulos de $A(x)$. Outra notação também pode ser usada: $A(x,a,m,b) = \max\{\min[(x-a)/(m-a), (b-x)/(b-m)], 0\}$. Este tipo de função serve para indicar apenas um único ponto de máximo no conjunto, definido no intervalo $[0,1]$, como descreve a Figura 4-1, onde $a = -2,5$, $b = 2,5$ e $m = 0$.

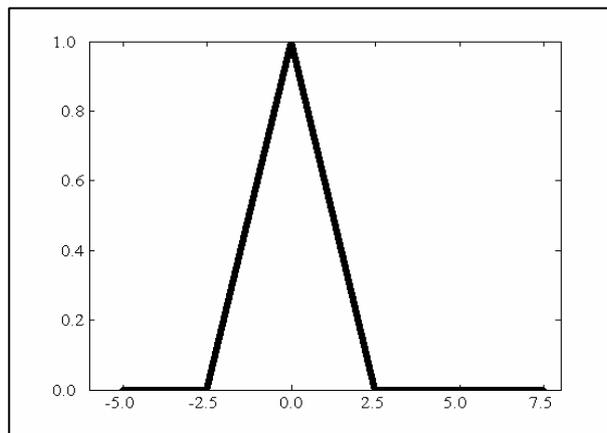


Figura 4-1: Representação Triangular

Função de Pertinência Trapezoidal

A função Trapezoidal é definida por:

$$A(x) = \begin{cases} 0, & \text{se } x < a \\ \frac{x-a}{m-a}, & \text{se } x \in [a,m] \\ 1, & \text{se } x \in [m,n] \\ \frac{b-x}{b-n}, & \text{se } x \in [n,b] \\ 0, & \text{se } x > b \end{cases}$$

Usando uma notação equivalente, obtemos: $A(x, a, m, b) = \max\{\min[(x-a)/(m-a), 1, (b-x)/(b-n)], 0\}$. Nesta função é possível representar todo um intervalo de pontos de máximo, também definido dentro do intervalo $[0,1]$, conforme mostrado na Figura 4-2, onde $a = -2,5$, $m = 0$, $n = 2,5$ e $b = 5,0$.

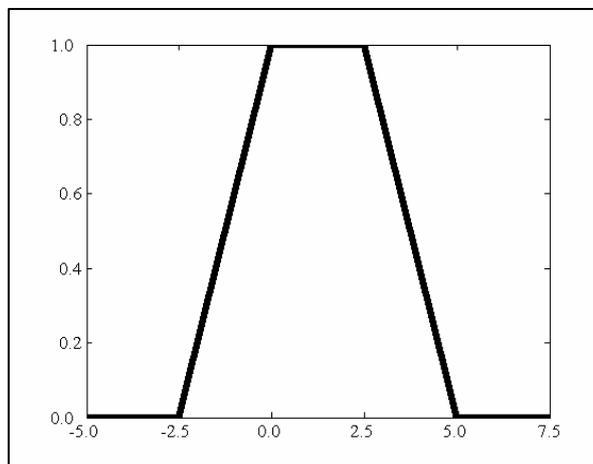


Figura 4-2: Representação Trapezoidal

4.1.1 Operações Básicas com Conjuntos Nebulosos

Com base na teoria dos conjuntos clássicos, Zadeh [Zadeh 1965] definiu as operações de união, interseção e complemento para conjuntos nebulosos, a partir da função de pertinência.

União

A função de pertinência $\mu_A \cup \mu_B$ é definida ponto a ponto para todo $u \in U$ por:

$$\mu_{A \cup B}(u) = \max \{ \mu_A(u), \mu_B(u) \} \text{ ou } \mu_{A \cup B}(u) = \mu_A(u) \vee \mu_B(u)$$

Intersecção

A função de pertinência $\mu_A \cap \mu_B$ é definida ponto a ponto para todo $u \in U$ por:

$$\mu_{A \cap B}(u) = \min \{ \mu_A(u), \mu_B(u) \} \text{ ou } \mu_{A \cap B}(u) = \mu_A(u) \wedge \mu_B(u)$$

Complemento

A função de pertinência $\mu_{\neg A}$ do complemento de um conjunto nebuloso A é definida ponto a ponto para todo $u \in U$ por:

$$\mu_{\neg A} = 1 - \mu_A(u)$$

4.1.2 Variáveis Lingüísticas

Uma variável lingüística pode ser definida, de uma maneira informal, como uma variável cujos valores são palavras ou sentenças, ao invés de números [Pedrycz e Gomide 1998]. Zadeh [Zadeh 1975], definiu formalmente uma variável lingüística como sendo uma quintupla denotada por:

$$\langle X, T(X), U, G, M \rangle$$

onde,

- X é o nome da variável, cuja variável base é x ;
- $T(X)$ é o conjunto de termos lingüísticos. Cada elemento de $T(X)$ representa um rótulo L dos termos que a variável X pode assumir;
- U é o universo de discurso da variável lingüística X ;
- G é a gramática para a geração dos termos ou rótulos;

- M é a regra semântica que associa a cada rótulo L, um conjunto nebuloso no universo U, representando o seu significado M(L).

Por exemplo, considere a variável lingüística velocidade ($X = \text{velocidade}$) com universo $U = [0,150]$ e variável base $x \in X$. Um possível conjunto de termos associados à variável velocidade poderia ser:

$$T(\text{velocidade}) = \{\text{muito baixa, baixa, média, alta, muito alta}\}.$$

Assim, o valor de uma variável lingüística é geralmente um termo composto, resultado da concatenação de outros termos, os quais podem ser divididos em 4 categorias:

- termos primários: rótulos de conjuntos nebulosos especificados em um determinado universo, como por exemplo: alto, baixo, pequeno, médio, grande;
- a negação ‘não’ e os conectivos ‘e’, ‘ou’ e ‘mas’;
- modificadores: muito, pouco, levemente, positivo, negativo;
- e marcadores, tais como parênteses.

Os conectivos ‘e’ e ‘ou’ são equivalentes a operações de união e intersecção de conjuntos, respectivamente, podendo dar origem a conjuntos complexamente definidos, porém representados lingüisticamente de maneira simples.

A Figura 4-3, retrata o termo Variável Lingüística e o que seriam seus valores quando associadas à funções de pertinência, utilizando a função de representação trapezoidal.

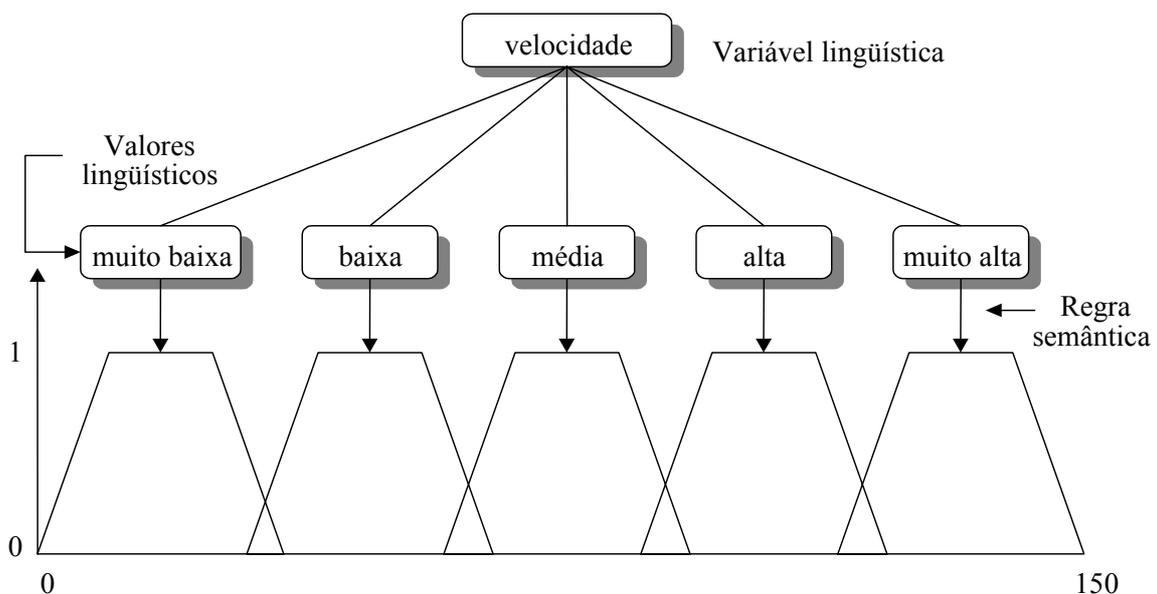


Figura 4-3: Representação de Variável Lingüística

4.2 Sistema Nebuloso

Com base nas informações apresentadas, é possível construir um Sistema Nebuloso. O Sistema Nebuloso tem como objetivo modelar o modo aproximado de raciocínio, permitindo o desenvolvimento de sistemas que imitem a habilidade humana de tomar decisões racionais em um ambiente de incerteza e imprecisão. Deste modo, a lógica nebulosa oferece meios para capturar informações imprecisas, em linguagem natural, e convertê-las em uma forma numérica [Mendel 1995]. Este tipo de sistema baseia-se na teoria de Conjuntos Nebulosos e Regras Nebulosas, utilizando o conceito de pertinência para modelar as regras imprecisas e seu processamento. Na Figura 4-4 [Tanscheit 2002], é mostrado um diagrama de um Sistema Nebuloso.

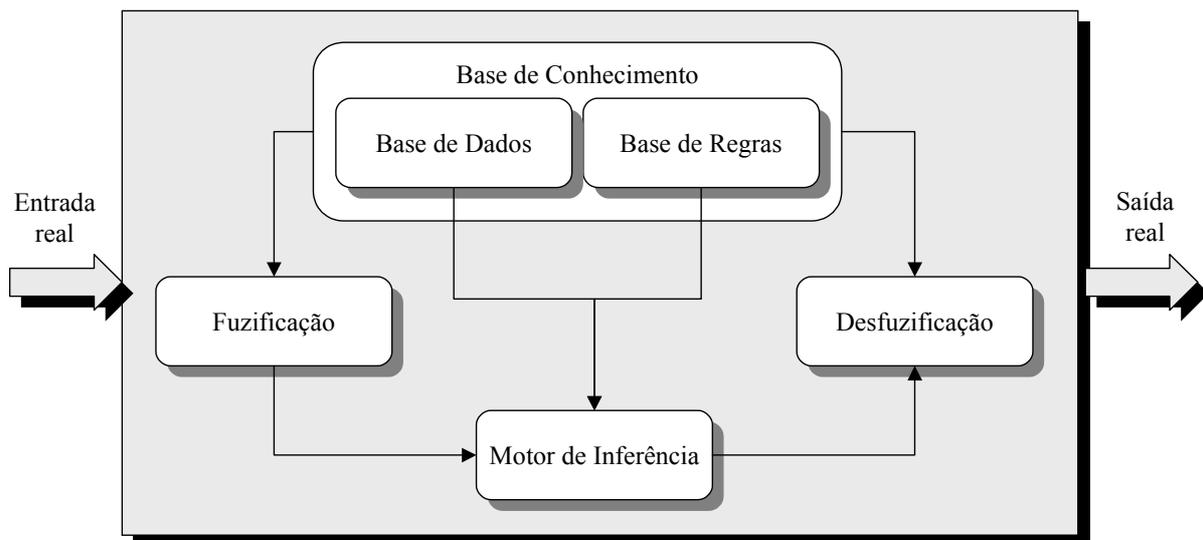


Figura 4-4: Sistema Nebuloso

O Sistema Nebuloso trata, portanto, das formas de imprecisão, que segundo Zadeh [Zadeh 1996] “é uma necessidade quando a informação disponível é muito imprecisa para justificar o uso de números, e quando há uma tolerância por imprecisão que pode ser explorada para alcançar tratabilidade, robustez, solução de baixo custo e concordância melhor com a realidade”.

O Sistema Nebuloso inclui módulos de interface que mapeiam valores de entrada em conjuntos nebulosos equivalentes e, posteriormente, os valores nebulosos gerados em valores

numéricos. As principais características desta técnica são a capacidade de modelar ambigüidades e o número menor de regras necessárias para definir um sistema, comparativamente às regras precisas, do tipo “sim” ou “não”. Também como uma vantagem desta abordagem, pode-se citar o fato de que em um Sistema Nebuloso o conhecimento é armazenado na forma em que o especialista transmite [Cox 1998];[Jager 1995].

4.2.1 Fuzificação

O Módulo de Fuzificação tem como função converter os valores de entrada do sistema (números reais) para termos lingüísticos representados por determinados conjuntos nebulosos. Esta conversão depende principalmente da função de pertinência, pois é ela que designará a qual conjunto Nebuloso um determinado valor pertence e qual o grau de pertinência desse valor em relação a este conjunto [Klir e Yuan 1995].

4.2.2 Base de Conhecimento Nebuloso

A Base de Conhecimento é o local onde todo o conhecimento sobre o domínio do problema em questão é armazenado. A Base de Conhecimento é formada por uma Base de Regras, que contém o conjunto de regras nebulosas e por uma Base de Dados, que define as funções de pertinência usadas nas regras nebulosas.

A maneira mais comum de armazenar informações em uma Base de Conhecimento Nebuloso é a representação por meio de Regras de Produção Nebulosas [Rezende 2003]. Uma regra de produção normalmente é formada por duas partes principais:

SE <antecedente> **ENTÃO** <conseqüente>

O antecedente é composto por um conjunto de condições que, quando satisfeitas, mesmo que parcialmente, determinam o processamento do conseqüente da regra por um mecanismo de inferência nebulosa, descrevendo assim o disparo de uma regra. Por sua vez, o conseqüente é composto de um conjunto de ações ou diagnósticos que são gerados com o disparo da regra. Os conseqüentes das regras disparadas são processados em conjunto para gerar uma resposta determinística para cada variável de saída do sistema.

As regras formam uma das partes fundamentais da estrutura de conhecimento em um sistema nebuloso de inferência e podem ser definidas manualmente com a ajuda de um ou mais especialistas no assunto em estudo ou serem geradas automaticamente por um sistema, caso haja dados disponíveis.

Os sistemas nebulosos que utilizam regras nebulosas para representar o conhecimento do domínio do problema são denominados Sistemas Nebulosos Baseados em Regras (SNBR).

4.2.3 Motores de Inferência

Existem vários modelos de motores de inferência. Na maioria dos casos, o antecedente é formado por proposições lingüísticas e a distinção entre os modelos se dá no conseqüente das regras nebulosas. Entre os modelos mais conhecidos estão Mamdani, Takagi-Sugeno e Tsukamoto [Delgado 2002].

O Modelo de Mamdani [Mamdani e Assilian 1975], utiliza conjuntos nebulosos também nos conseqüentes das regras. A saída final é representada por um conjunto nebuloso resultante da agregação da saída inferida de cada regra. Para se obter uma saída final *não*-nebulosa adota-se algum método de transformação da saída nebulosa em *não*-nebulosa. No Modelo de Takagi-Sugeno [Takagi e Sugeno 1983], o conseqüente é representado por uma função das variáveis de entrada. A saída final é obtida pela média ponderada das saídas inferidas de cada regra. Os coeficientes da ponderação são dados pelos graus de ativação das respectivas regras. O Modelo de Tsukamoto [Tsukamoto 1979], utiliza funções de pertinência monotônicas no conseqüente. Assim como no modelo Takagi-Sugeno, é inferido um valor *não*-nebuloso induzido pelo nível de ativação da regra e a saída final é obtida por média ponderada das saídas inferidas de cada regra.

Neste trabalho será utilizado o Modelo de Mamdani [Mamdani e Assilian 1975], em função de permitir obter como resposta um valor escalar e um termo lingüístico. Os modelos de Takagi-Sugeno e Tsukamoto oferecem como resposta apenas um valor escalar, o que não diferenciaria o trabalho de outros que utilizam métricas numéricas convencionais.

Baseado no trabalho de [Rezende 2003], será apresentada a seguir a definição de todos os processos envolvidos no Modelo de Inferência de Mamdani.

O Módulo de Inferência recebe valores nebulosos provenientes do módulo de Fuzificação, processa as regras existentes na Base de Regras e gera um Conjunto Nebuloso de saída para o módulo de Defuzificação, a partir da composição de todas as regras disparadas. A obtenção da conclusão de cada regra depende da semântica escolhida e dos resultados da combinação e agregação dos antecedentes. Da agregação das conclusões inferidas de cada regra, resulta um conjunto nebuloso. Por uma regra disparada, entende-se uma regra cujo processamento do antecedente para as entradas atuais gerou graus de pertinência não-nulos, ou seja, a relação nebulosa entre as entradas e os termos primários do antecedente é maior que zero.

A característica básica do modelo de Mamdani é o fato de utilizar conjuntos nebulosos nos conseqüentes das regras. A inferência tradicionalmente utilizada pelo modelo de Mamdani é chamada de Inferência Max-Min. Ela utiliza as operações de união e intersecção entre conjuntos.

Considerando uma regra genérica mostrada abaixo, o método de inferência Máx-Min proposto por Mamdani será detalhado a seguir.

$$\text{Se } x_1=A_i \text{ E } x_2=A_j \text{ E...E } x_p=A_j \text{ Então } y_1=B_i \text{ E } y_2=B_m$$

Nesta regra, x_i são as entradas do sistema, A_1, \dots, A_j são conjuntos nebulosos nos universos definidos de cada variável de entrada, y_1 e y_2 são variáveis de saída e B_1, \dots, B_m são conjuntos nebulosos nos universos definidos de cada variável de saída.

Durante o processo de Fuzificação, os antecedentes de cada regra são processados por meio da intersecção entre os graus de pertinência das entradas atuais nos termos primários definidos em cada uma. Este processo gera um grau de pertinência de disparo para cada regra de produção. Em outras palavras, é calculado para a k -ésima regra da Base de Regras um coeficiente de disparo $D^{(k)}$, onde os índices k nos conjuntos nebulosos denotam os termos primários que compõem a regra k na Base de Regras. Este processamento transforma informações quantitativas em informações qualitativas, e é considerado um processo de generalização, sendo definido abaixo.

$$D^{(k)} = T[\mu_{A_1^k}(x_1), \mu_{A_2^k}(x_2), \dots, \mu_{A_p^k}(x_p)] = \min[\mu_{A_1^k}(x_1), \mu_{A_2^k}(x_2), \dots, \mu_{A_p^k}(x_p)]$$

Todas as regras para as quais o coeficiente de disparo for maior que zero são ditas regras que dispararam para as entradas atuais. Isto quer dizer que elas vão contribuir para o cálculo da saída correspondente do sistema de inferência. Os coeficientes de disparo, por sua

vez, vão limitar os valores máximos dos conjuntos nebulosos de saída gerados por estas regras. Finalmente, uma operação global de união vai compor um conjunto nebuloso para cada variável de saída, contendo informações sobre todas as regras disparadas para as entradas atuais.

Abaixo, é mostrada a composição deste conjunto para o caso da saída y_2 , apresentada na regra de produção nebulosa genérica anterior. O universo de discurso desta variável de saída é composto pelos elementos y e U_{y_2} .

$$\mu_{B_i}(y) = S_{k=1..n} [T(D^{(k)}, \mu_{B_i}(y))] = \max_{k=1..n} [\min(D^{(k)}, \mu_{B_i}(y))] \forall y \in U_{y_2}$$

Todo este processo é ilustrado pela Figura 4-5. Nesta ilustração, são considerados quatro conjuntos nebulosos A_1 , A_2 , B_1 e B_2 e dois valores associados às variáveis lingüísticas X_a e X_b . Analisando os valores de X_a e X_b nos conjuntos A_1 e B_1 , respectivamente, o Modelo de Inferência de Mamdani processa as regras existentes na base de regras, tais como “SE (X_a é A_1) E (X_b é B_1) ENTÃO (α_1 é C_1)” e seleciona o mínimo valor de pertinência obtido (Min), no caso de X_b , gerando o conjunto nebuloso C_1 . Do mesmo modo com A_2 e B_2 , o menor valor de pertinência é o de X_a , o que resulta no conjunto C_2 . O conjunto nebuloso de saída C' será a união dos conjuntos C_1 e C_2 (Max).

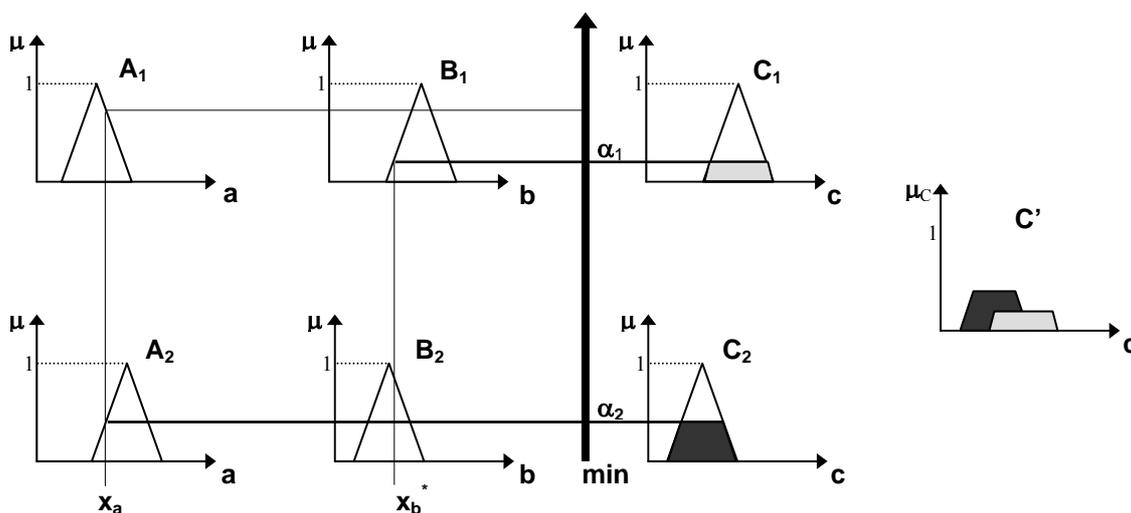


Figura 4-5: Modelo de Inferência de Mamdani

O processo de inferência acima descrito transforma uma informação qualitativa em outra informação qualitativa, por meio de uma conversão. O conjunto nebuloso gerado

durante o processo de inferência pode então ser utilizado diretamente em um diagnóstico qualitativo de tomada de decisão, ou será convertido em um valor escalar proporcional para atuação externa por meio de atuadores convencionais, no caso de um sistema de controle. Esta conversão é feita através do módulo de Desfuzificação.

4.2.4 Desfuzificação

O Módulo de Desfuzificação transforma informações qualitativas em uma informação quantitativa, para se obter uma saída final *não*-nebulosa, sendo um processo de especificação. Dentre os métodos mais utilizados para o processo de Desfuzificação, destacam-se o Método do Centro de Massa ou Centróide e o Método do Centro de Máxima.

O Método do Centro de Massa ou Centróide calcula para um dado conjunto nebuloso de saída (**C'**), proveniente de uma Base de Regras processada, a abscissa (no universo de discurso definido para a variável em questão) do ponto de centro de massa correspondente (**COG**), e se utiliza como valor escalar de saída, como mostra a Figura 4-6. O centro de massa pode ser definido como o ponto no centro da distribuição de massa de um objeto, onde se pode considerar que esteja concentrada toda a sua massa.

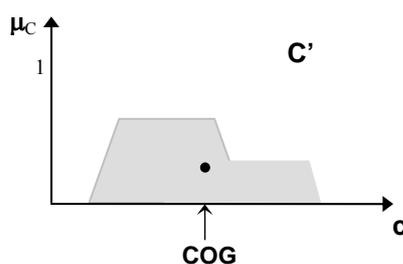


Figura 4-6: Método do Centro de Massa

No Método do Centro de Máxima o valor desfuzificado é definido como a média aritmética entre o menor e o maior valor encontrado para μ_c . A Figura 4-7 mostra o esboço deste método, no qual o valor desfuzificado resultante (**CM**) tende ao centro da área com maior valor de μ_c , quase que desprezando a área com menor valor de pertinência.

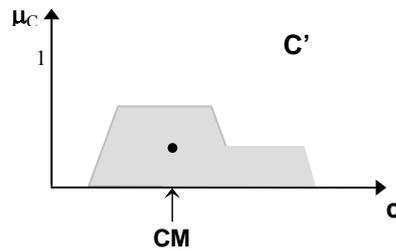


Figura 4-7: Método do Centro de Máxima

Foram implementados, a título de verificação, os dois métodos, Centróide e Centro de Máxima, apresentados. O Método do Centróide foi escolhido por mais se aproximar do resultado real, em função de não desprezar em nada a área com menor valor de pertinência, como ocorreu com o método do Centro de Máxima. Sua expressão analítica de implementação é apresentada abaixo:

$$Z = \frac{\sum_{i=0}^m \mu_c(Z_i) * Z_i}{\sum_{i=0}^m \mu_c(Z_i)}$$

Nesta expressão, Z é o valor desfuzificado, ‘m’ é o número de intervalos de quantização da saída, Zi é o valor da variável de saída para o intervalo de quantização ‘i’, e ‘μc (Zi)’ seu grau de pertinência.

4.3 Considerações Finais

Este capítulo apresentou definições de Lógica Nebulosa e descreveu os objetivos e as etapas de um Sistema Nebuloso. De forma concisa, como geralmente os dados de entrada do sistema são valores precisos, resultantes de medições ou observações, é necessário efetuar-se um mapeamento destes dados precisos para os conjuntos nebulosos de entrada, o que é realizado no Módulo de Fuzificação. No estágio seguinte, ocorre a ativação das regras relevantes para uma dada situação. Uma vez obtido o conjunto nebuloso de saída através do

processo de inferência, é efetuada uma interpretação dessa informação no estágio de Desfuzificação.

Em sua aplicação, o Sistema Nebuloso mostrou-se viável e operacional quando aplicado à avaliação de Qualidade de Serviço, resultando em duas respostas que representam o objetivo deste trabalho: um valor escalar e um termo lingüístico. O valor escalar é representativo da QoS no nível de rede e é obtido através do método de desfuzificação Centro de Massa, podendo ser analisado nos domínios dos conjuntos nebulosos de saída definidos ou ser utilizado por algum sistema que promova renegociação dinâmica, caso seja necessário. O termo lingüístico demonstra a qualidade de serviço de forma que haja entendimento por parte dos diversos níveis de usuário que possam estar utilizando o sistema.

Capítulo 5

Algoritmos Genéticos

Muitos problemas computacionais requerem uma busca através de um alto número de possíveis soluções, chamado de *espaço de busca*. O termo *espaço de busca* refere-se a uma coleção de soluções candidatas para um determinado problema [Mitchell 1996]. Um exemplo é o problema de encontrar computacionalmente uma forma para construir um conjunto de regras ou equações que predizem os altos e baixos de um mercado financeiro, tal como o valor de uma moeda estrangeira.

Os Algoritmos Genéticos (AG's) introduzidos por [Holland 1975] são algoritmos de busca baseados no mecanismo da seleção natural e da genética natural. A idéia que envolve os AG's é a de imitar o que a natureza faz. Em outras palavras, os algoritmos genéticos simulam processos naturais de sobrevivência e reprodução dos seres vivos que são essenciais em sua evolução. Na natureza, os indivíduos de uma mesma população competem entre si, buscando principalmente a sobrevivência, e, quanto melhor um indivíduo se adaptar ao meio em que vive, maior será sua chance de sobreviver e gerar descendentes.

Para reproduzir estes processos naturais, os algoritmos genéticos retiram cromossomos (cadeias de símbolos de um alfabeto) de uma população de cromossomos e os insere para uma nova população, usando um tipo de seleção natural juntamente com os operadores de cruzamento e mutação inspirados na genética. Cada cromossomo consiste de genes, que são parâmetros codificados nos cromossomos, ou seja, um elemento do vetor (bit) que representa o cromossomo, e cada gene representa uma instância de um alelo particular (valores que o gene pode assumir). O operador de seleção escolhe, dentre os cromossomos da população, aqueles que irão se reproduzir. Em média, os cromossomos mais fortes (melhores adaptados) produzem mais descendentes do que aqueles mais fracos (menos adaptados). Os

cromossomos selecionados podem sofrer modificações em suas características através dos operadores de cruzamento e mutação, gerando descendentes para a próxima geração [Mitchell 1996].

Em se tratando de propósitos computacionais, um algoritmo genético simples possui uma estrutura conforme o pseudocódigo a seguir [Michalewicz 1996]:

Algoritmo Evolução

```
t ← 0
inicializar P(t)
avaliar P(t)
enquanto (condição verdadeira) faça
    t ← t + 1
    gerar P(t) de P(t - 1)
    alterar P(t)
    avaliar P(t)
fim enquanto
```

Fim

Durante a iteração t , o algoritmo genético mantém uma população de soluções candidatas (cromossomos), $P(t) = \{x_1^t, \dots, x_n^t\}$. Cada solução x_i^t é avaliada para medir sua aptidão (*fitness*), ou seja, a qualidade da solução do problema representada por este cromossomo. Então, uma nova população (iteração $t + 1$) é formada pela seleção dos indivíduos (cromossomos) mais aptos. Alguns membros desta nova população serão selecionados para fazer parte da chamada ‘população intermediária’, a qual sofrerá alterações devido à ação dos operadores genéticos de cruzamento e mutação, enquanto outros permanecerão intactos. O cruzamento combina as características de dois cromossomos pais para formar dois cromossomos filhos. O objetivo da aplicabilidade do cruzamento é trocar informações entre soluções em potencial. A mutação altera aleatoriamente um ou mais genes de um cromossomo selecionado, com o intuito de introduzir informação extra para a população.

Segundo [Michalewicz 1996], para a implementação de um algoritmo genético deve-se primeiramente definir alguns aspectos importantes, tais como uma representação genética para as soluções do problema, uma forma de criar uma população inicial das soluções, uma função de avaliação que desempenha o papel do ambiente, operadores genéticos que alteram a composição dos filhos e os valores para os parâmetros utilizados pelo algoritmo genético.

Em se tratando da representação genética para as soluções do problema, a representação binária (cadeias de zeros e uns) é a tradicionalmente usada, uma vez que é de fácil utilização e manipulação, além de simples de analisar teoricamente. Contudo, apresenta algumas desvantagens quando aplicada a problemas multidimensionais e a problemas numéricos de alta precisão [Michalewicz 1996].

A escolha de uma população inicial maior que a população a ser utilizada nas gerações subsequentes pode melhorar a representação do espaço de busca. A população inicial pode ser gerada de várias maneiras. Se uma população inicial pequena for gerada aleatoriamente, provavelmente algumas regiões do espaço de busca não serão representadas [Lacerda e Carvalho 1980].

Uma função de avaliação deve desempenhar o papel do ambiente, ou seja, avaliar as soluções em termos de aptidão. Alguns cuidados devem ser tomados para que cromossomos idênticos não sejam avaliados mais de uma vez [Lacerda e Carvalho 1980], tais como: evitar gerar cromossomos idênticos na população inicial, verificar se foi aplicado cruzamento ou mutação nos pais, verificar se o filho é igual a um dos pais, manter a população com todos os cromossomos distintos entre si e, antes de avaliar um filho, verificar se um cromossomo idêntico já existe na população.

Os operadores genéticos que alteram a composição dos filhos são os operadores de Cruzamento e Mutação. O operador de cruzamento tem o intuito de trocar informações entre diferentes soluções em potencial. Já a aplicabilidade do operador de mutação, tem por objetivo introduzir uma variabilidade extra na população [Michalewicz 1996].

Dentre os parâmetros utilizados pelo AG destacam-se: tamanho da população, probabilidades de aplicação dos operadores genéticos e número de gerações. A influência de cada parâmetro no desempenho do algoritmo depende da classe de problemas que se está tratando. Assim, a determinação de um conjunto de valores otimizado para estes parâmetros dependerá da realização de um grande número de experimentos e testes. Na literatura os valores encontrados estão na faixa de 60% a 65% para a probabilidade de cruzamento e entre 0,1% e 5% para a probabilidade de mutação. O tamanho da população e o número de gerações dependem da complexidade do problema de busca e devem ser determinados experimentalmente. No entanto, deve ser observado que o tamanho da população e o número de gerações definem diretamente o tamanho do espaço de busca a ser coberto. Existem estudos que utilizam um AG como método de otimização para a escolha dos parâmetros de

outro AG [Grefenstette 1986] [Bramlette 1991], devido à importância da escolha correta destes parâmetros.

5.1 Métodos de Seleção

Dada uma população em que a cada indivíduo foi atribuído um valor de aptidão, existem vários métodos para selecionar os indivíduos sobre os quais serão aplicados os operadores genéticos de cruzamento e mutação. Estes indivíduos selecionados formarão uma população, que é conhecida como população intermediária. A maioria dos métodos de seleção são projetados para escolher preferencialmente indivíduos com maiores notas de aptidão, embora não exclusivamente, a fim de manter a diversidade da população. A seguir veremos dois destes métodos, com base em [Michalewicz 1996].

5.1.1 Seleção pela Roleta

O método de seleção da Roleta é um método de seleção muito utilizado, onde os indivíduos de uma geração são escolhidos para fazer parte da população intermediária através de um sorteio de roleta. Neste método cada indivíduo da população é representado na roleta proporcionalmente ao seu índice de aptidão, como é mostrado na Figura 5-1. Assim, aos indivíduos com alta aptidão é dada uma porção maior da roleta, enquanto aos de aptidão mais baixa é dada uma porção relativamente menor da roleta. Finalmente, a roleta é girada um determinado número de vezes, dependendo do tamanho da população, e são escolhidos como indivíduos que participarão da população intermediária aqueles sorteados na roleta.

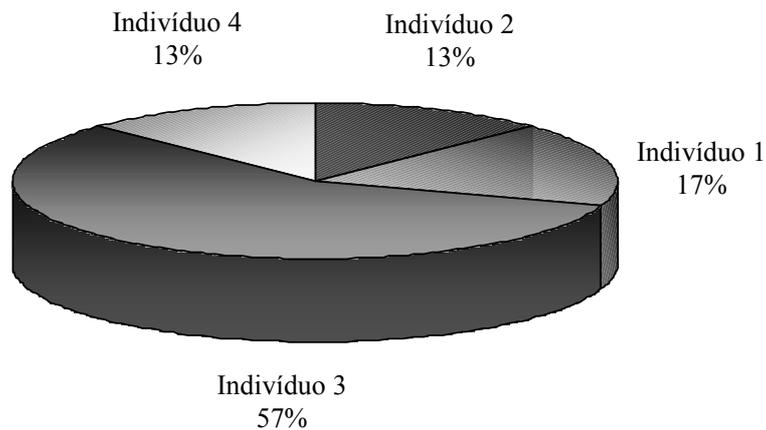


Figura 5-1: : Exemplo de uma Roleta de seleção

Um exemplo da implementação em pseudocódigo deste método, segundo [Mitchell 1996], é descrito a seguir:

Algoritmo Seleção

```

T ← soma dos valores de aptidão de todos os indivíduos da população
repita n vezes para selecionar n indivíduos
  r ← valor aleatório entre 0 e T
  Percorra seqüencialmente os indivíduos da população, acumulando
  em S o valor de aptidão dos indivíduos já percorridos
  se S >= r então
    selecione o indivíduo corrente
  fim se
fim repita

```

Fim

5.1.2 Seleção Elitista ou Elitismo

A seleção elitista ou elitismo, introduzido por [Dejong 1975], pode ser usado em conjunto com outro método de seleção, obrigando o algoritmo genético a reter um determinado número de melhores indivíduos em cada geração. A aplicação do elitismo previne que tais indivíduos sejam destruídos pela aplicação dos operadores de cruzamento ou mutação [Mitchell 1996].

5.2 Operadores Genéticos

Os operadores genéticos de cruzamento e mutação provocam alterações em uma população.

A mutação é necessária para a introdução e manutenção da diversidade genética da população, alterando arbitrariamente um ou mais indivíduos, fornecendo assim, meios para a introdução de novos indivíduos na população. O operador de mutação é aplicado aos indivíduos com uma probabilidade dada pela taxa de mutação. Uma baixa taxa de mutação pode assegurar a diversidade na população, mas o contrário, pode destruir toda a informação contida no indivíduo que foi adquirida durante as gerações passadas.

O operador de cruzamento é aplicado aos cromossomos de acordo com uma probabilidade, conhecida como taxa de cruzamento. Quanto maior for esta taxa, mais rapidamente novos indivíduos serão introduzidos na população. Mas se esta for muito alta, indivíduos com boas aptidões poderão ser retirados mais rapidamente, assim, a maior parte da população será substituída e poderá ocorrer a perda de indivíduos com alta aptidão. Com um valor baixo, o algoritmo pode tornar-se muito lento. Assim, as taxas de mutação e cruzamento devem ser definidas empiricamente para cada domínio de problema, levando em conta essas considerações.

A escolha dos operadores genéticos está intimamente ligada à codificação adotada para a representação genética [Delgado 2002], ou seja, existe uma variação nos comportamentos dos operadores, os quais estão relacionados à codificação empregada nos cromossomos. Em outras palavras, há operadores genéticos que foram concebidos para o uso com codificação binária e operadores genéticos para o uso com codificação real ou inteira.

Com base em [Michalewicz 1996], serão definidos apenas os operadores genéticos para codificação binária, sendo estes suficientes para o desenvolvimento deste trabalho.

5.2.1 Operadores Genéticos para Codificação Binária

Cruzamento de um-ponto

O cruzamento de um-ponto é a forma mais simples deste operador. É aplicado a um par de cromossomos retirados da população intermediária, gerando dois cromossomos filhos. Cada um dos cromossomos pais é cortado em uma posição aleatória produzindo duas cabeças e duas caudas. As caudas são trocadas gerando dois novos cromossomos. A Figura 5-2 ilustra o comportamento deste operador.

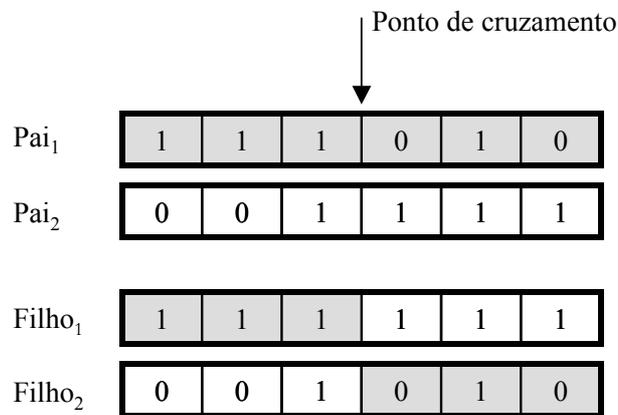


Figura 5-2: Cruzamento de um-ponto

Cruzamento de n-pontos

O cruzamento de n-pontos funciona da mesma forma que o cruzamento de um-ponto, porém, neste há vários pontos de corte nos cromossomos pais. Na Figura 5-3 é exemplificado o cruzamento de três-pontos.

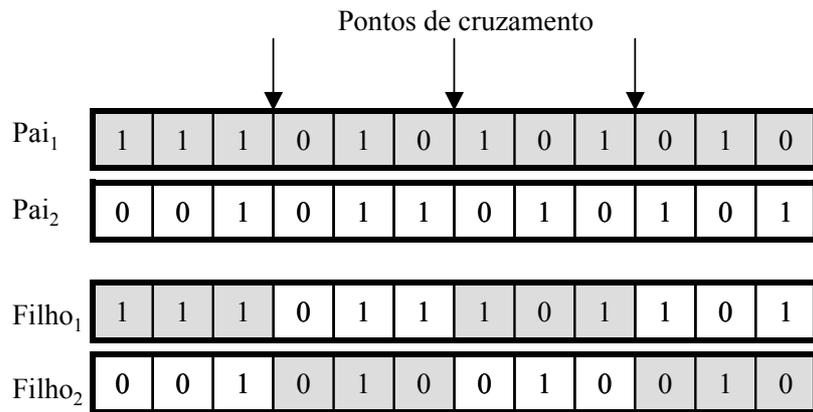


Figura 5-3: Cruzamento de três-pontos

Cruzamento uniforme

No cruzamento uniforme cada gene do cromossomo filho é criado copiando o gene correspondente de um dos pais, escolhido de acordo com uma máscara de cruzamento gerada aleatoriamente. Onde houver '1' na máscara de cruzamento, o gene correspondente será copiado do primeiro pai e onde houver '0' será copiado do segundo. O processo é repetido com os pais trocados para produzir o segundo descendente. Uma nova máscara de cruzamento é criada para cada par de pais. O número de pontos de troca de informação não é fixo, mas em geral é usado $L/2$ (onde L é o comprimento do cromossomo). Na Figura 5-4 o processo é demonstrado graficamente.

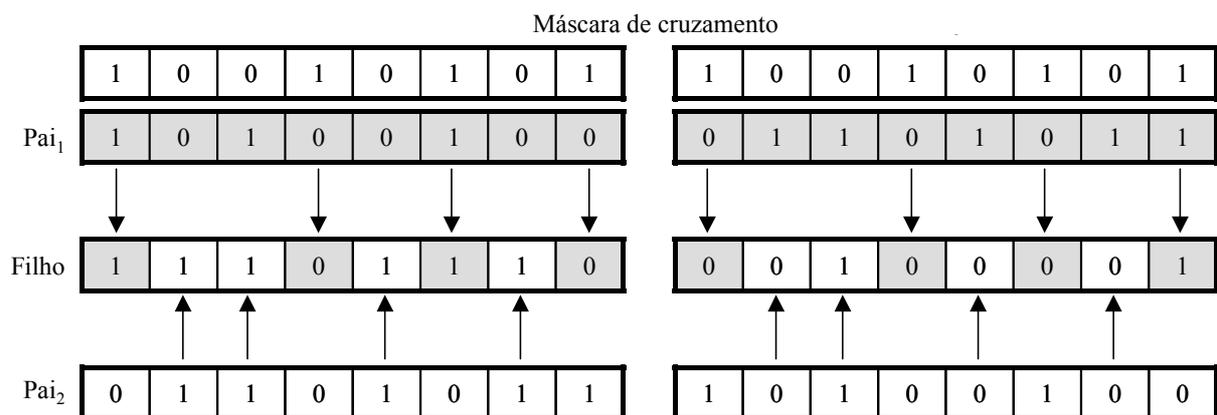


Figura 5-4: Cruzamento uniforme

Estes três tipos de cruzamento foram concebidos para uso com codificação binária. Entretanto, não existe nenhum impedimento quanto ao uso destes operadores junto a outros

tipos de codificação, existindo porém operadores de cruzamento específicos para problemas com restrições [Delgado 2002].

Mutação simples

O operador de mutação simples é aplicado aleatoriamente em alguns genes do cromossomo que sofrerá a mutação. Este processo inverte os valores dos genes, ou seja, muda o valor de um dado gene de '1' para '0' ou de '0' para '1'. Na Figura 5-5 é ilustrado o comportamento deste operador.

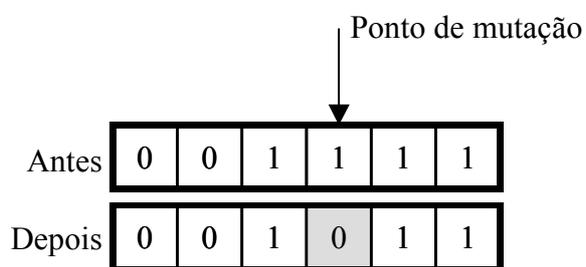


Figura 5-5: Mutação simples

5.3 Considerações Finais

Este capítulo descreveu uma visão geral sobre os Algoritmos Genéticos e seus componentes, pois esta metodologia será utilizada neste trabalho para otimizar a base de regras do sistema nebuloso, como mencionado anteriormente. Pesquisas em todo o mundo surgem objetivando explorar ainda mais esse método, de forma a otimizar os componentes de um sistema nebuloso [Jain et al. 1999]. Nos parágrafos seguintes, alguns exemplos são apresentados.

Uma arquitetura de sistema de gerenciamento para manter a QoS em um ambiente DiffServ é explicitada em [Fernandez et al. 2003]. O sistema nebuloso de controle, otimizado com a aplicação do algoritmo genético, mostrou-se mais eficiente que o controlador convencional para qualquer intervalo de operação. Isso pode ser evidenciado através dos valores obtidos para as duas variáveis de análise: delay e perda de pacotes. Suas taxas foram reduzidas em 66% e 33%, respectivamente. Assim, após a realização de diversos testes, foi

possível escolher as melhores combinações de parâmetros, ou seja, a melhor base de regras para problema em questão.

O trabalho de [Gonçalves et al. 2001] traz a implementação de uma ferramenta para a otimização de parâmetros de controladores nebulosos e de sistemas especialistas reativos utilizando algoritmos genéticos. Esta ferramenta permite a utilização de algoritmos genéticos para determinar os valores ótimos para os parâmetros dos controladores nebulosos. A análise comparativa é feita sobre uma função de erro estabelecida, tendo seu valor reduzido em 28% após a aplicação do algoritmo genético. Com a utilização do sistema nebuloso com menos regras, a possibilidade de confusão e erro por parte do mesmo diminui, tornando-o mais eficiente e rápido.

Em [Medeiros et al. 2001] verifica-se que a utilização de algoritmos genéticos realmente é uma técnica eficiente para determinar parâmetros de um sistema nebuloso, fornecendo resultados extremamente satisfatórios.

O trabalho desenvolvido por [Pal e Pal 2003], retrata um método para construção e otimização da bases de regras de sistemas nebulosos de controle, mostrando que a redução do número de regras aumenta o tempo de resposta e a eficiência do sistema.

No próximo capítulo, será apresentada a descrição do método desenvolvido neste trabalho, o qual utiliza os paradigmas de Inteligência Computacional, apresentados nos capítulos 4 e 5, aplicados à análise de qualidade de serviço em Redes de Computadores. Métodos que integram mais de um tipo de sistema inteligente, unindo os méritos de um para compensar deficiências isoladas, são ferramentas poderosas para superar desafios [Jain et al. 1999].

Capítulo 6

Método de Avaliação de Qualidade de Serviço em Sistemas Multimídia Utilizando um Sistema Genético Nebuloso Baseado em Regras

Existem duas motivações principais para a avaliação do ambiente de rede considerando QoS, quais sejam: primeiro, determinar quão bem pode-se entregar serviços que correspondem às necessidades particulares de um usuário. De fato, isto é a medição do nível de consistência entre o ambiente que oferece QoS e as necessidades do usuário; e segundo, o usuário pode estar interessado em medir o sucesso ou falha do provedor de rede em satisfazer os parâmetros de QoS acordados.

Um usuário de uma rede remota pode também estar interessado em conhecer a transitoriedade de um ambiente de QoS, de modo a verificar se os níveis de QoS de uma transação são mantidos quando esta atravessa a rede de diversos provedores e se os mecanismos de QoS utilizados em cada rede são compatíveis.

O administrador da rede também está interessado em avaliar parâmetros do ambiente com QoS, mas por razões diferentes do usuário. O administrador está interessado em analisar o desempenho que o ambiente com QoS está entregando ao usuário, para assegurar que os níveis de serviços estão dentro dos parâmetros acordados. Os usuários estão limitados a obter os efeitos da arquitetura de QoS através da avaliação de dados de desempenho do sistema final que gerou a transação na rede e não podem ver diretamente o comportamento dos dispositivos internos da rede.

Tipicamente, a qualidade fim-a-fim de uma conexão depende de todas as métricas dos parâmetros envolvidos. Como mencionado no item 3.3.2, os parâmetros mais relevantes, a serem considerados em aplicações multimídia são os Parâmetros de Rede, tais como: vazão, delay, jitter e perda de pacotes. Definir um único valor de QoS, em uma linguagem acessível ao usuário, se torna importante para refletir a qualidade de serviço que pode ser oferecida pela rede, possibilitando que o mesmo possa tomar decisões cabíveis ao tipo de aplicação que deseja executar na rede em questão.

6.1 Transformação dos Parâmetros de QoS

Para alcançar a satisfação do usuário no que se refere a um determinado serviço, é essencial a correta transformação dos parâmetros de QoS quantificados, desde o usuário final até a rede de transmissão e vice-versa, objetivando facilitar o entendimento dos parâmetros por parte dos usuários. Com este intuito, surgiu o conceito de Mapeamento de Parâmetros de QoS, o qual pode ser definido como: “Transformação da especificação de qualidade de serviço entre os diferentes níveis de qualidade de serviço” [Huard e Lazar 1997].

O fator essencial desse tópico e relevante para essa pesquisa é a necessidade da realização de um mapeamento ou transformação de parâmetros entre os diversos níveis de QoS. Essa transformação acontece desde a camada do usuário até a camada de rede.

Vale a pena ressaltar que, para uma correta e eficiente transformação de parâmetros de QoS, é necessário levar em conta o tipo da aplicação. Neste trabalho, a transformação dos parâmetros foi feita apenas para quatro aplicações multimídia: vídeo sob demanda, videoconferência, áudio sob demanda e áudioconferência, podendo ser facilmente estendida a outras aplicações com base em novas investigações a respeito de mapeamento de parâmetros de qualidade de serviço.

Em [Price 2001], é mostrada uma tabela contendo valores esperados para os parâmetros de QoS para que haja satisfação por parte do usuário. A Tabela 6-1 ilustra estes valores.

Tabela 6-1: Valores esperados de alguns parâmetros de QoS

	Parâmetro	Valor considerado aceitável
Vídeo sob Demanda	Perda de Pacotes(%)	Bem menor que 5 ($\cong 3$)
	Vazão(Kbits/s)	>30
Videoconferência	Delay(ms)	Entre 150 e 400
	Perda de Pacotes(%)	<2
	Vazão(Kbits/s)	>100
Áudio sob Demanda	Perda de Pacotes(%)	<2
	Vazão(Kbits/s)	>30
Audioconferência	Vazão(Kbits/s)	>100

No trabalho de [Busse et al. 1995] apud [Lunardi e Dotti 2001], são definidos domínios para o parâmetro perda de pacotes em uma aplicação de vídeo sob demanda. O estado da rede é considerado não carregado caso a perda de pacotes seja inferior a 2%. Se estiver entre 2% e 4%, é declarado como carregado, e caso seja superior a 4% é tido como congestionado.

Nas pesquisas realizadas por [ITU 2001] e [Cotroneo et al. 2003], foram definidas tabelas para expressar os valores aceitáveis dos parâmetros de qualidade de serviço. Essas tabelas se confirmam e são reunidas na Tabela 6-2.

Tabela 6-2: Valores dos parametros de QoS

	Parâmetro	Valor considerado aceitável
Vídeo sob Demanda	Delay(s)	<10
	Perda de Pacotes(%)	<1
Videoconferência	Delay(ms)	Entre 150 e 400
Áudio sob Demanda	Perda de Pacotes(%)	<3
	Vazão(Kbits/s)	>32
Audioconferência	Delay(ms)	Entre 150 e 400
	Perda de Pacotes(%)	<3
	Vazão(Kbits/s)	>64

[Sharda 2003] apresenta também apresenta alguns valores dos parâmetros relacionados a aplicações de vídeo, como mostrado na Tabela 6-3.

Tabela 6-3: Valores aceitáveis para os parâmetros de QoS

	Parâmetro	Valor considerado aceitável
Vídeo sob Demanda	Jitter(s)	<1
	Vazão(Kbits/s)	>112
Videoconferência	Delay(ms)	Entre 100 e 500
	Vazão(Kbits/s)	>112

Em [T-Mobil 1999] é mostrada uma tabela contendo faixas nas quais os parâmetros de QoS são considerados aceitáveis para uma transmissão. A Tabela 6-4 apresenta estas faixas.

Tabela 6-4: Faixas aceitáveis para os parâmetros de QoS

	Parâmetro	Valor considerado aceitável
Vídeo sob Demanda	Vazão(Kbits/s)	>64
Videoconferência	Delay(ms)	<200
	Jitter(ms)	Baixo (<200)
	Vazão(Kbits/s)	>64
Áudio sob Demanda	Delay(s)	<6
	Jitter(s)	<5
	Vazão(Kbits/s)	>32
Audioconferência	Jitter(ms)	<3
	Vazão(Kbits/s)	>64

Nos experimentos de [Dressler 2003b], foram considerados aceitáveis os valores dos parâmetros definidos pela Tabela 6-5.

Tabela 6-5: Valores dos parâmetros de QoS considerados aceitáveis

	Parâmetro	Valor considerado aceitável
Vídeo sob Demanda	Delay(s)	<4
	Jitter(s)	<3
	Perda de Pacotes(%)	<1
Videoconferência	Delay(ms)	<200
	Jitter(ms)	<150
	Perda de Pacotes(%)	<1

Por fim, os trabalhos de [Bochmann 2002] trazem os valores ideais para os parâmetros delay e perda de pacotes, os quais são respectivamente menores que 100 ms e menores que 1%.

Reunindo as pesquisas realizadas por [Price 2001]; [Busse et al. 1995] apud [Lunardi e Dotti 2001];[ITU 2001];[Cotroneo et al. 2003];[Sharda 2003];[T-Mobil 1999];[Dressler 2003b];[Bochmann 2002] foi definida a Tabela 6-6, que explicita os parâmetros de Qualidade de Serviço para as aplicações citadas acima, mapeados diretamente do nível do Usuário para o nível de Rede. Nestes mesmos trabalhos podem ser encontradas as definições dos parâmetros de QoS para outras aplicações multimídia, tais como videofone e jogos iterativos.

Tabela 6-6: Mapeamento dos parâmetros de QoS - Usuário/Rede

		QUALIDADE		
		Intolerável	Aceitável	Ótimo
Vídeo sob Demanda	Delay(s)	>6	3-6	0-3
	Jitter(s)	>5	1-5	0-1
	Perda de Pacotes(%)	>3	1-3	0-1
	Vazão(Kbits/s)	0-64	64-112	>112
Videoconferência	Delay(ms)	>350	150-350	0-150
	Jitter(ms)	>250	100-250	0-100
	Perda de Pacotes(%)	>2	1-2	0-1
	Vazão(Kbits/s)	0-64	64-112	>112
Áudio sob Demanda	Delay(s)	>6	3-6	0-3
	Jitter(s)	>5	1-5	0-1
	Perda de Pacotes(%)	>2	1-2	0-1
	Vazão(Kbits/s)	0-32	32-64	>64
Áudioconferência	Delay(ms)	>400	150-400	0-150
	Jitter(ms)	>3	1-3	0-1
	Perda de Pacotes(%)	>3	1-3	0-1
	Vazão(Kbits/s)	0-64	64-128	>182

Estes são os parâmetros típicos envolvendo as aplicações descritas na Tabela 6-6. Para uma outra classe de problemas, deve-se definir todos os parâmetros necessários para avaliação da QoS, bem como suas faixas de variação nas três classes de qualidade apresentadas.

A tabela de Mapeamento dos Parâmetros de QoS – Usuário/Rede é a primeira tarefa a ser realizada pelo método de avaliação de QoS proposto, sendo validada de forma experimental para cada um dos parâmetros individualmente. Esse mapeamento dos parâmetros de um valor numérico para uma linguagem de entendimento do usuário, e vice-versa, é feito com base em políticas ou regras estabelecidas. A especificação dessas políticas traduz uma decisão administrativa na gerência do sistema.

6.2 Aplicação do Sistema Genético Nebuloso Baseado em Regras na Avaliação de QoS

Por serem as regras de políticas abstratas e próximas da percepção humana, é muito difícil mapeá-las em regras computacionais, absolutas e exatas por natureza. Esta dificuldade está no fato de que definir intervalos exatos de valores para uma variável pode prejudicar ou até mesmo inutilizar um resultado obtido com base em uma análise feita nestes intervalos. A Lógica Nebulosa tem a característica de tratar variáveis semânticas com certo grau de imprecisão. Por isso, é quase intuitiva a aproximação das regras de especificação de políticas de gerenciamento, dos atributos de um Sistema Nebuloso [Fernandez et al. 2003].

6.2.1 Módulo de Fuzificação

As entradas para o sistema são os valores obtidos para os parâmetros de Qualidade de Serviço, através da medição da rede. Após serem mensurados, os valores referentes à vazão, delay, jitter e perda de pacotes são repassados ao Sistema Nebuloso e serão convertidos em termos lingüísticos. Para que esta conversão seja feita, três conjuntos nebulosos foram definidos para cada uma das variáveis de entrada, sendo eles: Intolerável, Aceitável e Ótimo. Estes conjuntos estão representados na forma trapezoidal (sendo possível

representar todo um intervalo de pontos de máximo), e são caracterizados por uma função de pertinência que mapeia os valores possíveis para os parâmetros de entrada no intervalo [0,1].

Os parâmetros das funções de pertinência levam em consideração os pontos da função trapezoidal, definidos pela Figura 4-2 e o mapeamento dos valores dos parâmetros de QoS apresentados na Tabela 6-6, podendo também ser definidos para outras aplicações multimídia apenas estendendo a Tabela 6-6, inserindo mapeamento dos parâmetros de qualidade de serviço para outras aplicações. A definição destes parâmetros é feita como mostrado a seguir. Considere, por exemplo, o parâmetro Vazão considerado na aplicação de videoconferência. Temos definido na Tabela 6-6, os intervalos considerados Intoleráveis (0 a 64), Aceitáveis (64 a 112) e Ótimos (>112). Com o objetivo de colocar esses intervalos como parâmetros para as funções trapezoidais mostradas na Figura 4-2, devem ser definidos valores para as variáveis **a**, **m**, **n** e **b**, distribuídos uniformemente dentro do domínio pretendido para cada trapézio. Assim, o valor '0' é considerado como o limite inferior da variável vazão e o valor 224 seu limite superior, tendo sido obtido como o dobro do maior valor definido para o ponto **m**, ou seja, 112. Os valores 32 e 88 representam os pontos médios dos intervalos (0 a 64) e (64 a 112), respectivamente. Para os demais parâmetros de qualidade de serviço foi empregada a mesma sistemática.

Nas tabelas abaixo são mostrados os parâmetros das funções de pertinência para cada variável de entrada, levando em consideração cada aplicação analisada nos estudos de caso, conforme definido no item 1.3.

Aplicação de Videoconferência

Tabela 6-7: Parâmetros das funções de pertinência para a variável de entrada VAZÃO

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	88	112	224	224
Conjunto Aceitável	32	64	88	112
Conjunto Intolerável	0	0	32	64

Tabela 6-8: Parâmetros das funções de pertinência para a variável de entrada DELAY

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	0	0	75	150
Conjunto Aceitável	75	150	275	350
Conjunto Intolerável	275	350	700	700

Tabela 6-9: Parâmetros das funções de pertinência para a variável de entrada JITTER

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	0	0	50	100
Conjunto Aceitável	50	100	200	250
Conjunto Intolerável	200	250	500	500

Tabela 6-10: Parâmetros das funções de pertinência para a variável de entrada PERDA DE PACOTES

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	0	0	0.5	1
Conjunto Aceitável	0.5	1	1.5	2
Conjunto Intolerável	1.5	2	4	4

Aplicação de Vídeo sob Demanda

Tabela 6-11: Parâmetros das funções de pertinência para a variável de entrada VAZÃO

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	88	112	224	224
Conjunto Aceitável	32	64	88	112
Conjunto Intolerável	0	0	32	64

Tabela 6-12: Parâmetros das funções de pertinência para a variável de entrada DELAY

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	0	0	1.5	3
Conjunto Aceitável	1.5	3	4.5	6
Conjunto Intolerável	4.5	6	12	12

Tabela 6-13: Parâmetros das funções de pertinência para a variável de entrada JITTER

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	0	0	0.5	1
Conjunto Aceitável	0.5	1	3	5
Conjunto Intolerável	3	5	10	10

Tabela 6-14: Parâmetros das funções de pertinência para a variável de entrada PERDA DE PACOTES

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	0	0	0.5	1
Conjunto Aceitável	0.5	1	2	3
Conjunto Intolerável	2	3	6	6

Com os parâmetros das funções de pertinência definidos, o processo de fuzificação ocorre como ilustrado na Figura 6-1. O eixo X dos gráficos é definido como Valores de Entrada de cada variável no sistema, e o eixo Y é definido como Valores de Pertinência de cada variável aos conjuntos nebulosos. As intersecções da reta que parte verticalmente do valor de entrada da variável (eixo X) com as retas que delimitam os conjuntos, devem ser projetadas no eixo Y, demonstrando o grau de pertinência da variável a cada conjunto nebuloso.

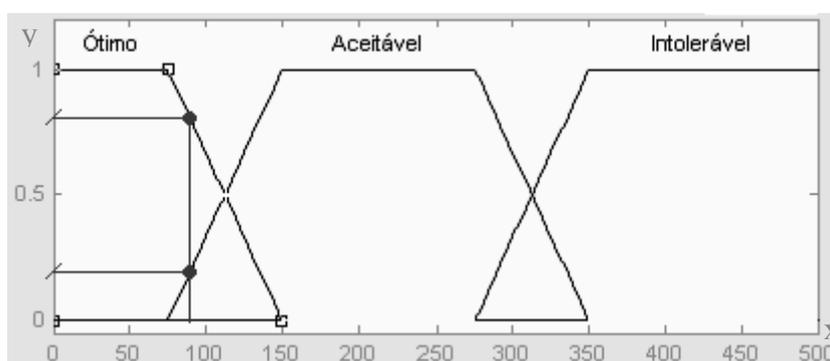


Figura 6-1. Exemplo de cálculo de pertinência

Supondo que o valor obtido para o delay seja de 90 milisegundos, seu grau de pertinência é de aproximadamente 0.85 no conjunto Ótimo, 0.2 no conjunto Aceitável e 0 no conjunto Intolerável. Essa análise é feita sucessivamente para cada um dos parâmetros e as pertinências são armazenadas para serem usadas posteriormente pelo motor de inferência.

6.2.2 Base de Regras

A partir da classificação de todos os valores dos parâmetros de QoS em três subconjuntos, é necessário quantificar de forma geral a qualidade da conexão para uma dada aplicação. Esse procedimento é parcialmente realizado pela base de regras. A base de regras é definida com a ajuda de um especialista e deve ser validada para garantir sua funcionalidade.

As regras definidas para o sistema desenvolvido possuem o seguinte formato (tendo como variantes apenas as classificações: Ótima, Aceitável e Intolerável):

SE (Vazão é Aceitável) E (Delay é Ótimo) E (Jitter é Ótima) E (Perda é Ótima)
ENTÃO (QoS é Ótima)

Deste modo, estas regras conseguem relacionar as condições aferidas nos diversos parâmetros de qualidade de serviço com uma avaliação global da QoS, podendo ser modeladas para as diversas aplicações multimídia existentes.

As Bases de Regras iniciais do sistema nebuloso, uma para a aplicação de Videoconferência e outra para a de Vídeo sob Demanda, foram construídas com a ajuda de um especialista em análise de desempenho em redes de computadores e posteriormente validadas, sendo moldadas por experimentação, utilizando-se o emulador de rede NIST Net [Carson e Santay 2003], uma ferramenta para emulação de redes IP baseada em Linux, de fácil instalação e de distribuição gratuita.

Por operar no nível de roteamento, o NIST Net pode emular características de desempenho fim-a-fim que limitam o funcionamento da rede. O conjunto de efeitos/parâmetros de rede configuráveis no NIST Net inclui: perda de pacotes, delay, jitter, limitações de largura de banda, entre outros; parâmetros estes necessários ao desenvolvimento deste trabalho.

A implementação do NIST Net consiste em um módulo de extensão para o kernel do sistema operacional Linux e uma interface gráfica, o XNIST Net, para o usuário da aplicação. Esta interface permite selecionar e monitorar o tráfego através da rede, possibilitando a aplicação de efeitos de desempenho aos pacotes IP do fluxo de informação. A Figura 6-2 mostra parcialmente a interface gráfica XNIST Net e explica a sua funcionalidade.

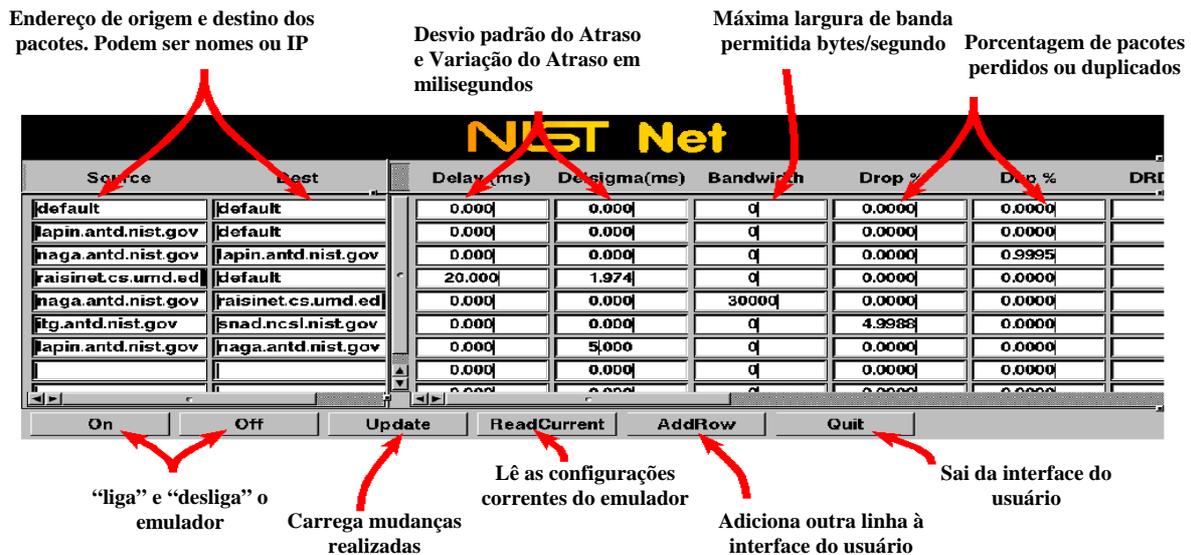


Figura 6-2: Interface gráfica XNIST Net

Os valores dos parâmetros são inseridos na tabela do XNIST Net, de acordo com o ambiente que se deseja emular, e o emulador é colocado para executar. Mudanças realizadas nos valores dos parâmetros podem ser aplicadas durante a execução.

Cada regra das bases de regras foi emulada individualmente e o resultado foi observado, sendo em seguida comparado ao que foi definido pelo especialista, com o objetivo de validar e ajustar as bases de regras iniciais. A Figura 6-3 apresenta um exemplo de enquadramento das imagens nas três categorias consideradas pelo sistema quando o parâmetro perda de pacotes estava sendo analisado. A primeira ilustra quais padrões de imagens foram considerados ótimos, a segunda, imagens consideradas aceitáveis e a terceira, os padrões definidos como intoleráveis. A exposição de todas as regras componentes das bases de regras e maiores comentários sobre as mesmas são apresentados no Apêndice A.

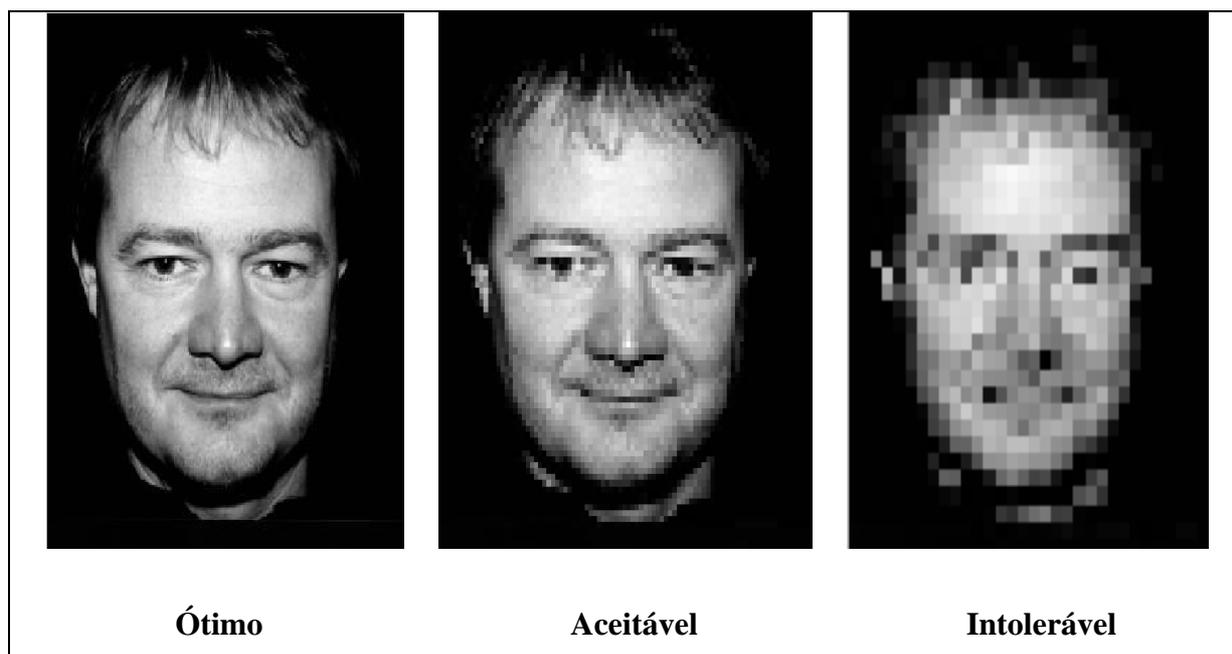


Figura 6-3: Imagens nas três categorias definidas pelo sistema

Estas bases de regras iniciais, apesar de validadas, são consideradas grandes em função do número de regras que possuem. O número de regras da base de regras é definido elevando-se o número de conjuntos nebulosos ao número de variáveis de entrada no sistema. Assim, considerando 3 conjuntos nebulosos (Ótimo, Aceitável e Intolerável) e 4 variáveis de entrada (vazão, delay, jitter e perda de pacotes), o número de regras é igual a 81. Estas 81 regras representam todas as combinações possíveis das variáveis de entrada do sistema com os conjuntos nebulosos definidos. Entretanto, algumas regras criadas não representam situações que verdadeiramente ocorrem no mundo real.

Portanto, quanto maior for a quantidade de conjuntos nebulosos e de variáveis de entrada, maior será o número de regras; o que aumenta exponencialmente a complexidade e o tempo de processamento do sistema nebuloso em função do aumento do espaço de busca do problema. Desta forma, considerou-se necessário desenvolver um mecanismo que reduzisse o número de regras das bases de regras e excluísse das mesmas regras não representativas, caso houvesse uma futura alteração no sistema inicial proposto, inserindo novos parâmetros para análise da QoS ou considerando novos conjuntos nebulosos. Para este fim, a metodologia de Algoritmos Genéticos foi utilizada.

Seleção de Regras Nebulosas por Algoritmos Genéticos

A combinação de Sistemas Nebulosos Baseados em Regras e Algoritmos Genéticos resultou em sistemas denominados Sistemas Genéticos Nebulosos Baseados em Regras, onde um algoritmo genético é empregado para ajustar ou otimizar a base de conhecimento de um SNBR. Em particular, um alto número de publicações explora o uso de algoritmos genéticos para o projeto de sistemas nebulosos durante a última década [Córdon et al. 2004].

O processo de seleção de regras tem por objetivo selecionar, dentre as regras previamente obtidas com o auxílio do especialista, um reduzido número de regras com alta habilidade para classificação, eliminando regras redundantes e desnecessárias. Na etapa de seleção o objetivo é extrair um conjunto mais compacto de regras que mantenham o mesmo desempenho pela eliminação de redundâncias e excluir regras que não representam uma situação do mundo real.

Em seguida são apresentados os componentes e parâmetros do Algoritmo Genético utilizado. Para tanto, primeiramente é feita uma representação cromossômica dos indivíduos da população. Cada indivíduo é uma possível solução do problema. O algoritmo inicia seu desenvolvimento criando aleatoriamente a primeira população. A partir daí, é calculada a adaptabilidade relativa de cada indivíduo ao sistema. Frente a esse cálculo, são selecionados (por sorteio, mas obedecendo a critérios probabilísticos) os elementos que pertencerão à geração seguinte. Para completar a população, efetua-se a reprodução dos pais selecionados, implementando-se, então, operadores genéticos, como Cruzamento e Mutação. Cada operador genético tem uma taxa de ocorrência apropriada, em metáfora à biologia. O processo se repete até que uma condição de parada especificada seja atendida. Essa condição pode se basear na estagnação da evolução da adaptabilidade média da população, na repetição sucessiva do melhor indivíduo, ou mesmo num limite de gerações pré-fixado.

Codificação

Nesta etapa os cromossomos $C_i=(c_{i1},c_{i2},\dots,c_{im})$ são codificados com seqüências de dígitos binários e possuem tamanho fixo 'm', sendo que 'm' é o número de regras obtidas. Cada gene c_{ij} (um dígito binário) é associado a uma regra. Se $c_{ij}=1$, então a regra

correspondente a este gene fará parte da base de regras selecionada, caso contrário, a regra não fará parte.

Nos experimentos realizados neste trabalho, foi atribuído a 'm' o valor 81, que representa o número de regras de cada uma das bases de regras.

Tamanho da população

O tamanho da população afeta o desempenho global e a eficiência dos algoritmos genéticos, já que uma população pequena fornece uma cobertura pequena do espaço de busca do problema. No entanto, para se trabalhar com grandes populações, são necessários maiores recursos computacionais e tempos muito longos para resolução de problemas simples. Assim, a escolha correta do tamanho de população, que produz respostas corretas e rápidas, depende do problema e da sua formulação. Geralmente, usa-se um tamanho de população proporcional ao tamanho do cromossomo, isto é, quanto maior for o cromossomo maior deverá ser o tamanho da população, para manter uma diversidade razoável.

O tamanho da população utilizada foi de 200 indivíduos, pois possibilitou uma melhor cobertura do espaço de busca e se mostrou eficiente em nossa experiência.

População inicial

A população inicial é gerada introduzindo um cromossomo que represente todas as regras previamente obtidas, ou seja, todos os genes do cromossomo recebem valor 1. Os cromossomos restantes foram gerados aleatoriamente.

Número de Gerações

O número de gerações define a quantidade de populações criadas até a resposta final. Com um valor baixo, pode-se encontrar rapidamente uma solução ótima local, porém ainda longe da solução global. Com um valor alto, tem-se garantia de encontrar a solução ótima, porém o tempo para se chegar à resposta pode ser longo, além de no caso de problemas simples, criarem-se novas gerações desnecessariamente, mesmo após a solução ótima ter sido encontrada. O valor ideal para este parâmetro depende da formulação do problema.

Foi estabelecido um número de 100 gerações como critério de parada, sendo suficiente para o alcance de uma base de regras otimizada, pois considerando um número de gerações superior ao utilizado, não havia variação na escolha do melhor indivíduo a partir da centésima geração.

Função de aptidão

A função de aptidão avalia cada cromossomo, ou seja, cada sub-conjunto de regras candidatas com base em dois critérios: número de exemplos corretamente classificados por ele e o número de regras desse sub-conjunto, ou seja, o número de genes com valor 1 nesse cromossomo. Para se obter os exemplos a serem classificados, necessários para o cálculo do valor de aptidão, foi necessária a construção de dois conjuntos de dados, um para cada aplicação envolvida nos experimentos, gerados a partir da base de regras inicial, sendo que cada conjunto consiste em 255 exemplos representativos de situações reais.

Terá o maior valor de aptidão o indivíduo que apresentar o maior valor de classificação correta de exemplos, com o menor número de regras. Desta forma, a função de aptidão pode ser expressa pela equação abaixo, definida por [Castro et al. 2003]:

$$\text{valor_aptidão}(C_i) = \text{NEC}(C_i) - \text{QTDR}(C_i)$$

onde $\text{NEC}(C_i)$ representa o número de exemplos classificados corretamente pelo sub-conjunto de regras representado no cromossomo i e $\text{QTDR}(C_i)$ indica a quantidade de regras existentes no sub-conjunto representado pelo cromossomo i .

Estes conjuntos de dados, criados poderão ser disponibilizados para o uso em treinamento ou teste de outros AG's com o mesmo fim, assim como existem conjuntos de dados tais como o Iris Plants (dados de classes de plantas do tipo Íris) e o Wine (dados resultantes de uma análise química de vinhos de uma mesma região da Itália), disponíveis em [Blake e Merz 1998].

Desta forma, o algoritmo genético pode ser aplicado aos conjuntos de dados para selecionar as regras previamente obtidas, excluindo as regras desnecessárias.

População Intermediária

O método de seleção utilizado neste trabalho para selecionar a população intermediária foi o método da Roleta e os operadores genéticos aplicados foram os de Cruzamento, Mutação e Elitismo.

A taxa de cruzamento define a probabilidade de ocorrer um cruzamento. Quanto maior for essa taxa, mais rapidamente novas estruturas serão introduzidas na população. Porém, estruturas com boas aptidões poderão ser substituídas mais rapidamente, perdendo-se boas oportunidades para achar uma solução ótima. Com um valor baixo, o algoritmo pode tornar-se muito lento, demorando a encontrar a solução ideal. A maior parte da literatura recomenda uma taxa de cruzamento entre 50% e 95% [Mitchell 1996]. A taxa de cruzamento utilizada foi de 70%, valor usual que se mostrou adequado para nossos experimentos.

A taxa de mutação define a probabilidade de ocorrência de mutação. Uma taxa de mutação baixa diminui a inclusão de indivíduos novos na população, provocando o encontro de resultados baseados em máximos locais, pois restringe o espaço de busca. Uma taxa muito alta transforma a busca do algoritmo genético em busca essencialmente aleatória. A maior parte da literatura recomenda uma taxa de mutação entre 0,1% e 5% [Mitchell 1996]. O valor da taxa de mutação utilizada foi de 1%, estando este dentro da faixa recomendada.

Utilizou-se também o método de seleção Elitista, com uma porcentagem de elitismo de 10%. A aplicação do elitismo previne que os melhores indivíduos de cada geração sejam destruídos pela aplicação dos operadores de cruzamento ou mutação, possibilitando que estes permaneçam na população e contribuam para uma convergência mais rápida do algoritmo.

Após a aplicação do Algoritmo Genético, a base de regras construída para a aplicação de Videoconferência foi reduzida em aproximadamente 40%, passando de 81 regras para 49. A base de regras da aplicação de Vídeo sob Demanda foi simplificada em aproximadamente 46%, passando a ter, ao invés de 81 regras, apenas 37.

6.2.3 Módulo de Desfuzificação

Como mencionado, foi utilizado neste trabalho o método de desfuzificação Centro de Massa, definido no item 4.2.4. No módulo de desfuzificação, as pertinências são calculadas nos parâmetros dos conjuntos nebulosos definidos para as variáveis de saída, que neste trabalho trata-se de uma única que representa a QoS. Para ambas as aplicações consideradas nos experimentos, os parâmetros dos conjuntos nebulosos de saída foram os mesmos e são apresentados na Tabela 6-15. Esses parâmetros foram definidos com base nas funções trapezoidais, mostradas na Figura 4-2, da mesma forma feita no módulo de fuzificação, analisando a QoS como Intolerável no intervalo (0 a 40), Aceitável (40 a 80) e Ótimo (>80).

Tabela 6-15: Parâmetros das funções de pertinência para a variável de saída QoS

	Ponto a	Ponto m	Ponto n	Ponto b
Conjunto Ótimo	60	80	100	100
Conjunto Aceitável	20	40	60	80
Conjunto Intolerável	0	0	20	40

Para um melhor entendimento, considere a Figura 6-4 a seguir.

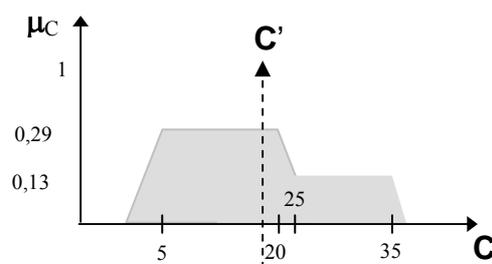


Figura 6-4: Exemplo do cálculo de desfuzificação

Considere também a expressão analítica de implementação do método do Centro de Massa definida anteriormente e ilustrada abaixo.

$$Z = \frac{\sum_{i=0}^m \mu_C(Z_i) * Z_i}{\sum_{i=0}^m \mu_C(Z_i)}$$

Na Figura 6-4, os valores de $\mu_C(Z_i)$, ou seja, os mínimos valores de pertinência para cada regra disparada, são 0,29 e 0,13 e os Z_i são os valores correspondentes no eixo definido por C, que se encontra dividido em intervalos de 5 em 5. Essa divisão de Z_i em intervalos menores ou maiores é dependente do grau de especificidade da resposta que se deseja obter. A divisão em intervalos de 5 foi suficiente para se obter a resposta esperada do sistema desenvolvido neste trabalho.

No dividendo, cada valor de $\mu_C(Z_i)$ deve ser multiplicado pelo respectivo Z_i (variante em intervalos de 5) e somado com os demais. No divisor, os valores de $\mu_C(Z_i)$ devem ser somados quantos forem os Z_i .

Substituindo estes valores na expressão analítica temos:

$$C' = \frac{0,29 * (5 + 10 + 15 + 20) + 0,13 * (25 + 30 + 35)}{(4 * 0,29 + 3 * 0,13)}$$

Assim, o valor desfuzificado **C'** é aproximadamente 16,9.

Este valor desfuzificado representa o “melhor compromisso” entre os valores de $\mu_C(Z_i)$ obtidos [Constantin 1995]. Para fins de explicação, considere a seguinte analogia.

Imagine que você está em seu apartamento às onze horas da noite e gostaria de ouvir uma música. Isso somente lhe seria prazeroso se a música estivesse em um volume elevado, mas seus vizinhos se aborreceriam e você não quer isso. Agora, quando você for configurar o volume do seu aparelho de som, você terá que combinar estes dois objetivos, nebulosos e conflitantes, em um valor real (desfuzificado) que represente um valor de volume para seu aparelho de som, de modo que haja um “ponto de equilíbrio” entre os dois objetivos. Deste modo, o valor desfuzificado é um balanceamento dos objetivos pretendidos, um meio termo.

6.3 Considerações Finais

Neste capítulo foram apresentadas todas as etapas de desenvolvimento do método de quantificação de QoS proposto, envolvendo a descrição das técnicas utilizadas e a definição dos parâmetros com os quais estas técnicas foram executadas para as duas aplicações em estudo de caso.

Os resultados obtidos com a aplicação do sistema nebuloso na avaliação de qualidade de serviço estavam dentro do domínio esperado e próximos dos obtidos pelas métricas numéricas convencionais utilizadas como comparação. Tais resultados e comparações serão apresentados no capítulo seguinte. A base de regras, definida pelo especialista e ajustada por observação experimental, conseguiu representar de forma satisfatória as situações passíveis de ocorrência no ambiente em análise. Sua otimização com o algoritmo genético sofreu alguns problemas em função de se ter inicialmente um conjunto de dados muito pequeno para que o algoritmo o classificasse no processo de cálculo do valor de aptidão, mencionado no item 6.2.2. Isso provocava a eliminação de regras importantes para a classificação da QoS, pois haviam poucos exemplos (amostras) representando situações de ocorrência frequente no ambiente em questão. O problema foi solucionado aumentando o número de amostras significativas no conjunto de dados, de 110 para 255. Com o uso do algoritmo genético o sistema nebuloso manteve seu poder de avaliação e caso fosse considerada uma base de regras de tamanho bastante superior a 81 regras, o sistema nebuloso conseguiria uma resposta de avaliação de QoS mais rapidamente, pois o espaço de busca do problema ficaria reduzido. Além disso, os valores de otimização das bases de regras em torno de 45% está próximo dos valores obtidos com o emprego do algoritmo genético para fins de otimização em outras aplicações, conforme apresentado no item 5.3.

O Apêndice B traz a modelagem do sistema desenvolvido, incluindo diagramação de classes e casos de uso, com a respectiva seqüência de processamento das informações no sistema.

No próximo capítulo serão apresentados os experimentos realizados, os resultados obtidos, e os comparativos com intuito de validação e demonstração de melhoria na forma da resposta final obtida.

Capítulo 7

Experimentos e Resultados

Com a finalidade de testar e validar o método de avaliação de QoS desenvolvido, foram realizados dois estudos de caso utilizando aplicações com características diferentes: uma de Videoconferência e outra de Vídeo sob Demanda .

7.1 Estudo de Caso 1: Videoconferência

Videoconferência é a transmissão de imagem e voz, em tempo real, entre dois ou mais locais separados fisicamente, utilizando câmeras de vídeo, microfones, monitores de vídeo e caixas de som [CookBook 2000]. Nesta aplicação, três clientes estão conectados a um ISP (Internet Service Provider) e todos eles concordaram com o mesmo SLA (Service Level Agreements) com o ISP, conforme ilustração na Figura 7-1.

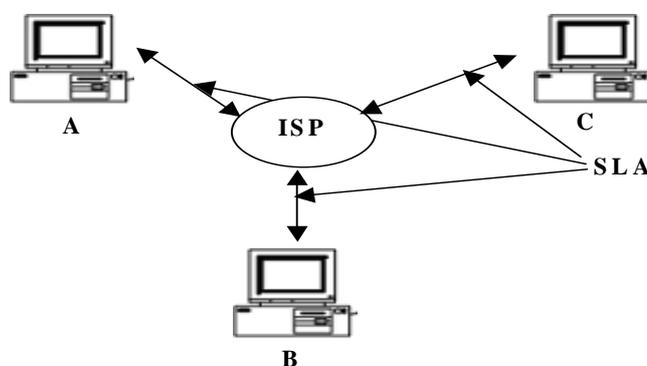


Figura 7-1: Aplicação de Videoconferência

As medidas dos parâmetros de QoS foram coletadas utilizando um Monitor de Qualidade de Serviço, denominado MQM – Multicast Quality Monitor, desenvolvido por [Dressler 2003c], que introduz diferentes formas de medir parâmetros em uma rede multicast. A Tabela 7-1 mostra os resultados das medidas para a aplicação de videoconferência.

Tabela 7-1: Videoconferência - Resultados de medidas dos parâmetros

Parâmetros	Conexão A-B	Conexão A-C
Vazão	27Mbps	1.2Mbps
Delay	20ms	150ms
Jitter	10ms	80ms
Perda de Pacotes	5%	0,1%

Considere a Tabela 6-6 para compreensão da análise a seguir.

Na conexão A-B, os valores para vazão, delay e jitter estão dentro do domínio considerado ótimo. Em contraste, o parâmetro perda de pacotes, considerado de maior relevância em aplicações desse tipo, está com um valor considerado intolerável.

No caso da conexão A-C, todos os parâmetros estão com seus valores incluídos no domínio ótimo considerado para o sistema, estando apenas o delay próximo do limiar entre ótimo e aceitável.

7.2 Estudo de Caso 2: Vídeo sob Demanda

Um serviço de vídeo sob demanda envolve o armazenamento de dados multimídia tais como textos, imagens, áudio e vídeo em servidores. O projeto de tais servidores difere do projeto de servidores convencionais fundamentalmente por requerer a transmissão dos quadros de um filme em tempo real. Ademais, um servidor de vídeo requer alta capacidade de armazenamento e alta largura de banda para a leitura em disco e transferência pela rede dos dados multimídia [Campos e Almeida 2000].

Na aplicação ilustrada pela Figura 7-2, existe um servidor de vídeo e dois clientes, C1 e C2, conectados a ele, solicitando transmissões remotas de vídeo.

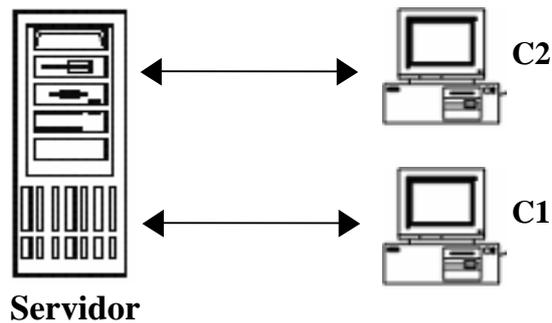


Figura 7-2: Aplicação de Vídeo sob Demanda

A Tabela 7-2 mostra os resultados das medidas para a aplicação de vídeo sob demanda.

Tabela 7-2: Vídeo sob Demanda - Resultados de medidas dos parâmetros

Parâmetros	Conexão C1-Servidor	Conexão C2-Servidor
Vazão	13Mbps	900Kbps
Delay	5s	7s
Jitter	0,6s	4s
Perda de Pacotes	2,5%	5%

Considerando novamente a Tabela 6-6, percebe-se que na conexão C1-Servidor os parâmetros vazão e jitter são considerados ótimos e os parâmetros delay e perda de pacotes, embora estejam no limite entre aceitável e intolerável, são ainda classificados como aceitáveis.

Em se tratando da conexão C2-Servidor, a vazão é ótima, o jitter aceitável e os demais parâmetros, delay e perda de pacotes são classificados como intoleráveis.

Com base nestas medições e observações, uma análise da qualidade de serviço pode ser efetuada pelo sistema nebuloso com o objetivo de avaliá-la. Essa análise será apresentada nos itens seguintes.

7.3 Resultados dos Estudos de Caso

Considerando os intervalos de valores que denominam os níveis de qualidade de serviço (Tabela 6-6), e com os valores discretos dos parâmetros de QoS coletados (Tabela 7-1 e Tabela 7-2), é possível estabelecer uma relação entre eles, utilizando métricas numéricas

convencionais ou as metodologias de Inteligência Computacional apresentadas anteriormente, com o intuito de obter um único valor resultante que expresse o nível de QoS da conexão como um todo. As métricas diferenciam-se desde o método utilizado no cálculo, até o formato do resultado obtido.

Com o Sistema Nebuloso, dois resultados para cada aplicação foram obtidos (Tabela 7-3 e Tabela 7-4): primeiramente utilizando a base de regras inicial e posteriormente usando a base de regras otimizada com a aplicação do algoritmo genético.

Tabela 7-3. Videoconferência - Resultados de avaliação de QoS utilizando um SGNBR

	Conexão A-B		Conexão A-C	
	Valor Desfuzificado	Classificação	Valor Desfuzificado	Classificação
Base de Regras Inicial	10	100% Intolerável	90	60% Ótima
Base de Regras Otimizada	10	100% Intolerável	70	40% Aceitável

Na conexão A-B, os resultados foram os mesmos, tanto para a base de regras inicial quanto para a base otimizada, declarando a QoS totalmente intolerável para a aplicação de videoconferência. Isso já era esperado, uma vez que a perda de pacotes, o parâmetro mais importante neste tipo de aplicação, era extremamente intolerável (na faixa de 5%). Além disso, o valor desfuzificado ‘10’ está totalmente inserido no conjunto nebuloso de saída Intolerável. Os parâmetros para os conjuntos nebulosos de saída são definidos pela Tabela 6-15.

Na conexão A-C, os resultados com a base de regras inicial e com a base de regras otimizada se complementam. Observando novamente os comentários feitos no item 7.1, que traz uma análise dos parâmetros mensurados para a aplicação de videoconferência, é possível verificar que apenas o parâmetro delay está com um valor entre o ótimo e o aceitável. Os demais parâmetros são classificados como ótimos. Assim, poderia se dizer que a QoS é mais ótima do que aceitável, o que ambas as classificações demonstram. Quando se diz que a QoS é 60% ótima, com um valor desfuzificado de ‘90’, subentende-se que os outros 40% restantes, estão dentro da faixa considerada aceitável, uma vez que o valor desfuzificado ‘70’ está no limiar entre os conjuntos nebulosos de saída Aceitável e Ótimo, não sofrendo nenhuma influência do conjunto nebuloso de saída Intolerável.

Tabela 7-4: Vídeo sob Demanda - Resultados de avaliação de QoS utilizando um SGNBR

	Conexão C1-Servidor		Conexão C2-Servidor	
	Valor Desfuzificado	Classificação	Valor Desfuzificado	Classificação
Base de Regras Inicial	$\cong 24.6$	50% Aceitável	10	90% Intolerável
Base de Regras Otimizada	$\cong 41.57$	50% Aceitável	10	90% Intolerável

Analisando a aplicação de vídeo sob demanda, conexão C1-Servidor, percebe-se que também há uma igualdade na classificação da qualidade de serviço com ambas as bases de regras. O valor desfuzificado de aproximadamente 41.57 representa melhor a situação da QoS, pois os parâmetros mais importantes para esse tipo de aplicação, delay e perda de pacotes, estão com seus valores de medição na faixa aceitável mas com grande influência do conjunto nebuloso de saída Intolerável. Deste modo, pode-se dizer que a QoS é aceitável, mas com inclinação para intolerável.

Na conexão C2-Servidor, a QoS é considerada totalmente intolerável para as duas bases de regras. Está classificação era esperada, já que o delay e a perda de pacotes, os parâmetros mais importantes neste tipo de aplicação, assumiam valores extremamente intoleráveis. Além disso, o valor desfuzificado '10' está totalmente inserido no conjunto nebuloso de saída Intolerável.

7.4 Análise da Variação da QoS

Outras análises foram feitas em alguns intervalos de tempo para ambas as aplicações, considerando todas as combinações dos parâmetros de qualidade de serviço, iniciando da situação onde todos estavam em baixa, até a situação onde todos estavam em alta. Agrupando os quatro parâmetros nas duas situações consideradas (em alta ou em baixa) foram geradas 16 combinações, onde apenas 9 foram analisadas por serem situações possíveis de ocorrência no ambiente em questão.

As figuras a seguir ilustram graficamente as oscilações da qualidade de serviço, em função de seus parâmetros, para as aplicações de videoconferência e vídeo sob demanda, uma

vez que ambas apresentam praticamente o mesmo comportamento quando o valor do parâmetro perda de pacotes, o mais significativo para estas aplicações, está aumentando ou diminuindo.

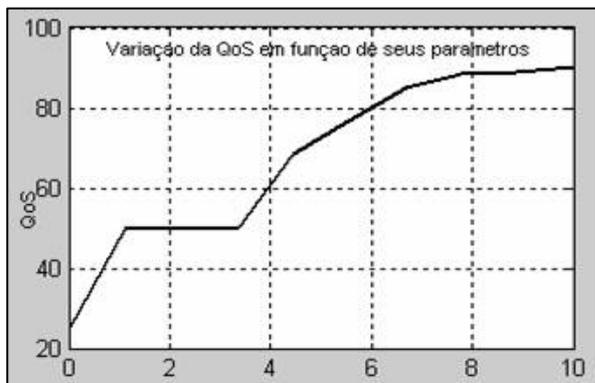


Figura 7-3: Variação da QoS – Análise 1

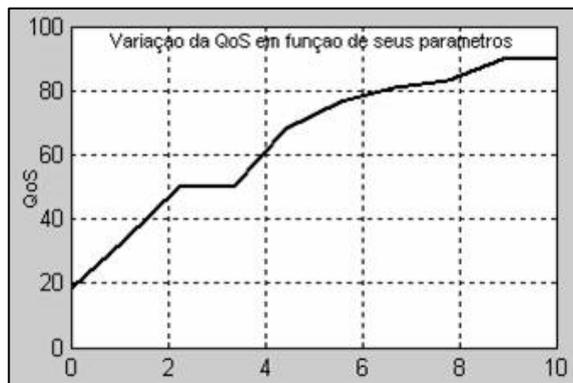


Figura 7-4: Variação da QoS – Análise 2

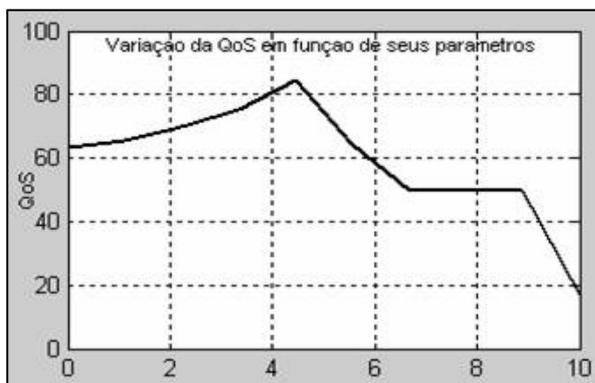


Figura 7-5: Variação da QoS – Análise 3

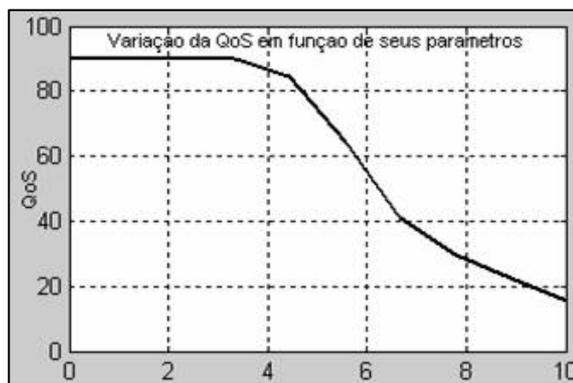


Figura 7-6: Variação da QoS – Análise 4



Figura 7-7: Variação da QoS – Análise 5

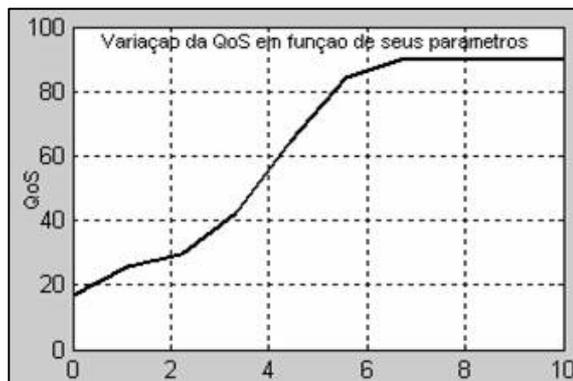


Figura 7-8: Variação da QoS – Análise 6

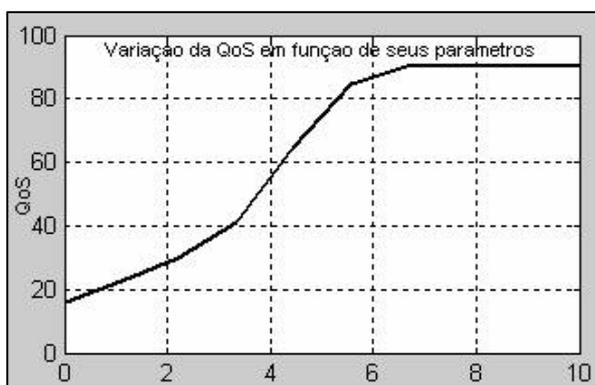


Figura 7-9: Variação da QoS – Análise 7

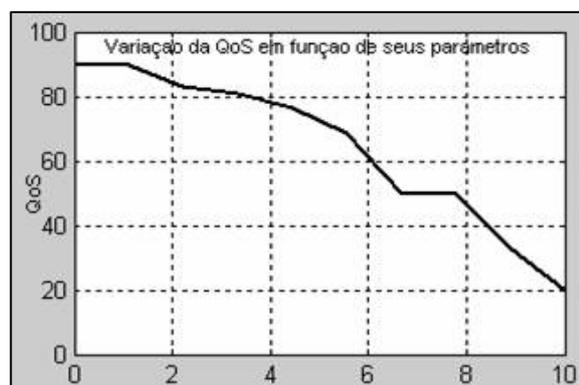


Figura 7-10: Variação da QoS – Análise 8

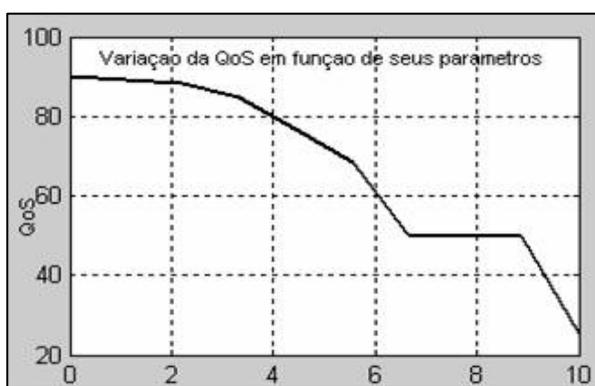


Figura 7-11: Variação da QoS – Análise 9

Para facilitar a compreensão, a análise destas medidas será feita com base no valor desfuzificado obtido para a qualidade de serviço, uma vez que este valor reflete a condição de cada parâmetro envolvido e sua importância no processo. O eixo horizontal dos gráficos ilustrados pelas figuras acima rotula as dez avaliações feitas para a qualidade de serviço mediante a submissão dos parâmetros ao sistema genético nebuloso. Após esta submissão, cada avaliação, de 1 até 10, gera um valor de QoS expresso no eixo vertical, sendo analisado utilizando os parâmetros dos conjuntos nebulosos de saída que representam a situação da qualidade de serviço, definidos pela Tabela 6-15. As intersecções das avaliações rotuladas com seus correspondentes valores de qualidade de serviço produzem as curvas apresentadas nas 9 análises da variação da QoS.

Na Figura 7-3, todos os parâmetros iniciam com um valor alto e vão diminuindo no decorrer do tempo. A vazão vai sendo reduzida, mas ao mesmo tempo, delay, jitter e perda de pacotes também vão diminuindo, o que faz a avaliação da QoS sair de intolerável e chegar próxima de 100% ótima.

No caso da Figura 7-4, vazão, delay e perda de pacotes estão diminuindo, ao passo que o jitter aumenta. A avaliação da QoS também passa de intolerável para ótima, mas apenas com um limite de aproximadamente 85% ótima.

A Figura 7-5 expressa uma QoS variante entre ótima e aceitável durante algum período de tempo, caindo posteriormente para intolerável. Nesta análise, a vazão e o delay estavam em queda e o jitter e a perda de pacotes em alta. Assim, à medida que o delay se tornava menor, mas por algum motivo sua variação e a perda de pacotes sofriam um incremento, a QoS começava a oscilar, até chegar a um valor considerado intolerável em função da quantidade de pacotes perdidos ser muito alta. A diminuição da vazão também influenciava negativamente o processo.

A situação mostrada pela Figura 7-6, reflete uma avaliação de qualidade de serviço que sai de uma análise ótima para intolerável, pois apesar do jitter estar decaindo no decorrer do tempo, a vazão está diminuindo e os parâmetros delay e perda de pacotes estão em alta, influenciando em muito a avaliação da QoS como negativa.

Na Figura 7-7, temos dois extremos para a qualidade de serviço, onde a vazão vai sendo reduzida e os demais parâmetros incrementados. No início da análise, com uma alta vazão e baixos delay, jitter e perda de pacotes, a QoS é extremamente ótima. Conforme esta situação vai se invertendo, a situação da QoS vai se tornando negativa, até atingir o ponto de extremo intolerável.

Uma situação inversa da apresentada na Figura 7-7 é mostrada na Figura 7-8. Enquanto a vazão vai aumentando e os demais parâmetros vão diminuindo, a QoS passa de um extremo intolerável pra um extremo ótimo.

Para a Figura 7-9, tem-se os parâmetros vazão e jitter aumentando e delay e perda de pacotes diminuindo. Com o aumento da vazão e diminuição do delay e da perda de pacotes, a QoS vai se tornando ótima, mas em função de um alto jitter, fica limitada a aproximadamente 85% ótima.

A Figura 7-10 é representativa de uma qualidade de serviço em baixa, onde os parâmetros delay e perda de pacotes estão sendo incrementados. A vazão está em alta e o jitter diminuindo, mas a forte influencia da elevação na perda de pacotes vai transformando lentamente a QoS de ótima para intolerável.

A avaliação da QoS na Figura 7-11 é contrária a feita na Figura 7-3. Nesta avaliação, todos os parâmetros em análise estão sofrendo aumento em seus valores. A vazão está em

alta, mas os demais parâmetros, também em elevação, atingem a qualidade de serviço de forma extremamente negativa, fazendo-a assumir valores bastante intoleráveis.

O objetivo destas análises foi o de complementar a avaliação de QoS, cobrindo todas as situações passíveis de ocorrência no ambiente em questão, também como forma de comprovar a validade e aplicabilidade do método desenvolvido.

7.5 Validação do Método Proposto

Com a utilização do NIST Net, foi possível visualizar a qualidade da aplicação em exibição, tendo também como parâmetro de comparação a avaliação feita pelo usuário. De acordo com a qualidade visualizada, os parâmetros foram usados para ‘alimentar’ o Sistema Nebuloso, devendo este gerar uma resposta coerente ao que foi observado. Como exemplo, considere a seguinte situação: são definidos valores para os parâmetros perda de pacotes, delay, jitter e vazão. Emulador de rede e aplicação são colocados em execução e observa-se uma péssima qualidade no que está sendo exibido. Assim, o Sistema Nebuloso receberá os mesmos valores dos parâmetros e deverá retornar uma resposta confirmando o intolerável nível de QoS.

Esses mesmos valores foram também inseridos nas métricas numéricas que serão apresentadas a seguir, objetivando ajustes e mais comparações.

7.5.1 Métrica Numérica proposta por Dressler

Esta métrica, proposta por [Dressler 2003c], define um método de cálculo que proporciona comparar parâmetros com características em comum ou não. Tipicamente, a qualidade fim-a-fim de uma conexão depende de todas as métricas dos parâmetros envolvidos. Assim, a qualidade de uma conexão ‘c’ pode ser escrita como um vetor, conforme mostrado abaixo:

$$c = \begin{bmatrix} m_0 \\ m_1 \\ \dots \\ m_n \end{bmatrix}$$

Onde:

m_i = métricas dos diferentes parâmetros;

$i = 0, 1, \dots, n$.

Uma representação numérica da qualidade da conexão como um todo, é descrita como um produto das métricas individuais de ‘c’, como mostrado na fórmula abaixo, cabendo ao usuário interpretar este valor para poder tomar suas decisões.

$$c = m_0 * m_1 * \dots * m_n$$

Desta forma, se o valor de ‘c’ é igual a 1, a conectividade é provida com um serviço perfeito de QoS; os valores entre 0 e 1 descrevem uma qualidade média, com a QoS podendo ser garantida apenas parcialmente; e o valor 0, significa que o serviço não está disponível.

Assim, a análise é feita para ambas as aplicações como mostram as tabelas Tabela 7-5 e Tabela 7-6:

Tabela 7-5. Videoconferência - Análise de QoS segundo Dressler

Conexão A-B	Conexão A-C
$m_{Vazão} = 1$ (valor medido e melhor que o ótimo);	$m_{Vazão} = 1$ (valor medido e melhor que o ótimo);
$m_{Delay} = 1$ (valor medido e melhor que o ótimo);	$m_{Delay} = 1$ (valor medido e melhor que o ótimo);
$m_{Jitter} = 1$ (valor medido e melhor que o ótimo);	$m_{Jitter} = 1$ (valor medido e melhor que o ótimo);
$m_{Perda\ de\ Pacotes} = 0$ (valor medido e pior que o máximo tolerável);	$m_{Perda\ de\ Pacotes} = 1$ (valor medido e melhor que o ótimo);
$c_{AB} = 1*1*1*0 = 0$; o que significa que a conexão de A para B <u>não pode</u> ser usada.	$c_{AC} = 1*1*1*1 = 1$; o que significa que a conexão de A para B <u>pode ser utilizada</u> para o serviço requisitado com uma QoS fim-a-fim de 100%.

Tabela 7-6: Vídeo sob Demanda - Análise de QoS segundo Dressler

Conexão C1-Servidor	Conexão C2-Servidor
<p>$m_{Vazão} = 1$ (valor medido é melhor que o ótimo);</p> <p>$m_{Delay} = 0.5$ (valor medido está entre o máximo tolerável e o ótimo);</p> <p>$m_{Jitter} = 1$ (valor medido é melhor que o ótimo);</p> <p>$m_{Perda\ de\ Pacotes} = 0.5$ (valor medido está entre o máximo tolerável e o ótimo);</p> <p>$c_{C1-Servidor} = 1*0.5*1*0.5 = 0.25$; o que significa que a conexão de C1 para o Servidor <u>não deveria</u> ser usada.</p>	<p>$m_{Vazão} = 1$ (valor medido é melhor que o ótimo);</p> <p>$m_{Delay} = 0$ (valor medido é pior que o máximo tolerável);</p> <p>$m_{Jitter} = 0.5$ (valor medido está entre o máximo tolerável e o ótimo);</p> <p>$m_{Perda\ de\ Pacotes} = 0$ (valor medido é pior que o máximo tolerável);</p> <p>$c_{C2-Servidor} = 1*0*0.5*0 = 0$; o que significa que a conexão de C2 para o Servidor <u>não pode</u> ser usada.</p>

7.5.2 Métrica Numérica proposta por Lunardi e Dotti

O algoritmo utilizado nesta métrica é baseado na informação de *pacotes perdidos* encontrada em *receiver blocks* dos pacotes de controle do protocolo RTCP (Real Time Control Protocol) e foi definido por [Busse et al. 1995] apud [Lunardi e Dotti 2001]. O RTCP é usado para troca de informações de controle entre fonte e destino.

Somente a taxa de perda de pacotes é utilizada para identificar congestionamento na rede. Uma fórmula para calcular a taxa de perda, suavizar estatísticas e evitar oscilações de QoS é apresentada abaixo:

$$\lambda^{atual} = (1-\alpha)\lambda^{anterior} + \alpha b$$

Onde:

λ = taxa de perda de pacotes/estado da rede;

b = novo valor referente à perda de pacotes;

α = fator de suavização, que determina o peso que é dado à taxa de perdas anterior, sendo $0 \leq \alpha \leq 1$

A rede é avaliada constantemente pelo algoritmo que verifica seu comportamento de acordo com a informação de pacotes perdidos, obtida através do *sequence number* contido no cabeçalho dos pacotes RTCP no receptor. O estado anterior da rede sempre é levado em consideração a cada novo cálculo da taxa de perdas de modo a evitar tomadas de decisões de adaptação não necessárias em períodos curtos de pico da rede. Também, procurou-se descobrir um valor adequado para α . Segundo testes, o valor ótimo que resultou em uma taxa de perdas próxima da realidade foi de 0,5.

O estado da rede, baseado em perdas, é determinado como Não Carregado (até 2%), Carregado (de 2% a 4%) e Congestionado (acima de 4%).

Avaliando as medidas coletadas das duas aplicações em estudo, observou-se que na conexão A-C, a rede não está carregada e a conexão pode ser provida com uma garantia de QoS de 100%. Na conexão C1-Servidor, a rede está carregada, sendo a QoS fornecida com aproximadamente 50% de garantia. Nas conexões A-B e C2-Servidor, existe congestionamento e nenhuma qualidade pode ser oferecida.

Assim, podemos sintetizar, nas tabelas Tabela 7-7 e Tabela 7-8, os resultados obtidos através das métricas numéricas apresentadas nos itens 7.4.1 e 7.4.2:

Tabela 7-7: Videoconferência - Resultados de avaliação de QoS utilizando métricas numéricas

	Conexão A-B	Conexão A-C
	Classificação	Classificação
DRESSLER	0	1
LUNARDI e DOTTI	0% (Congestionado)	100% (Não Carregado)

Tabela 7-8: Vídeo sob Demanda - Resultados de avaliação de QoS utilizando métricas numéricas

	Conexão C1-Servidor	Conexão C2-Servidor
	Classificação	Classificação
DRESSLER	0.25	0
LUNARDI e DOTTI	50% (Carregado)	0 (Congestionado)

As métricas apresentadas realizam a avaliação da qualidade de serviço em um ambiente multimídia, assim como o sistema genético nebuloso desenvolvido neste trabalho. Essas métricas utilizam-se da matemática convencional para obtenção de seus resultados, o que as difere do método desenvolvido, o qual extrai suas conclusões baseado em metodologias de inteligência computacional.

Ambas as metodologias, convencional e inteligente, obtiveram a mesma classificação de qualidade de serviço para as aplicações em estudo de caso, como mostram a Tabela 7-3, em comparação com a Tabela 7-7 e a Tabela 7-4 em relação à Tabela 7-8, servindo as métricas numéricas como forma de comparar e validar resultados do sistema nebuloso.

7.6 Análise dos Resultados

A métrica numérica desenvolvida por [Dressler 2003b] apresenta um valor numérico final da QoS, resultante de uma intersecção dos valores individuais dos parâmetros. Este valor pode não significar nada para um determinado tipo de usuário, devendo ser submetido a um interpretador que defina a QoS de uma forma compreensível a qualquer nível de usuário. Além disso, se apenas um único parâmetro estiver com um valor considerado “ruim”, ele simplesmente exclui a possibilidade de transmissão por aquela conexão, mesmo que os demais parâmetros estejam em uma faixa aceitável. Isso ocorre porque não foi definida nenhuma forma de ponderar os valores medidos dos parâmetros e assim, todos possuem o mesmo grau de importância, o que não é correto. Conforme observado, um dos fatores mais importantes, mas não o único ao qual deve ser dada atenção, é a perda de pacotes.

A métrica proposta por [Busse et al. 1995] apud [Lunardi e Dotti 2001], tenta classificar o valor de QoS obtido em um termo lingüístico, mas que é ainda bastante técnico e distante do entendimento de alguns tipos de usuário. Outra desvantagem é o fato de considerar apenas um único parâmetro de rede como forma de avaliar a QoS de toda uma conexão, desconsiderando outros parâmetros que, apesar de menos importantes, exercem influência na qualidade de uma conexão.

A observação dessas desvantagens tem levado um grande número de pesquisadores a procurar alternativas não analíticas de técnicas de modelagem e desenvolvimento de sistemas de controle, em função de que a exatidão de um sistema puramente numérico nem sempre classifica corretamente um problema [Sekercioglu et al. 2001]. Um Sistema Nebuloso é justificado pela observação de que os parâmetros de QoS tratam de domínios diferentes, e que as regiões e as variáveis dentro da aplicação são discretas e não-lineares. Em tal cenário, a lógica nebulosa permite a especificação de tal processo de tomada de decisão através de regras nebulosas.

Além disso, a Lógica Nebulosa e os Algoritmos Genéticos por serem tolerantes à imprecisão, incerteza e verdades parciais, são uma alternativa para a resolução de problemas que atualmente não possuem solução por meio dos métodos tradicionais ou que as soluções encontradas por estes não apresentam bons resultados, sejam nos valores ou na forma de apresentação dos mesmos, o que é o caso das métricas numéricas apresentadas neste trabalho.

O Sistema Nebuloso é mais apropriado quando se trata de um ambiente com não-linearidades, com o processo de mapeamento de entradas e saídas sendo expresso por um número de regras lingüísticas armazenadas em uma base de regras. As regras que controlam o processo de mapeamento são configuráveis e podem ser intuitivamente especificadas para satisfazer as necessidades da aplicação. O modelo fornece uma flexível especificação das necessidades de uma aplicação, o que é desejável, sem perda de desempenho, configurabilidade e generalidade [Nahrstedt e Li 2000].

7.7 Considerações Finais

Neste capítulo foram apresentados os resultados obtidos na avaliação de qualidade de serviço em duas transmissões envolvendo aplicações multimídia. Essa avaliação se deu com a utilização de um Sistema Genético Nebuloso Baseado em Regras, em comparativo com a análise feita através de métricas numéricas convencionais.

No próximo capítulo serão apontadas as conclusões, trabalhos relacionados e trabalhos futuros.

Capítulo 8

Conclusões, Trabalhos Relacionados e Trabalhos Futuros

8.1 Conclusões

Com o objetivo de tornar possível analisar a atual disponibilidade de qualidade de serviço (fim-a-fim) em uma rede, algumas métricas têm sido desenvolvidas. Esse trabalho apresentou um método para avaliação de QoS em uma conexão multimídia, utilizando um Sistema Genético Nebuloso Baseado em Regras, e apresentou resultados comparativos com métodos de modelos quantitativos convencionais em dois estudos de caso: uma aplicação multimídia de Videoconferência e um Vídeo sob Demanda. Os métodos diferenciam-se desde os parâmetros considerados no cálculo, até o formato do resultado obtido.

Comparando-se os resultados obtidos através da utilização dos métodos apresentados, juntamente com o Sistema Genético Nebuloso, é possível observar a semelhança nos valores finais resultantes, demonstrativos da QoS da conexão. Isto evidencia que um Sistema Genético Nebuloso Baseado em Regras pode ser utilizado para se analisar a Qualidade de Serviço em uma conexão de rede, podendo ser estendido a qualquer tipo de aplicação multimídia, ainda com a vantagem de obter uma resposta com um maior grau de especificidade, devolvendo um valor final de qualidade de serviço em uma linguagem acessível a qualquer tipo de usuário, interpretando o valor numérico de saída do sistema, associando a ele uma semântica. Se o usuário desejar ter acesso também ao valor numérico, este pode ser facilmente fornecido.

O Sistema Genético Nebuloso desenvolvido utiliza-se dos recursos de aprendizado dos algoritmos genéticos para adaptar o sistema nebuloso inicial. Os algoritmos genéticos devem a sua popularidade à possibilidade de percorrer espaços de busca irregulares e extensos. Assim, viabilizam o tratamento da questão de construção de bases de regras nebulosas como um problema de busca e otimização e, como os resultados dos testes mostraram, torna possível fazer uma redução significativa no número de regras, obtendo uma base de regras mais compacta, que apresenta o mesmo desempenho da base original, em termos de taxa de classificação correta.

A combinação genético-nebuloso tem grande aceitação na comunidade científica, pois adere ao princípio de balanceamento de vantagens da inteligência computacional [Rezende 2003], onde metodologias diferentes colaboram entre si potencializando a utilidade e aplicabilidade dos sistemas resultantes. Os sistemas genéticos nebulosos apresentam a transparência e facilidade de compreensão típica dos sistemas nebulosos lingüísticos ao mesmo tempo em que se beneficiam das possibilidades de aprendizado dos algoritmos genéticos.

8.2 Trabalhos Relacionados

O trabalho desenvolvido em [Ghinea e Angelides 2004], apresenta um sistema de comércio eletrônico relacionado a multimídia que tem como fator de análise de desempenho a qualidade perceptível ao usuário, levando em consideração a subjetividade inerente a essa análise. Este trabalho também considera a avaliação de QoS em aplicações multimídia por parte do usuário, participando este do processo de construção e validação da base de regras do sistema nebuloso e recebendo como resposta deste sistema uma análise da qualidade de serviço em uma linguagem de seu entendimento.

Em [Koliver et al. 2004] é desenvolvido um ambiente que promove adaptação de qualidade de serviço para sistemas multimídia utilizando um sistema de controle nebuloso. São considerados cinco conjuntos nebulosos (Excelente, Bom, Médio, Fraco e Insuficiente) e duas variáveis de entrada (Perda de pacotes e Qualidade de Visualização). O termo lingüístico de saída é associado a seu valor desfuzificado, que é utilizado para realimentar o sistema na tentativa de adaptação da qualidade de serviço. Este trabalho realiza os mesmos

procedimentos, mas não promove a adaptação de QoS, dando prioridade à análise de qualidade no nível do usuário e criando mecanismos de otimização do sistema nebuloso como forma de melhorar seu desempenho.

Em [Fernandez et al. 2003] foi implementado um controlador nebuloso para o gerenciamento de políticas de QoS, criando um mecanismo de provisionamento dinâmico, reconfigurando os nós de acordo com o tráfego entrante. O controlador nebuloso foi comparado com um controlador convencional, assim como neste trabalho e, da mesma forma, o controlador nebuloso obteve resultados mais satisfatórios.

O sistema genético nebuloso desenvolvido em [Mondragon 1999], é utilizado para controle de congestionamento em uma rede ATM. As variáveis de entrada são o tamanho da fila, a variação do tamanho da fila e a probabilidade de descarte de pacotes. O valor de saída desfuzificado é usado para tomada de decisão nos roteadores. Da mesma forma com efetuado neste trabalho, um algoritmo genético é utilizado para reduzir o número de regras da base de regras do sistema nebuloso, alcançando também benefícios sobre o sistema inicial.

8.3 Trabalhos Futuros

Uma possível continuação deste trabalho seria a de tornar dinâmico e automatizado o processo de tomada de decisão, ou seja, “o que fazer” mediante a obtenção do nível de QoS em uma conexão de rede. Assim, é proposto como trabalho futuro a construção de um framework para adaptação dinâmica de qualidade de serviço, o qual recebe o valor da qualidade disponível na rede em um determinado instante e, se esta não estiver de acordo com o especificado pelo usuário, o framework tenta adaptar automaticamente a QoS, de forma que uma queda de qualidade no serviço oferecido não venha a ser percebida pelo usuário. Assim, o usuário define a QoS desejada e o framework se encarrega em tentar mantê-la, eliminando a responsabilidade por parte do usuário sobre a atitude a ser realizada.

Outras questões se referem às metodologias de Inteligência Computacional utilizadas. Como foram realizados testes com dois diferentes métodos de desfuzificação, poderia ser feito o mesmo em outras etapas do trabalho, tanto no sistema nebuloso quanto no algoritmo genético, com o objetivo de inserir melhorias no sistema desenvolvido.

Na etapa de fuzificação, poderiam ser utilizados um número maior de conjuntos nebulosos, o que aumentaria o grau de especificidade do sistema nebuloso. Poderiam também ser analisados um número maior de parâmetros de QoS, envolvendo outros níveis, como por exemplo o Atendimento, um parâmetro de qualidade de serviço no nível de usuário.

Referências Bibliográficas

- [Almeida et al. 2004] Almeida, J.; Eager, D. L.; Vernon, M. K.; Wright, S. J. "*Minimizing Delivery Cost in Scalable Streaming Content Distribution Systems*". IEEE Transactions on Multimedia, Special Issue on Streaming Media, Abril, 2004.
- [Belchior 1995] Belchior, A. D. "*Um Modelo Fuzzy para Avaliação da Qualidade de Software*". RT ES 344/95, COPPE/UFRJ - Rio de Janeiro, 1995.
- [Bittencourt 1998] Bittencourt, G. Inteligência Artificial: Ferramentas e Teorias. Editora da UFSC, 1998.
- [Blake e Merz 1998] Blake, C. L.; Merz, C. J. "*UCI Repository of Machine Learning Databases*". Department of Information and Computer Science - University of California - Irvine, C.A., 1998. Disponível em: <http://www.ics.uci.edu/~mlearn/MLRepository.html>.
- [Bochmann 2002] Bochmann, G. V. "*Multimedia Applications Protocols*". Course Notes of Quality of Service Management for Multimedia Applications - University of Ottawa, 2002.
- [Bochmann et al. 2001] Bochmann, G. V.; Kerhervé, B.; Lutfiyya, H.; Salem, M.; Ye, H. "*Introducing QoS to Electronic Commerce Applications*". ISEC - International Symposium on Electronic Commerce, vol.1, pp.138-147, Hong Kong - China, Abril, 2001.
- [Bramlette 1991] Bramlette, M. F. "*Initialization, Mutation and Selection Methods in Genetic Algorithms for Functions Optimization*". Proceedings of the Fourth International Conference on Genetic Algorithms, In R. K. Belew and L. B. Booker, eds., Morgan Kaufmann, 1991.
- [Busse et al. 1995] Busse, I.; Deffner, B.; Schulzrinne, H. "*Dynamic QoS Control of Multimedia Applications based on RTP*". First International Workshop on High Speed Networks and Open Distributed Platforms, St. Petersburg, Russia, Junho, 1995.

- [Campos e Almeida 2000] Campos, S.; Almeida, J. "*Descrição de um Serviço de Vídeo sob Demanda*", 2000. Disponível em: <http://www.vod.dcc.ufmg.br/vod/docs/descricao/dccvod.html>
- [Canan e Raabe 2004] Canan, R.; Raabe, A. "*Um Ambiente para Transmissão de Vídeos Instrucionais sob Demanda*". CINTED - UFRGS - Revista de Novas Tecnologias na Educação, vol.2, no.1, Março, 2004.
- [Carson e Santay 2003] Carson, M.; Santay, D. "*NIST Net – A Linux-based Network Emulation Tool*". ACM SIGCOMM Computer Communications Review, vol.33, pp.111-126, Julho, 2003.
- [Castro et al. 2003] Castro, P. A. D.; Pires, M. G.; Camargo, H. A. "*Aprendizado e Seleção de Regras Fuzzy usando Algoritmos Genéticos*". VI Simpósio Brasileiro de Automação Inteligente, Bauru, São Paulo, Brasil, Setembro, 2003.
- [Constantin 1995] Constantin, V. A. *Fuzzy Logic and Applications Explained*. 1ª Edição. Editora Prentice Hall, 1995.
- [CookBook 2000] CookBook "*Video Conferencing Cookbook*", 2000. Disponível em: <http://www.vide.gatech.edu/cookbook2.0>.
- [Córdon et al. 2004] Córdon, O.; Gomide, F.; Herrera, F.; Hoffmann, F.; Magdalena, L. "*Ten Years of Genetic Fuzzy Systems: Current Framework and New Trends*". *Fuzzy Sets and Fuzzy Systems*, vol.141, pp.5-31, 2004.
- [Cotroneo et al. 2003] Cotroneo, D.; Flora, C.; Paolillo, G.; Russo, S. "*Modeling and Detecting Failures in Next-generation Distributed Multimedia Applications*". 22º Symposium on Reliable Distributed Systems, vol.13, pp.379-388, Florença-Itália, Outubro, 2003.
- [Cox 1998] Cox, E. *The Fuzzy Systems Handbook: A Practitioner's Guide to Building, Using, & Maintaining Fuzzy Systems*. 2ª Edição. Editora Academic Press, 1998.
- [Dejong 1975] Dejong, K. A. "*An Analysis of the Behavior of a Class of Genetic Adaptive Systems*". Ph.D. thesis, Department of Computer and Communication Sciences, University of Michigan, 1975.
- [Delgado 2002] Delgado, M. R. "*Projeto Automático de Sistemas Nebulosos: Uma Abordagem Co-Evolutiva*". Tese de Doutorado - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação, 2002.
- [Dobson 2004] Dobson, G. "*Quality of Service: a Service-Centric Perspective*". Computing Department, Lancaster University. <http://www.lancs.ac.uk/ug/dobsong/qos.htm>, 2004.
- [Dressler 2003a] Dressler, F. "*Considerations on Selection Criteria for Sources of Multimedia Traffic*". SCI - 7th World Multiconference on Systematics, Cybernetics and Informatics, vol.3, pp.34-38, Orlando, Flórida - Estados Unidos, Julho, 2003.

- [Dressler 2003b] Dressler, F. "A Metric for Numerical Evaluation of the QoS of an Internet Connection". ITC18 - 18th International Teletraffic Congress, vol.5b, pp.1221-1230, Berlin - Alemanha, Agosto, 2003.
- [Dressler 2003c] Dressler, F. "Monitoring of Multicast Networks for Time-Synchronous Communication". Ph.D. Thesis - Department of Computer Science - University of Erlangen-Nuremberg, 2003. Disponível em: <http://bsd.rze.uni-erlangen.de/~fd/publications/dissertation.pdf>.
- [Fernandez et al. 2003] Fernandez, M. P.; Pedroza, A. C.; Rezende, J. F. "Implementação de Políticas de Gerenciamento através de lógica Fuzzy visando melhoria da Qualidade de Serviço (QoS)". 21º Simpósio Brasileiro de Redes de Computadores, Natal - RN - Brasil, Maio, 2003.
- [Ghinea e Angelides 2004] Ghinea, G.; Angelides, M. "A User Perspective of Quality of Service in m-Commerce". Multimedia Tools and Applications - Kluwer Academic Publishers, vol.22, no.1, pp.187-206, 2004.
- [Gonçalves et al. 2001] Gonçalves, E. M.; Costa, A. L.; Bittencourt, G. "Otimização de controladores nebulosos e sistemas especialistas reativos utilizando algoritmos genéticos". III Encontro Nacional de Inteligência Artificial, Fortaleza - CE - Brasil, 2001.
- [Grefenstette 1986] Grefenstette, J. J. "Optimization of Control Parameters for Genetic Algorithms". IEEE Transaction on Systems, Man and Cybernetics, vol.16, no.1, pp.122-128, 1986.
- [Holland 1975] Holland, J. H. Adaptation in Natural and Artificial Systems. Editora Ann Arbor: The University of Michigan Press, 1975.
- [Huard e Lazar 1997] Huard, J.; Lazar, A. "On End-to-End QoS Mapping." IFIP - Fifth International Workshop on Quality of Service (IWQoS'97), vol.1, pp.303-314, Nova York, Maio, 1997.
- [ITU 2001] ITU "ITU-T Study Group 12 - Workshop on QoS and user-perceived transmission quality in evolving networks". ITU Telecommunication Standardization Sector, 2001.
- [Jager 1995] Jager, R. "Fuzzy Logic in Control". Ph.D. Thesis - Department of Electrical Engineering - Delft Technische Universiteit, 1995. Disponível em: <http://users.pandora.be/jati/renej/phd/rjphd.pdf>.
- [Jain et al. 1999] Jain, L. C.; Jain, R. K.; . Vol. 11; Scientific., W. "Advances in Fuzzy Systems – Applications and Theory". Hybrid Intelligent Engineering Systems - World Scientific, vol.11, 1999.

- [Kamienski e Sadok 2000] Kamienski, C. A.; Sadok, D. "*Qualidade de Serviço na Internet*". Minicurso apresentado na SBRC - Sociedade Brasileira de Redes de Computadores - Belo Horizonte, Minas Gerais - Brasil, 2000.
- [Klir e Yuan 1995] Klir, G.; Yuan, B. *Fuzzy Sets and Fuzzy Logic. Theory and Applications*. Editora Prentice Hall, 1995.
- [Koliver et al. 2004] Koliver, C.; Farines, J.; Nahrstedt, K. "*QoS Adaptation Based on Fuzzy Theory*". In *Soft Computing for Communications*, Editora Springer-Verlag, pp. 245-267, 2004.
- [Lacerda e Carvalho 1980] Lacerda, E. G. M.; Carvalho, A. C. P. L. F. "*Introdução aos Algoritmos Genéticos*". XIX Congresso Nacional da Sociedade Brasileira de Computação, vol.2, pp.51-126, 1980.
- [Lu 1996] Lu, G. *Communication and Computing for Distributed Multimedia Systems*. 1ª Edição. Editora Artech House, Inc., 1996.
- [Lunardi e Dotti 2001] Lunardi, S.; Dotti, F. "*Uma Camada de Adaptação à Qualidade de Serviço na Internet para Aplicações Multimídia*". XI Simpósio Brasileiro de Redes de Computadores, vol.1, Florianópolis, SC, Brasil, Outubro, 2001.
- [Mamdani 1974] Mamdani, E. H. "*Application of Fuzzy Algorithms for Control of Simple Dynamic Plant*". IEE (Control and Science), vol.121, no.12, pp.298-316, Dezembro, 1974.
- [Mamdani 1977] Mamdani, E. H. "*Application of Fuzzy Logic to Approximate Reasoning using Linguistic Synthesis*". IEEE Transactions on Computers, vol.26, no.12, pp.1182-1191, Dezembro, 1977.
- [Mamdani e Assilian 1975] Mamdani, E. H.; Assilian, S. "*An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller*". International Journal of Man-Machine Studies, vol.7, no.1, pp.1-13, 1975.
- [Mbise e Woods 2003] Mbise, M.; Woods, J. "*Multimedia Interaction and a Quality of Experience Metric Based on Application and Network Level Considerations*". Packet Video, vol.1, École Polytechnique de L'université de Nantes, France, Abril, 2003.
- [Medeiros et al. 2001] Medeiros, A. V.; Maitelli, A. L.; Filho, O. G. "*Otimização das Funções de Pertinência de um Controlador Nebulosos utilizando Algoritmos Genéticos*". V SBAI - 5º Simpósio Brasileiro de Automação Inteligente, Sessão 21, Bauru - São Paulo - Brasil, 2001.
- [Melo 2001] Melo, E. T. "*Qualidade de Serviço em Redes IP com DiffServ: Avaliação através de Medições*". Dissertação de Mestrado - Universidade Federal de Santa Catarina - Florianópolis, Santa Catarina, Brasil - 2001. Disponível em: <http://lrg.ufsc.br/teses/Melo.pdf>.

- [Mendel 1995] Mendel, J. M. "*Fuzzy Logic Systems for Engineering: A Tutorial*". IEEE Special Issue on Fuzzy Logic, vol.83, no.3, pp.345-377, Março, 1995.
- [Michalewicz 1996] Michalewicz, Z. Genetic Algorithms + Data Structures = Evolution Programs. Editora Springer, 1996.
- [Mitchell 1996] Mitchell, M. An Introduction to Genetic Algorithms. MIT Press, 1996.
- [Mondragon 1999] Mondragon, E. H. "*The Design of a Fuzzy Traffic Controller for ATM Networks*". HCS Research Lab - University of Florida, 1999.
- [Monteiro et al. 2002] Monteiro, J.; Sampaio, L.; Figueredo, M. "*Qualidade de Serviço: Diagnóstico e Alternativas*". Relatório Parcial do Grupo de Trabalho de Qualidade de Serviço (GT-QoS) da RNP2, 2002.
- [Nahrstedt e Li 2000] Nahrstedt, K.; Li, B. "*A Control-Based Middleware Framework for Quality of Service Adaptations*". IEEE Journal on Selected Areas in Communications, vol.17, no.9, pp.1632-1650, Setembro, 2000.
- [Nahrstedt e Steinmetz 1995] Nahrstedt, K.; Steinmetz, R. "*Resource Management in Networked Multimedia Systems*". IEEE Computer Society - Electronic Edition, vol.28, no.5, pp.52-63, Maio, 1995.
- [Oliveira 2001] Oliveira, J. C. "*Sistemas Multimídia*". Material para o curso de Sistemas Multimídia do IME-USP, 2001.
- [Pal e Pal 2003] Pal, T.; Pal, N. R. "*SOGARG: A Self-Organized Genetic Algorithm-Based Rule Generation Scheme for Fuzzy Controllers*". IEEE Transactions on Evolutionary Computation, vol.7, no.4, Agosto, 2003.
- [Pedrycz e Gomide 1998] Pedrycz, W.; Gomide, F. An Introduction to Fuzzy Sets: Analysis and Design. Editora The MIT Press, 1998.
- [Price 2001] Price, D. "*Requirements to Support Video and Audio Applications*". JANET QoS Workshop, London, England, Julho, 2001.
- [Rezende 2003] Rezende, S. O. Sistemas Inteligentes - Fundamentos e Aplicações. 1ª Edição. Editora Manoele, 2003.
- [Sekercioglu et al. 2001] Sekercioglu, Y.; Ptsillides, A.; Vasilakos, A. "*Computational intelligence in management of ATM networks: A survey of current state of research*". Soft Computing Journal, vol.5, no.4, pp.257-263, Agosto, 2001.
- [Sharda 2003] Sharda, N. "*Multimedia: Fundamentals and Design Principles*". Notas de aula da disciplina Introduction to Multimedia - Computer Science and Multimedia School of Computer Science and Mathematics - Victoria University, Australia, 2003.

- [Shenker e Wroclawski 1997] Shenker, S.; Wroclawski, J. "*General Characterisation Parameters for Integrated Service Network Elements*". RFC 2215, 1997.
- [Silveira 2000] Silveira, R. M. "*Técnica para Ajuste dos Coeficientes de Quantização do Padrão MPEG em Tempo-Real*". Tese de Doutorado - Departamento de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo - São Paulo, Brasil, 2000.
- [Stallings 1999] Stallings, W. SNMP, SNMPv2, SNMPv3 and Rmon 1 and 2. 1ª Edição. Editora Addison Wesley, 1999.
- [Takagi e Sugeno 1983] Takagi, T.; Sugeno, M. "*Derivation of Fuzzy Control Rules from Human Operator's Control Actions*". Proceedings of the IFAC Symp. on Fuzzy Information, Knowledge Representation and Decision Analysis, pp.55–60, Marseilles, France, 1983.
- [Tanenbaum 1996] Tanenbaum, A. S. Computer Networks. 3ª Edição. Editora Prentice Hall, 1996.
- [Tanscheit 2002] Tanscheit, R. "*Fundamentos de Lógica Fuzzy e Controle Fuzzy*". Notas de aula do curso de Sistemas Inteligentes do Núcleo de Pesquisa em Inteligência Computacional Aplicada - Departamento de Engenharia Elétrica - PUC-Rio de Janeiro, 2002. Disponível em: http://www.ica.ele.puc-rio.br/cursos/download/SI-Logica_Control_Fuzzy.pdf.
- [Teixeira e Barbar 2003] Teixeira, M. A.; Barbar, J. S. "*Gerenciamento de QoS para as Aplicações Multimídia no Sistema Final*". V Semana de Informática - Seminfo e 1º ErCOMP - FACIC/FUOM, vol.1, Formiga, Minas Gerais - Brasil, Novembro, 2003.
- [T-Mobil 1999] T-Mobil 6.3.10 - TSG-SA Working Group 1. "*Requirements of IP based Multimedia-Applications over UMTS*". Surrey, UK, Maio, 1999.
- [Tsukamoto 1979] Tsukamoto, Y. "*An Approach to Fuzzy reasoning Methods*". Advances in Fuzzy Set Theory, pp.137-149, In Ragade, R. & Yager, R. editors, Amsterdam - North-Holland, 1979.
- [Vogel et al. 1995] Vogel, A.; Kerherve, B.; Bochmann, G. V.; Gecsei, J. "*Distributed Multimedia Applications and Quality of Service: - A Survey*". IEEE Multimedia Journal, vol.2, no.2, pp.10-19, Agosto, 1995.
- [Zadeh 1965] Zadeh, L. A. "*Fuzzy Sets*". Information and Control, vol.8, no.3, pp.338-353, Junho, 1965.
- [Zadeh 1973] Zadeh, L. A. "*Outline of a new Approach to the Analysis of Complex Systems and Decision Processes*". IEEE Transactions on Systems, Man and Cybernetics - SMC, vol.3, no.1, pp.28-44, Outubro, 1973.

- [Zadeh 1975] Zadeh, L. A. “*The Concept of Linguistic Variable and its Application to Approximate Reasoning*”. Information Sciences, vol.8, pp.199-249 (part I), 301-57 (part II), 1975.
- [Zadeh 1996] Zadeh, L. A. “*Fuzzy Logic - Computing With Words*”. IEEE Transactions On Fuzzy Systems, vol.4, no.2, pp.103-111, Maio, 1996.
- [Zimmermann 1990] Zimmermann, H.-J. Fuzzy Set Theory and its Applications. Edição 2^a. Editora Kluwer Academic Publishers, 1990.

Apêndice A:

Análise das Bases de Regras do Sistema Nebuloso

Neste apêndice serão expostas e comentadas as regras das bases de regras do sistema nebuloso para as duas aplicações em estudo de caso, definidas pelo especialista e posteriormente validadas com o uso do emulador NIST Net.

Em [ITU 2001], [T-Mobil 1999] e [Price 2001] encontram-se considerações a respeito de aplicações multimídias, suas características e restrições. Segundo estes trabalhos, uma aplicação de videoconferência pode ter sua qualidade de serviço afetada pelos parâmetros vazão, delay, jitter e perda de pacotes, entre outros, mas é muito sensível à perda de pacotes. No caso da aplicação de vídeo sob demanda, existe tolerância com relação ao delay, mas é também muito sensível à perda de pacotes.

Apesar de serem duas aplicações com características distintas, a videoconferência e o vídeo sob demanda podem usar a mesma base de regras, pois o parâmetro de maior relevância para a QoS é o mesmo para as duas aplicações. Na aplicação de vídeo sob demanda, em situações onde o delay não deve ser tão considerado quando está com um valor classificado como “ruim” para uma transmissão, ocorrem dois fatos: ou a situação não existe no mundo real ou a perda de pacotes é intolerável, o que torna a QoS intolerável. Deste modo, apenas uma base de regras será analisada.

A seguir, são apresentadas as regras que produzem uma QoS considerada Ótima. São regras em que um balanceamento dos parâmetros gera uma situação em que é possível transmitir sem atrasos ou perdas que influenciem de forma significativa a qualidade do que está sendo exibido. Assim, o vídeo transmitido através do NIST Net pôde ser visualizado de

forma contínua e quando ocorriam perdas, provocando defeitos na imagem, esses eram quase imperceptíveis.

	Vazão	Delay	Jitter	Perda de Pacotes	QoS
1.	Ótimo	Ótimo	Ótimo	Ótimo	Ótima
2.	Ótimo	Ótimo	Ótimo	Aceitável	Ótima
3.	Ótimo	Ótimo	Aceitável	Ótimo	Ótima
4.	Ótimo	Ótimo	Aceitável	Aceitável	Ótima
5.	Ótimo	Ótimo	Intolerável	Ótimo	Ótima
6.	Ótimo	Aceitável	Ótimo	Ótimo	Ótima
7.	Ótimo	Aceitável	Aceitável	Ótimo	Ótima
8.	Aceitável	Ótimo	Ótimo	Ótimo	Ótima
9.	Aceitável	Ótimo	Aceitável	Ótimo	Ótima
10.	Aceitável	Ótimo	Aceitável	Aceitável	Ótima
11.	Aceitável	Ótimo	Intolerável	Ótimo	Ótima
12.	Aceitável	Aceitável	Ótimo	Ótimo	Ótima
13.	Aceitável	Aceitável	Aceitável	Ótimo	Ótima
14.	Intolerável	Ótimo	Ótimo	Ótimo	Ótima

As regras seguintes representam uma QoS classificada como Aceitável. São regras em que uma perda de pacotes e um delay mais acentuados provocam algumas distorções (sombras e partes faltando) e certo congelamento temporário no vídeo que está sendo exibido. Apesar disso, ainda se consegue uma boa visualização e entendimento do vídeo apresentado.

	Vazão	Delay	Jitter	Perda de Pacotes	QoS
1.	Ótimo	Ótimo	Intolerável	Aceitável	Aceitável
2.	Ótimo	Aceitável	Ótimo	Aceitável	Aceitável
3.	Ótimo	Aceitável	Aceitável	Aceitável	Aceitável
4.	Ótimo	Aceitável	Intolerável	Ótimo	Aceitável
5.	Ótimo	Aceitável	Intolerável	Aceitável	Aceitável
6.	Ótimo	Intolerável	Aceitável	Aceitável	Aceitável
7.	Ótimo	Intolerável	Intolerável	Aceitável	Aceitável
8.	Aceitável	Ótimo	Ótimo	Aceitável	Aceitável
9.	Aceitável	Ótimo	Intolerável	Aceitável	Aceitável
10.	Aceitável	Aceitável	Ótimo	Aceitável	Aceitável
11.	Aceitável	Aceitável	Aceitável	Aceitável	Aceitável
12.	Aceitável	Aceitável	Intolerável	Ótimo	Aceitável
13.	Aceitável	Aceitável	Intolerável	Aceitável	Aceitável
14.	Aceitável	Intolerável	Aceitável	Aceitável	Aceitável
15.	Aceitável	Intolerável	Intolerável	Aceitável	Aceitável
16.	Intolerável	Ótimo	Ótimo	Aceitável	Aceitável
17.	Intolerável	Ótimo	Aceitável	Ótimo	Aceitável
18.	Intolerável	Ótimo	Aceitável	Aceitável	Aceitável
19.	Intolerável	Ótimo	Intolerável	Ótimo	Aceitável
20.	Intolerável	Ótimo	Intolerável	Aceitável	Aceitável
21.	Intolerável	Aceitável	Ótimo	Ótimo	Aceitável
22.	Intolerável	Aceitável	Ótimo	Aceitável	Aceitável
23.	Intolerável	Aceitável	Aceitável	Ótimo	Aceitável
24.	Intolerável	Aceitável	Aceitável	Aceitável	Aceitável
25.	Intolerável	Aceitável	Intolerável	Ótimo	Aceitável
26.	Intolerável	Aceitável	Intolerável	Aceitável	Aceitável

Por fim, as regras que demonstram uma QoS Intolerável. Nestas regras, o delay e a perda de pacotes são intoleráveis, produzindo uma exibição deficiente, fazendo com que não seja possível discernir o que está sendo exibido, nem entender as seqüências das ações no vídeo.

	Vazão	Delay	Jitter	Perda de Pacotes	QoS
1.	Ótimo	Intolerável	Ótimo	Intolerável	Intolerável
2.	Ótimo	Intolerável	Aceitável	Intolerável	Intolerável
3.	Ótimo	Intolerável	Intolerável	Intolerável	Intolerável
4.	Ótimo	Ótimo	Ótimo	Intolerável	Intolerável
5.	Aceitável	Ótimo	Ótimo	Intolerável	Intolerável
6.	Aceitável	Intolerável	Ótimo	Intolerável	Intolerável
7.	Aceitável	Intolerável	Aceitável	Intolerável	Intolerável
8.	Aceitável	Intolerável	Intolerável	Intolerável	Intolerável
9.	Intolerável	Intolerável	Ótimo	Intolerável	Intolerável
10.	Intolerável	Intolerável	Aceitável	Aceitável	Intolerável
11.	Intolerável	Intolerável	Aceitável	Intolerável	Intolerável
12.	Intolerável	Intolerável	Intolerável	Aceitável	Intolerável
13.	Intolerável	Intolerável	Intolerável	Intolerável	Intolerável

Somando-se todas as regras mostradas acima, tem-se um total de 53 regras. Considerando que a base de regras tem 81 regras, significa que 28 regras não representam situações possíveis de ocorrer no mundo real e, portanto, não necessitam fazer parte da base. Estas regras são apresentadas abaixo. Trata-se de regras cujas classificações dos parâmetros que as formam não representam uma situação coerente, como por exemplo, uma regra na qual a vazão e o delay são intoleráveis e a perda de pacotes de pacotes é classificada como ótima.

Mesmo com este conhecimento, deixou-se a cargo do algoritmo genético eliminar estas regras.

	Vazão	Delay	Jitter	Perda de Pacotes
1.	Ótimo	Ótimo	Aceitável	Intolerável
2.	Ótimo	Ótimo	Intolerável	Intolerável
3.	Ótimo	Aceitável	Ótimo	Intolerável
4.	Ótimo	Aceitável	Aceitável	Intolerável
5.	Ótimo	Aceitável	Intolerável	Intolerável
6.	Ótimo	Intolerável	Ótimo	Ótimo
7.	Ótimo	Intolerável	Ótimo	Aceitável
8.	Ótimo	Intolerável	Aceitável	Ótimo
9.	Ótimo	Intolerável	Intolerável	Ótimo
10.	Aceitável	Ótimo	Aceitável	Intolerável
11.	Aceitável	Ótimo	Intolerável	Intolerável
12.	Aceitável	Aceitável	Ótimo	Intolerável
13.	Aceitável	Aceitável	Aceitável	Intolerável
14.	Aceitável	Aceitável	Intolerável	Intolerável
15.	Aceitável	Intolerável	Ótimo	Ótimo
16.	Aceitável	Intolerável	Ótimo	Aceitável
17.	Aceitável	Intolerável	Aceitável	Ótimo
18.	Aceitável	Intolerável	Intolerável	Ótimo
19.	Intolerável	Ótimo	Ótimo	Intolerável
20.	Intolerável	Ótimo	Aceitável	Intolerável
21.	Intolerável	Ótimo	Intolerável	Intolerável
22.	Intolerável	Aceitável	Ótimo	Intolerável
23.	Intolerável	Aceitável	Aceitável	Intolerável
24.	Intolerável	Aceitável	Intolerável	Intolerável
25.	Intolerável	Intolerável	Ótimo	Ótimo
26.	Intolerável	Intolerável	Ótimo	Aceitável
27.	Intolerável	Intolerável	Aceitável	Ótimo
28.	Intolerável	Intolerável	Intolerável	Ótimo

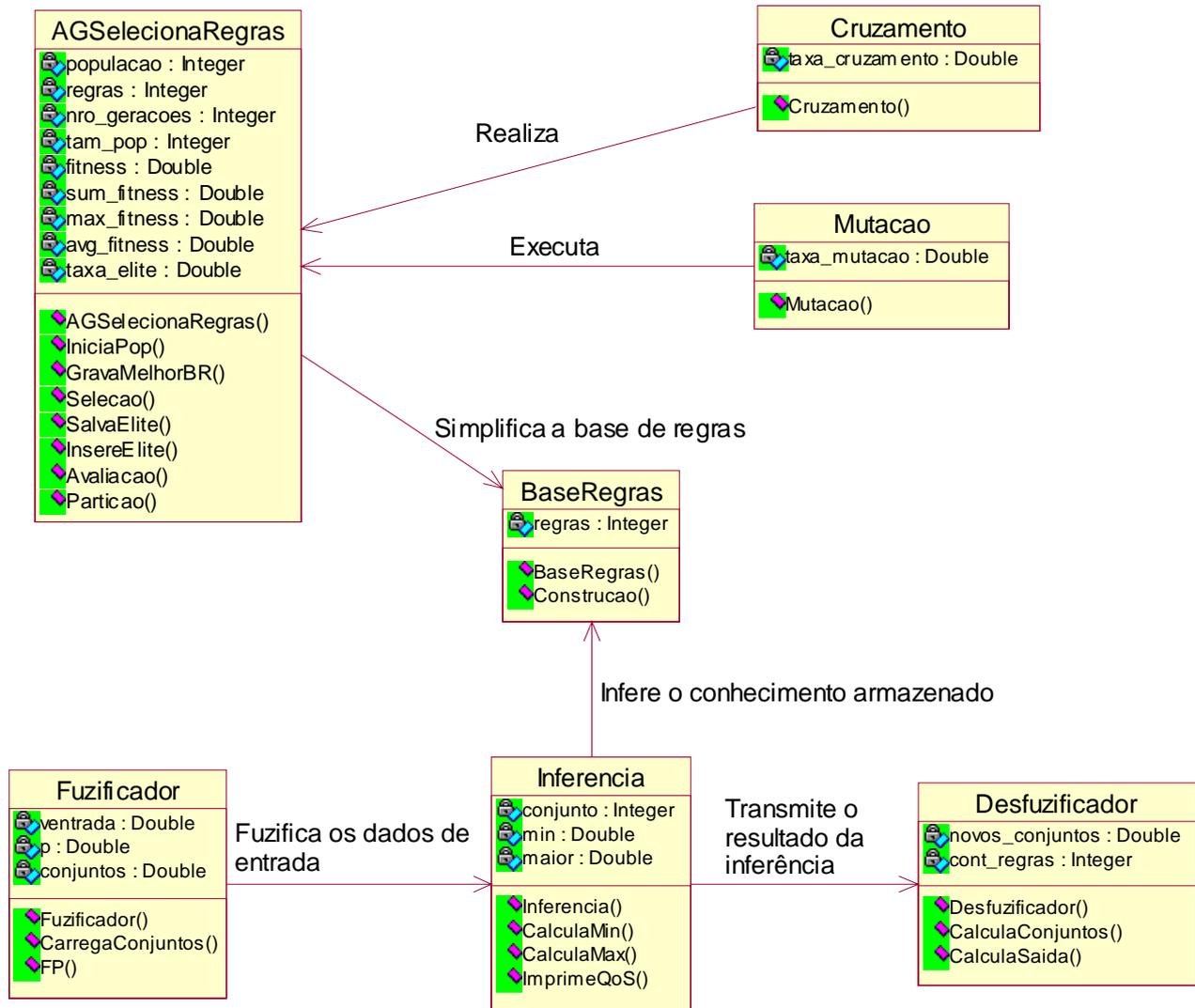
Apêndice B:

Modelagem do Sistema Genético Nebuloso

Baseado em Regras

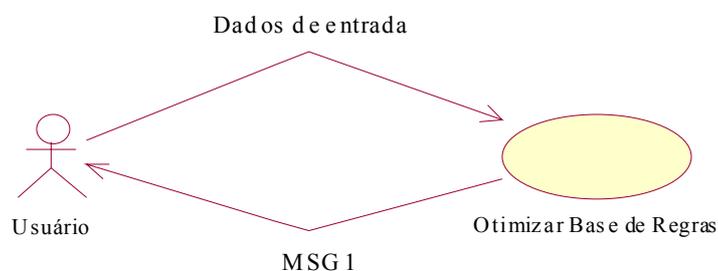
O objetivo deste apêndice é apresentar a documentação do Sistema Genético Nebuloso desenvolvido. São mostrados o diagrama de classes e os casos de uso, seguidos da descrição de seqüência de processamento das informações, devidamente enumeradas.

B.1 Diagrama de Classes



B.2 Diagramas de Casos de Uso

Caso de Uso 1: Otimizar Base de Regras



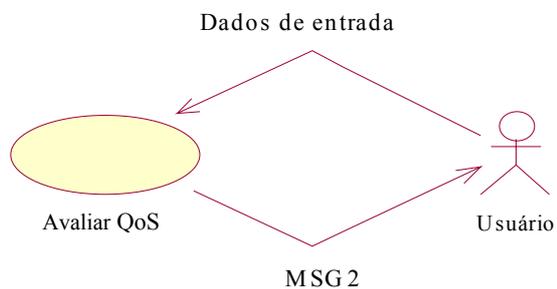
Curso Normal

1. Usuário informa que deseja otimizar uma base de regras.
2. Sistema solicita os dados referentes a esta base de regras.
3. Sistema solicita a base de dados.
4. Usuário informa a base de dados.
5. Sistema solicita a base de regras a ser otimizada.
6. Usuário fornece a base de regras.
7. Sistema solicita a partição dos conjuntos nebulosos.
8. Usuário fornece a partição dos conjuntos nebulosos.
9. Sistema verifica as regras redundantes e/ou desprezíveis.
10. Sistema emite MSG 1 informando que a base de regras foi otimizada com sucesso.

Curso Alternativo

1. Usuário informa que deseja otimizar uma base de regras.
2. Sistema solicita os dados referentes a esta base de regras.
3. Sistema solicita a base de dados.
4. Usuário informa a base de dados.
5. Sistema solicita a base de regras a ser otimizada.
6. Usuário fornece a base de regras.
7. Sistema solicita a partição dos conjuntos nebulosos.
8. Usuário fornece a partição dos conjuntos nebulosos.
9. Sistema verifica que não há regras redundantes e/ou desprezíveis.
10. Sistema emite MSG 1 informando que a base de regras já estava otimizada.

Caso de Uso 2: Avaliar QoS



Curso Normal

1. Usuário informa que deseja avaliar a QoS de uma determinada conexão.
2. Sistema solicita os parâmetros da conexão através dos quais a QoS será avaliada.
3. Usuário informa os parâmetros da conexão: delay, jitter, vazão e perda de pacotes.
4. Sistema solicita os parâmetros dos conjuntos nebulosos referentes a cada variável da aplicação.
5. Usuário informa os parâmetros dos conjuntos nebulosos.
6. Sistema solicita a base de regras.
7. Usuário informa a base de regras previamente otimizada.
8. Sistema calcula a QoS da aplicação a partir dos dados fornecidos.
9. Sistema emite MSG 2 informando a QoS obtida pelo sistema.