

Modelo de simulação de centrais de
teleatendimento baseado em sistemas
multiagentes

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.



PUC Minas
Programa de Pós-graduação em Informática

FOLHA DE APROVAÇÃO

Modelo de simulação de centrais de teleatendimento baseado em sistemas multiagentes

CLÁUDIO LÚCIO DO VAL LOPES

Dissertação defendida e aprovada pela seguinte banca examinadora:

Ricardo Luiz Martins Ferreira

Prof. Dr. Ricardo Luiz Martins Ferreira - Orientador

Doutor em Automática e Otimização (Universite de Toulouse III, França)

Luis Enrique Zarate Galvez

Prof. Luis Enrique Zarate Galvez (PUC Minas)
Doutor em Engenharia Metalúrgica e de Minas (UFMG)

Prof. Dr. Kleber Gonçalves do Patrocínio Junior

Prof. Dr. Kleber Gonçalves do Patrocínio Junior (PUC Minas)
Doutor em Ciência da Computação (UFMG)

Belo Horizonte, 27 de março de 2007.

FICHA CATALOGRÁFICA
Elaborada pela Biblioteca da
Pontifícia Universidade Católica de Minas Gerais

L864m	<p>Lopes, Cláudio Lúcio do Val Modelo de simulação de centrais de teleatendimento baseado em sistemas multiagentes. Belo Horizonte, 2007. 100f.</p>
	<p>Orientador: Ricardo Martins Ferreira Poley Dissertação (Mestrado) – Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Informática Bibliografia.</p>
	<p>1. Centros de atendimento ao cliente. 2. Simulação (Computadores). 3. Agentes inteligentes (Software). I. Poley, Ricardo Martins Ferreira. II. Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação Em Informática. III. Título.</p>
	CDU: 681.3.092

As centrais de teleatendimento são sistemas sócio-técnicos em que o comportamento dos clientes e dos atendentes são relacionados a indicadores de desempenho. Cada vez mais, as centrais, têm ganhado espaço nas organizações visto a necessidade de estabelecer e manipular um grande número de contatos com os clientes. Isto tem feito com que empresas busquem melhorar seus processos, serviços e tecnologia nas centrais de teleatendimento.

Uma das atribuições da área de conhecimento denominada engenharia de serviços é desenvolver princípios e ferramentas que suportem e equilibrem indicadores de desempenho e eficiência sob as várias perspectivas conflitantes dos envolvidos na dinâmica da central. O processo de tomada de decisão em centrais de teleatendimento e seus desafios são temas que têm sido tratados tradicionalmente pela teoria de filas. Os modelos auxiliam os gestores das centrais na busca de um modo de operação eficiente e de acordo com as metas estabelecidas na central. Algumas abordagens utilizadas são modelos analíticos, modelos de simulação e modelos de programação matemática originados da teoria de filas.

Este trabalho propõe um modelo de simulação que procura representar a complexidade e a dinâmica das centrais de teleatendimento. O modelo foi desenvolvido utilizando uma abordagem que faz interseção com a Inteligência Artificial Distribuída (sistemas multiagentes) que se baseia no comportamento social de indivíduos, suas ações e interações. A utilização desta aborda

The call centers are socio-technical system in which the behaviour of clients and service reps is intertwined with performance indicators. The companies have necessity to establish and to maintain contact with their customers, due to this, the call center are gaining more attention and importance. This fact have done with the call centers try to improve your process, services, and technologies.

The service engineering is developing tools to solve the problem to obtain better call center performance indicators subject to costs constraints and conflicting perspectives. The decision making process in call centers are issues that have been treating by queuing theory models. These models help call centers managers to reach the quality and costs targets established. Different approaches originated from queuing theory have been used: analytical models, discrete simulation models, and optimization models.

This work proposes a call center simulation model based on multiagent systems. The model objective is to represent the complexity and the work dynamics of an income call center. This work describes the conceptual architecture of the simulation model. The model is based on the agents (service reps, managers, and customers) interactions and their social behavior.

The viability of the multiagent systems approach for call center simulation is verified. It will be shown that this approach is capable to treat some problems that other “classical” models cannot lead. Experiments were done to test the simulation model efficacy. Comparisons were done with analytical models and with real call center data. Some case studies and examples of complex situations treatment will be showed:

- Increase customer patience tax and the impacts of this kind of behaviour in call center performance indicators;
- Skill based routing policies used to route calls to service reps. Others services policies (queue with priorities) and routing policies (based on service reps performance and marginal cost).

A Deus, pela força que me guia, aos
meus pais e a Michelle

Agradecimentos

O resultado desta jornada é fruto de muito trabalho. Trabalho não só do autor, mas de muitas outras pessoas. Pelo apoio e esforço destas pessoas menções e agradecimentos especiais.

Ao Professor Doutor Ricardo Poley por todas as colaborações, pelas reuniões em horário de almoço, pelas críticas e por toda orientação no decorrer deste agradável período de convivência. Minha profunda admiração e respeito, não só por seus conhecimentos que utilizamos nestes anos, mas também pelo ser humano, que dentre outras coisas, me fez compreender o que é necessário para “orientar” pessoas.

Aos nossos queridos amigos e parceiros de trabalho da iniciação científica: Hugo, Rafael e Gustavo. A participação de vocês agregou muito neste trabalho.

Aos colegas e professores do mestrado. Aos funcionários da secretária, principalmente a Giovana, sempre solícita.

Em especial, agradecimento aos meus pais, Lúcia e Luciano, por serem a causa de tudo que faço. Meu amor a vocês é incondicional. Agradeço pelos valores que me deram (meus alicerces) e também pelos pequenos gestos, como levantar mais cedo para fazer café para alguém que tem de sair às 06:30 da manhã para assistir aula.

A minha avó e irmã, sempre queridas.

Ao meu amor Michelle. Seu carinho, amizade, companheirismo e compreensão sempre me ajudaram. Obrigado por me apoiar e estar ao meu lado nos momentos de angústia, por me suportar nos momentos de mau humor (quando algum experimento não dava resultados esperados). Sem sua força esta etapa da minha vida não seria transposta.

Listas de Figuras.....	8
Listas de Tabelas.....	10
Lista de Siglas.....	12
1 Introdução.....	13
2 Revisão Bibliográfica.....	16
2.1 Sistemas Multiagentes.....	16
2.1.1 Inteligência Artificial Distribuída.....	16
2.1.2 Sistemas Multiagentes.....	17
2.1.3 Aplicações e ferramentas computacionais.....	20
2.2 Modelos de simulação baseados em sistemas multiagentes.....	21
2.2.1 Swarm - um “framework” para simulação multiagente.....	23
2.3 As centrais de teleatendimento.....	25
2.3.1 Processo e funcionamento das centrais.....	26
2.3.2 Indicadores de desempenho.....	28
2.3.3 Níveis de decisão e desafios gerenciais.....	31
2.3.4 Modelos “clássicos” para auxílio à tomada de decisão	32
2.3.5 Simulação em centrais de teleatendimento.....	36
3 Modelo de Simulação de uma central de teleatendimento.....	39
3.1 Descrição e apresentação.....	42
3.2 Detalhamento de alguns agentes.....	46
3.3 Descrição de parâmetros de entrada e de saída do modelo	48
3.3.1 Parâmetros de entrada.....	48
3.3.2 Saídas do modelo.....	50
4 Resultados Experimentais.....	52
4.1 Os Modelos de fila clássicos e o modelo de simulação.....	52
4.2 O modelo de Stanford e Grassmam e o modelo de simulação.....	56
4.3 Modelo de simulação – comparação com dados reais de uma central de teleatendimento.....	58
4.4 Capacidades adicionais do modelo de simulação.....	61
4.4.1 Simulação do abandono e o comportamento de espera dos clientes.....	61
4.4.1.1 Análise do abandono na perspectiva de uma base de dados real.....	61
4.4.1.2 Comportamento crescente para o tempo de espera na fila e suas implicações nos indicadores da central de teleatendimento.....	64
4.4.2 Roteamento baseado em habilidades.....	66
4.4.2.1 Definições sobre política de roteamento baseado em habilidades.....	67
4.4.2.2 Estudo de caso.....	68
4.4.3 Filas com prioridades e políticas de roteamento baseada em custo marginal e	

desempenho dos atendentes.....	71
4.4.3.1 Caso base - custos das ligações obtidos com o modelo de simulação.....	71
4.4.3.2 Políticas de atendimento e roteamento de chamadas para reduzir o custo com ligações.....	72
4.4.3.3 Resultados e comparações.....	76
5 Conclusões e comentários.....	78
5.1 Trabalhos futuros.....	79
5.2 Comentários	80
Apêndice A – Diagrama de classes do modelo	82
Apêndice B – Base de dados e programas utilizados para obtenção de resultados.....	85
Apêndice C – Considerações sobre simulação.....	89
Apêndice D – Algoritmo de alocação de tarefas em “cluster” de computadores baseado no custo marginal.....	92
Anexo I – Mensagens eletrônicas enviadas e recebidas.....	95
6 Bibliografia.....	96

Listas de Figuras

Figura 2.1: Esquema genérico para representar agentes autônomos proposto por [Wooldridge, 1999].....	18
Figura 2.2: Classificação de modelos de interação social entre agentes autônomos [Sichman,2005].....	19
Figura 2.3: Esquema geral simplificado de modelo de simulação Swarm.....	24
Figura 2.4: Um exemplo de processo de funcionamento de uma central de teleatendimento..	28
Figura 2.5 Gráfico que demonstra a variação das quantidades de chamadas por hora do dia, obtidos em uma central de teleatendimento real [Valladares, 2007].....	35
Figura 2.6 Esquema de simulação em centrais de teleatendimento, adaptado [Mehrotra e Fama, 2003].....	37
Figura 3.1 Descrição conceitual do modelo com os agentes, interações e objetos da simulação para as centrais de teleatendimento	42
Figura 3.2 Descrição do modelo de simulação em etapas. As etapas de A até F representam a evolução gradual na simulação das centrais de teleatendimento.....	43
Figura 3.3 Diagrama de estados sob a perspectiva das chamadas.....	44
Figura 3.4 Diagrama de estados sob a perspectiva do agente cliente.....	45
Figura 3.5 Descrição conceitual do modelo implementado neste trabalho com os agentes, interações e objetos da simulação para as centrais de teleatendimento	46
Figura 3.6 Diagrama de classe para o tipo de agente cliente - “(F)5.6753(3)-0.295585(.)-0.147792(5)-0.29	

g,

Figura 4.7 Resultados para o indicador do tempo médio de espera na fila - W_i em segundos, para cada um dos cenários do estudo de caso utilizando a política de roteamento baseada em habilidades.....	70
Figura 4.8 Resultados para o indicador de nível de serviço - ρ_i , para cada um dos cenários do estudo de caso utilizando a política de roteamento baseada em habilidades.....	70
Figura 4.9 Pseudo-algoritmo RBD – roteamento baseado em desempenho aplicado no estudo de caso para reduzir o custo total com ligações na central de teleatendimento.....	73
Figura 4.10 Pseudo-algoritmo RBCM – roteamento baseado em custo marginal aplicado no estudo de caso para reduzir o custo total com ligações na central de teleatendimento.....	76
Figura C.1: Esquema com macros passos utilizado na criação de modelos para processos de tomada de decisão, adaptado de [Goldbarg e Luna, 1999].....	90
Figura D.1 Comparação utilizando exemplo de um “cluster” fictício sem (gráfico 1) e com (gráfico 2) a utilização da proposta apresentada em [Barak et al., 2000].....	93
Figura D.2 Algoritmo ASSIGN-U para alocação de recursos baseado no conceito de custo marginal, adaptado da proposta apresentada em [Barak et al., 2000].....	94

Listas de Tabelas

Tabela 2.1: Resumo comparativo entre arquiteturas multiagentes reativas e deliberativas [Sichman, 2005]	18
Tabela 2.2: Resumo de áreas que utilizam simulação baseada em sistemas multiagentes [Macal e North, 2005].....	23
Tabela 2.3: Resumo de outros indicadores utilizados nas centrais de teleatendimento e suas unidades de medidas, adaptado de [Saltzman e Mehrotra, 2004].....	30
Tabela 4.1 Indicadores utilizados para experimentos comparativo entre resultados do modelo de simulação e modelos analíticos do tipo “Erlang”	52
Tabela 4.2 Parâmetros descritivos dos modelos utilizados no experimentos de comparação com modelo analíticos do tipo “Erlang”.....	53
Tabela 4.3 Resultados com as médias dos indicadores na execução do modelo de simulação com 10.000, 20.000, 40.000, 100.000, 500.000 e 1.000.000 de chamadas.....	53
Tabela 4.4 Resultados com o desvio padrão dos indicadores na execução do modelo de simulação com 10.000, 20.000, 40.000, 100.000, 500.000 e 1.000.000 chamadas.....	54
Tabela 4.5 Diferenças percentuais entre os resultados dos indicadores obtidos com a execução do modelo de simulação com 100.000 de chamadas e os modelos analíticos.....	56
Tabela 4.6 Parâmetros descritivos de um modelo utilizado no experimento representando uma central com dois tipos de serviços diferentes [Stanford e Grassmann, 1998].....	56
Tabela 4.7 Resumo comparativo de resultados entre o modelo de simulação e o trabalho de [Stanford e Grassmann, 1998] para o caso de uma central com dois tipos de serviços diferentes.....	57
Tabela 4.8 Dados reais com número de chamadas, número de atendentes e tempo médio de atendimento de uma central de teleatendimento, valores correspondentes a um dia “típico” de funcionamento [Valladares, 2007].....	59
Tabela 4.9 Os dez maiores grupos de clientes por percentual do total. Contém o número de clientes e o percentual de abandono para cada um dos grupos apresentados. Informação derivada da base de dados disponibilizada em [Mandeulbaum, 2006].....	63
Tabela 4.10 Indicadores obtidos através da realização de três conjuntos de experimentos, estes experimentos levam em conta a questão do comportamento do cliente quanto a taxa de crescimento da paciência que acontece em situações que o cliente faz chamadas durante um mesmo dia.....	66
Tabela 4.11 Parâmetros para representar o estudo de caso de um central de teleatendimento fictícia utilizada no estudo de caso com a abordagem de política de roteamento baseada em habilidades.....	68
Tabela 4.12 Valores por minuto para cada tipo de chamada em um central fictícia idealizada para avaliação de fila com prioridades e políticas de roteamento de chamadas.....	72

Tabela 4.13 Resultados obtidos com o modelo de simulação refletindo o funcionamento da central de teatendimento, incluindo os valores de custo das ligações e análise do percentual de ocupação dos atendentes.....72

Tabela 4.14 Resultados obtidos com o modelo de simulação utilizando as alternativas propostas inclui os valores de custo das ligações, indicadores e análise do percentual de ocupação dos atendentes.....77

Lista de Siglas

ANEEL - Agência Nacional de Energia Elétrica

DAC – Distribuidor Automático de Chamadas

IA – Inteligência Artificial

IAM – Inteligência Artificial Monolítica

IAD – Inteligência Artificial Distribuída

ICT – Integração Computador Telefone

PA – Posto de Atendimento

RBD – Roteamento baseado em Desempenho

RBCM – Roteamento baseado em Custo Marginal

SDP – Solução Distribuída de Problemas

SMA – Sistemas Multiagentes

URA – Unidade de Respostas Audíveis

1 Introdução

As centrais de teleatendimento, também conhecidas como “*call centers*”, são atualmente uma das principais formas de interação entre as empresas e seus clientes. Uma central de teleatendimento é um conjunto de recursos (hardware, software e pessoal para atendimento) que está habilitado a prover serviços via telefone [Gans et al., 2006]. Uma central é um sistema sócio-técnico em que o comportamento dos clientes e dos atendentes são relacionados a indicadores que representam níveis de qualidade de serviço [Koole e Mandelbaum, 2002].

As centrais de teleatendimento tem ganhado cada vez mais espaço nas organizações e sua importância econômica é evidenciada com várias pesquisas, como por exemplo, o site <http://www.callcenter.inf.br> que destaca o valor de receita US\$ 1,6 bilhão obtida no Brasil em 2005 pelas centrais de teleatendimento [L'Ecuyer e Avramidis, 2005], [Mandelbaum et al., 2001] e [Sharp, 2003]. A necessidade de estabelecer e manipular um grande número de contatos com os clientes têm feito as empresas buscarem melhorar seus processos, serviços e tecnologia nas centrais de teleatendimento.

O estudo das centrais de teleatendimento está concentrado em uma área de conhecimento denominada engenharia de serviços. A engenharia de serviços “busca desenvolver, com base científica, princípios e ferramentas (frequentemente culminando em softwares) que suportem e equilibrem qualidade de serviço e eficiência sob perspectivas conflitantes de clientes, atendentes, gerentes e, muito comumente, a sociedade” [Mandelbaum et al., 2001]. A teoria de filas constitui-se parte da base da engenharia de serviços. As centrais de teleatendimento podem ser modeladas como uma rede de filas [Whitt, 2002].

O processo de tomada de decisão em centrais de teleatendimento e seus desafios são temas que tem sido tratados tradicionalmente pela teoria de filas. Estes modelos auxiliam os gestores das centrais na busca de um modo de operação eficiente e de acordo com as metas estabelecidas na central. Diferentes tipos de modelos são utilizados para representar as centrais. Entre as abordagens utilizadas estão os modelos analíticos, modelos de simulação e modelos de programação matemática.

A maioria dos modelos encontrados na literatura são analíticos e são utilizados como ferramenta para auxílio ao suporte a decisão, no entanto a literatura aponta algumas limitações:

- Dificuldade de representar o comportamento humano [Koole, 2005] (os atendentes - fatores como pausas durante o trabalho, experiência, absenteísmo, rotatividade de pessoal; os clientes – fatores como paciência, qualidade de atendimento e serviço);
- O modelos são “Orientados a qualidade e eficiência” [L'Ecuyer e Avramidis, 2005] e pressupõem que todas as chamadas são atendidas imediatamente, não há atrasos.
- Religações, as religações (ou rechamadas) são comuns nas centrais de teleatendimento e não são conhecidos bons modelos para este problema [Koole, 2005];
- Tipos de chamadas, os modelos analíticos tratam apenas um tipo de serviço/chamada na central, quando na realidade existem vários tipos de serviços prestados pelas centrais, cada um com suas próprias particularidades;
- Variação temporal [L'Ecuyer e Avramidis, 2005]; é comum que exista uma variação do número de chamadas por hora do dia. As curvas de frequência de chamadas são

diferentes entre os dias, meses e anos nas centrais.

As limitações dos modelos analíticos tem sido uma das principais motivações para pesquisas. As incertezas típicas e complexidades operacionais das centrais criam uma “distância” entre muitos modelos e a realidade [L'Ecuyer e Avramidis, 2005]. Uma das alternativas para estas questões tem sido o uso de simulação.

Este trabalho propõe um modelo de simulação que procura representar a complexidade e a dinâmica das centrais de teleatendimento. O modelo foi desenvolvido utilizando uma abordagem que faz interseção com a Inteligência Artificial Distribuída (sistemas multiagentes) que se baseia no comportamento social de indivíduos, suas ações e interações. A utilização desta abordagem se traduz em um modelo que representa atendentes, clientes, equipamentos e gestores utilizando o conceito de “agentes” que estão inseridos em um ambiente da simulação.

Um dos principais objetivos da Inteligência Artificial Distribuída é a construção de sistemas inteligentes formados por entidades autônomas (agentes), caracterizadas e dirigidas por objetivos próprios. O preceito é distribuir inteligência entre um conjunto de agentes. Os sistemas multiagentes formam uma sub-área da Inteligência Artificial Distribuída e concentram-se no estudo de agentes autônomos em um universo multiagente. O termo autônomo designa o fato de que os agentes têm uma existência própria, particular e independente da existência de outros agentes. Cada agente possui um conjunto de capacidades comportamentais que definem seu estado e suas ações a fim de alcançar seus objetivos. A decisão de qual ação a ser realizada é determinada pelo agente considerando as mudanças acontecidas no ambiente em que atua, suas crenças e objetivos.

O enfoque de um sistema multiagente é que um comportamento global organizado pode ser alcançado a partir do comportamento individual dos agentes. A utilização de sistemas multiagentes parte da premissa de que os agentes irão interagir em um domínio dinâmico e complexo.

No modelo, os agentes tomam suas decisões, interagem com outros agentes e objetos, e expressam seus interesses particulares e necessidades específicas de acordo com seus objetivos. A abordagem adotada representa de forma única cada atendente, cada cliente, cada equipamento, cada gestor e suas interações. Os comportamentos e interações foram modelados em um nível granular: o agente. O universo da simulação possui um conjunto de agentes e objetos com suas características; estes configuram-se como os componentes do modelo para simular o funcionamento da central. Estes componentes e suas propriedades existem na simulação com a possibilidade de representar situações reais que acontecem no dia a dia da operação da central e desta forma demonstram o comportamento geral do funcionamento da central de teleatendimento.

O objetivo deste trabalho é verificar a viabilidade da abordagem baseada em sistemas multiagentes para simulação das centrais de teleatendimento; averiguar, pela utilização do modelo e de forma empírica, diferentes cenários partindo das clássicas perguntas “e se” para verificar a correlação entre indicadores de desempenho da central; e mostrar que esta abordagem oferece tratamento para problemas que os modelos “clássicos” não fornecem.

Esta forma de representação apresenta desafios e dificuldades encontrados durante a modelagem dos agentes; por exemplo, no que se refere a simular comportamento humano (“paciência” dos clientes, preferências de atendentes, dentre outros). A representação “atômica” de cada elemento no sistema e suas interações originam um sistema complexo.

Entretanto, esta abordagem de modelagem granular mínima deve fazer o funcionamento geral do sistema convergir para um comportamento que se aproxime da realidade. Apesar das dificuldades encontradas durante o trabalho o modelo apresentou resultados que demonstram a sua eficácia. O trabalho apresenta um conjunto de experimentos realizados:

- Comparação de resultados obtidos com experimentos entre os modelos analíticos e os resultados obtidos com o modelo de simulação baseado em sistemas multiagentes. Os resultados expressos como indicadores de desempenho da central de teleatendimento apresentam uma diferença menor que 6 %. Uma explicação possível para tal diferença também é apresentada;
- Comparação de indicadores reais de uma central de teleatendimento contra os indicadores obtidos com o modelo de simulação. Para este caso o modelo apresenta uma diferença entre os indicadores reais e simulados de 1,38%.

Outros conjuntos de experimentos que demonstram a flexibilidade e aderência do modelo para casos reais são:

- Tratamento de situações em que há um taxa crescente de paciência por parte dos clientes enquanto esperam por atendimento e seus respectivos impactos nos indicadores de desempenho da central;
- Estudo de caso que envolve a melhoria do nível de serviço da central, alterando-se os tipos de habilidades dos atendentes e fazendo uso de uma política de roteamento de chamadas baseada em habilidades;
- Estudo de caso que envolve a utilização de políticas de atendimento (fila com prioridades) e roteamento (baseado em desempenho ou baseado em custo marginal). A proposição destas políticas tem o intuito de buscar formas mais eficientes de rotear e atender as chamadas, possibilitando também a avaliação de suas consequências nos indicadores de desempenho da central.

O restante deste documento está organizado da seguinte forma: o capítulo que se segue relaciona o material bibliográfico utilizado durante a pesquisa e se concentra em sistemas multiagentes, simulação computacional e as centrais de teleatendimento. No capítulo 3 o modelo de simulação é descrito e apresentado (inclusive, com detalhamento de alguns agentes e os parâmetros de entrada e saída do modelo). No capítulo 4 os experimentos realizados são apresentados e os resultados discutidos. Finalmente, o último capítulo, trata das conclusões, trabalhos futuros e comentários.

2 Revisão Bibliográfica

Este capítulo apresenta uma revisão dos principais trabalhos analisados durante a pesquisa. Os conceitos, ferramentas, utilização e vantagens dos sistemas multiagentes são apresentados. A inter-relação entre sistemas multiagentes e modelos de simulação estendem a discussão da aplicação deste paradigma. Um “framework”, *Swarm*, é apresentado como um exemplo de ferramenta que permite o desenvolvimento de modelos de simulação baseado em sistemas multiagentes. Por fim, uma discussão sobre as centrais de teleatendimento é apresentada tratando temas sobre: processo de funcionamento, indicadores de desempenho, níveis de decisão e modelos utilizados para o suporte à tomada de decisão.

2.1 Sistemas Multiagentes

Sistemas multiagentes são compostos por um ambiente com vários agentes autônomos e seus mecanismos e regras de interação. Este paradigma é uma metáfora natural para criação de sistemas sociais artificiais. Sistemas multiagentes são um novo campo da ciência da computação que vêm ganhando atenção desde os anos 90 em função das possibilidades de aplicação em sistemas distribuídos [Wooldridge, 1999]. Esta seção apresenta uma visão geral do tema.

2.1.1 Inteligência Artificial Distribuída

As seguintes conceituações de diferentes autores sobre a inteligência artificial são relevantes no contexto deste trabalho:

- Uma máquina é inteligente se ela é capaz de solucionar uma classe de problemas que requerem inteligência para serem solucionados por seres humanos [McCarthy e Hayes, 1969].
- Inteligência artificial é a parte da ciência da computação que compreende o projeto de sistemas computacionais que exibam características associadas, quando presentes no comportamento humano, à inteligência [Barrand e Feigenbaum, 1981].
- E finalmente, a inteligência artificial é o estudo das faculdades mentais através do uso de modelos computacionais [Charniak e McDermott, 1985].

A inteligência artificial possui várias segmentações. Pode-se ressaltar as seguintes linhas de pesquisa sobre sistemas inteligentes [Bittencourt, 2001]:

- Simbólica: utiliza formalismos do tipo lógico para simular o comportamento inteligente expresso através da linguagem. Utilizada em sistemas especialistas e sistemas baseados em lógica nebulosa;
- Conexionista: visa à modelagem da inteligência humana através de representação dos componentes do cérebro, isto é, de seus neurônios, suas interligações. Base para as Redes Neurais.
- Evolutiva: simula a evolução natural para encontrar soluções de problemas complexos. Utilizada em métodos de otimização. Exemplo: Algoritmos Genéticos e Vida Artificial.

Uma outra segmentação para inteligência artificial é quanto a sua localização espacial:

- A IAM – Inteligência Artificial Monolítica diz respeito a sistemas simples, ausência de modularidade. Um exemplo deste segmento são os sistemas especialistas.
- Já a IAD – Inteligência Artificial Distribuída depende de um determinado conjunto de módulos que são utilizados para resolver, através de diversos tipos de interações, um determinado problema.

As principais motivações para o trabalho em IAD [Bittencourt, 2005] são:

- Explorar a distribuição inerente ao domínio de alguns tipos de aplicações;
- Permitir a integração de sistemas inteligentes existentes de maneira a aumentar a eficiência na solução de problemas.

No contexto da IAD e seu conjunto de módulos é que se apresenta o conceito de agentes como sendo a representação de um “módulo”. Não existe um consenso na área sobre o conceito de agente. Este trabalho irá utilizar do conceito explicitado por Jennings et al. [Jennings et al., 1998]: agente é um sistema computacional situado em um ambiente e que é capaz de realizar ações autônomas e flexíveis no intuito de alcançar seus objetivos. Seguindo esta conceituação é possível ainda, de acordo com esta referência, descrever sistemas multiagentes: sistemas baseados ou formados por um conjunto de agentes e que possuem mecanismos sofisticados de interações entre eles.

2.1.2 Sistemas Multiagentes

A IAD divide-se em duas sub-abordagens: SDP - Solução Distribuída de Problemas e SMA – Sistemas Multiagentes. A SDP tem como objetivo a utilização da capacidade de processamento e da robustez oferecidas pela tecnologia de redes para abordar e resolver problemas complexos de forma distribuída. É caracterizada por agentes pré-programados para interagir e com métodos que visam garantir uma interação eficiente. Já a sub-abordagem de SMA tem como objetivo o estudo das questões sobre agentes que garantam a possibilidade de ação interativa em sociedade. É caracterizada por agentes “racionais”, que agem no sentido de maximizar seus benefícios e/ou minimizar suas perdas. Na abordagem SDP o foco é a solução de um problema, na SMA o foco é o estudo das interações entre os agentes que levem o sistema a um estado que pode corresponder a solução de um problema. No entanto não existe consenso sobre a interseção e abrangência de cada uma destas sub-abordagens dentro da inteligência artificial distribuída [Durfee, 1991].

A área de sistemas multiagentes é multidisciplinar e envolve conceitos provenientes de diversas disciplinas, como por exemplo: biologia, economia, teoria da complexidade, teoria dos jogos, ciências sociais, etologia [Jennings et al., 1998]. Os sistemas multiagentes possuem dois aspectos importantes. O primeiro deles são os próprios agentes (com sua representação interna, racionalidade, autonomia e adaptabilidade) e o segundo trata (representa) as interações entre estes agentes (competição, comunicação, cooperação, coordenação, negociação e organização social). A figura 2.1 mostra um esquema genérico que representa a estrutura interna de um agente [Wooldridge, 1999].

Uma escala crescente de autonomia e complexidade pode ser apresentada com os seguintes tipos de agentes [Sichman, 2005]: agentes reativos, agentes baseados em modelo, agentes baseados em metas, agentes baseados em utilidade e agentes aprendizes. Ainda segundo este trabalho, para os sistemas multiagentes as seguintes arquiteturas são válidas: arquiteturas reativas, deliberativas (também conhecidas como cognitivas) e híbridas.

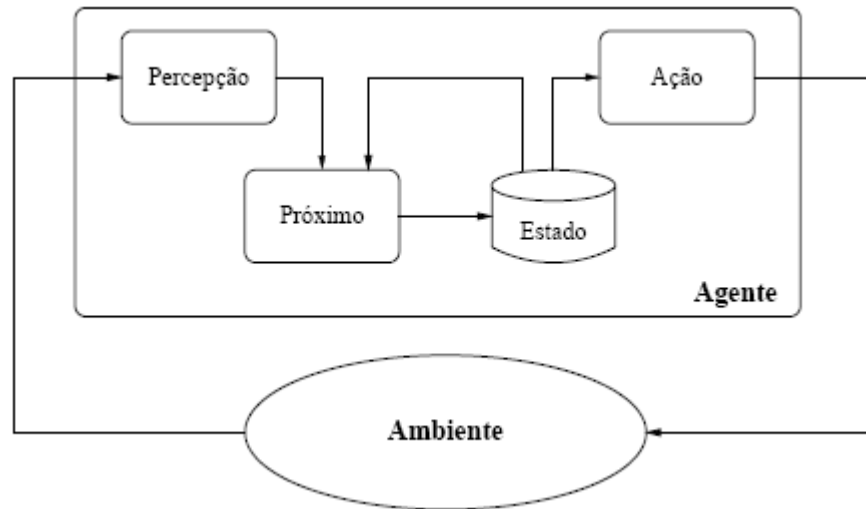


Figura 2.1: Esquema genérico para representar agentes autônomos proposto por [Wooldridge, 1999].

Na tabela 2.1 apresentam-se algumas das características das arquiteturas. Sendo que existem as arquiteturas híbridas que trazem características reativas e deliberativas. No entanto, existe um problema nestas arquiteturas: a definição de que característica deve prevalecer, e qual possui precedência sobre a outra [Sichman, 2005]. Este é um problema que envolve políticas de controle. No entanto, segundo este mesmo autor, a arquitetura híbrida é a que vem sendo mais estudada.

Tabela 2.1: Resumo comparativo entre arquiteturas multiagentes reativas e deliberativas [Sichman, 2005]

Deliberativas	Reativas
Representação explícita	Representação implícita
Possui histórico	Sem histórico (não possuem memória)
Comunicação direta	Comunicação indireta
Controle deliberativo	Controle não deliberativo
Organização social	Organização etológica
Poucos agentes (algumas dezenas, por exemplo)	Muitos agentes (algumas dezenas de milhares, por exemplo)

Alguns conceitos e definições úteis para área de sistemas multiagentes tratam de questões relativas a interação entre os agentes e sobre a representação interna dos agentes.

a) Cooperação entre agentes: uma das interações mais utilizadas em sistemas multiagentes é a cooperação. Um mecanismo de cooperação contém os seguintes itens: agrupamento, especialização, divisão de tarefas e resultados, coordenação de ações, resolução de conflitos, acordos, dentre outros. Um exemplo é uma rede contratual (pode-se pensar em uma analogia com leilões quando se fala neste protocolo), maiores detalhes podem ser vistos em [David, 1998].

b) Comunicação entre agentes: o primeiro aspecto da comunicação entre agentes é quando um agente decide-se comunicar com outros agentes. A teoria de atos de fala é uma proposta para este tema [Cohen e Levesque, 1990]. O segundo aspecto da comunicação refere-se à linguagem utilizada pelos agentes para a comunicação. Existem duas principais propostas utilizadas: KQML (resultado do projeto “*Knowledge Sharing Initiative*” financiado pelo DARPA) e a FIPA (“*Foundation for Intelligent Physical Agents*” – similar ao KQML). Ressalta-se que nestas propostas, o conceito de ontologias está implícito, pois elas são fundamentais para a comunicação e linguagem. O terceiro aspecto sobre a comunicação diz

respeito às mensagens trocadas entre os agentes, ou seja, o protocolo de interação entre os mesmos. Exemplos destes protocolos são: “*FIPA Request Protocol*”, “*FIPA Contract Net Protocol*” dentre outros.

c) Coordenação de agentes: a coordenação de agentes é, em essência, um problema de otimização [Lesser, 2005]. A coordenação é um processo de gerenciamento de dependências entre atividades. Ou seja, cada agente deve decidir que atividade deve fazer, como deve fazer e quando. Esta decisão, a busca de políticas ótimas, no seu caso mais geral é um problema de otimização combinatória. Segundo Lesser [Lesser, 2005] esta temática é nova na área de sistemas multiagentes – no relativo a políticas ótimas.

d) Modelos de interação social entre agentes: a figura 2.2 apresenta uma classificação que pode ser feita entre os modelos de interação [Sichman, 2005]. As interações “*Top down*” são caracterizadas por: cooperação pré-estabelecida, coordenação para atingir objetivo comum, subjetivismo social, dentre outros. Baseada na teoria organizacional, as interações “*Bottom-up*” estão centradas em utilidade, por exemplo, em teoria dos jogos e complementaridade, como a teoria das dependências.

Figura 2.2: Classificação de modelos de interação social entre agentes autônomos [Sichman,2005].

e) Raciocínio Social: o raciocínio social é baseado

aplicações nesta área, destacam-se: particionamento de tarefas entre robôs, transporte cooperativo entre robôs, particionamento e desenho de grafos, agrupamento e ordem de robôs, heurísticas para tratamento de problema como o “caixeiro viajante”, problemas de roteamento em redes (AntNet), problema de caminhos mais curtos dentre outros [Middendorf, 2005].

Outros conceitos são “clássicos” na área de sistemas multiagentes e remetem a itens como: planejamento, aprendizagem por reforço, redes neuronais dentre outros.

2.1.3 Aplicações e ferramentas computacionais

Algumas das áreas que estão utilizando agentes autônomos e sistemas multiagentes são [Jennings et. al, 1998]: manufatura, controle de processos, telecomunicações, controle de tráfego aéreo, sistemas de transportes, gerenciamento de informações (obtenção e seleção de informações), comércio eletrônico, gerência de processos empresariais, jogos eletrônicos, cinema interativo, acompanhamento de pacientes (área de saúde e medicina).

Alguns exemplos de aplicações implementadas são:

- “*Space Shuttle Malfunction (fault diagnosis) Handling*” – NASA [Malin et al., 2000];
- Gerenciamento da cadeia de suprimentos – Daimler Chrysler [Busmann e Sieverding, 2001];

Em [Pechoucek e Marik, 2006] é apresentado um compêndio com outras aplicações e utilizações industriais da tecnologia multiagentes, envolvem simulação e implementação.

O paradigma de sistemas multiagentes tem sido utilizado para implementações, conforme citado anteriormente. No entanto, este paradigma também tem sido utilizado na área da simulação, o que será discutido na próxima seção. As ferramentas computacionais que utilizam este paradigma irão diferir em seus propósitos, ou seja, algumas são orientadas para o desenvolvimento e implementação de sistemas computacionais e outras em simular utilizando a abordagem multiagente¹.

Alguns “frameworks” de programação utilizados no desenvolvimento e implementação de sistemas multiagentes [Coelho, 2005]: *Cougaar*², *Cybele*³, *JACK*⁴, *Jadex*⁵, *JAM*⁶, *Jason*⁷, *Quicksilver*⁸, *SIM_Agent*⁹, *Zeus*¹⁰, dentre outros. Um outro framework para implementação de sistemas multiagentes é o *JADE*¹¹ que é construído em java – “*Java Agent DEvelopment framework*”. Seu intuito é facilitar a implementação de sistemas multiagentes através de um “*middleware*” que implementa as especificações da FIPA. Possui um conjunto de ferramentas gráficas e é um projeto de software livre – código aberto.

Na simulação existem outras ferramentas e “frameworks”. Dentre estes “frameworks”, dois

1 Existem casos em que utiliza-se a simulação multiagente para testar sistemas desenvolvidos neste paradigma. Dado a complexidade destes sistemas a simulação torna-se uma ferramenta útil e flexível para testes.

2 <http://www.cougaar.org/>

3 <http://www.cybelepro.com/index.shtml>

4 <http://www.agent-software.com/>

5 <http://jade.tilab.com/>

6 <http://www.marcush.net/IRS/index.html>

7 <http://jason.sourceforge.net/>

8 <http://quicksilver.tigris.org/>

9 <http://www.cs.bham.ac.uk/research/projects/poplog/packages/simagent.html>

10 <http://sourceforge.net/projects/zeusagent>

11 <http://jade.tilab.com/>

se destacam em especial, foram analisados e considerados para desenvolvimento deste trabalho. Ambos possuem conceitos e esquema geral de funcionamento parecidos.

- *Swarm*: é um software que simula vida artificial, baseada em multiagentes. É mantido por uma comunidade científica. Possui funcionalidades como: modelagem espacial, visualizações de atividades em tempo real. É um projeto de software livre e foi feito pelo Instituto Santa Fé, no Novo México, utiliza linguagem Objective-C ou Java. Possui extensões como *EVO* e *MAML*. Será novamente discutido, visto que o modelo de simulação deste trabalho foi construído utilizando este software [Swarm, 2006].
- *Repast*: o acrônimo deriva do nome “*Recursive Porous Agent Simulation Toolkit*”. Possui alguns conceitos oriundos do *Swarm*, também tem o propósito de fazer modelos de simulação baseados em sistemas multiagentes. Uma diferença é que no *Repast* é possível o trabalho em várias linguagens de programação. É um projeto de código aberto da Universidade de Chicago. Possui três implementações: *Repast* para Java (*Repast J*), *Repast* para Microsoft .Net “framework” (*Repast.Net*) e *Repast* para Python Scripting (*Repast Py*) [Repast, 2005].

2.2 Modelos de simulação baseados em sistemas multiagentes

A simulação baseada em sistemas multiagentes é uma metodologia que simula experimentos construídos com um conjunto de agentes autônomos e que possuem em sua coletividade um conjunto de interações possíveis em um determinado ambiente, no intuito de replicar uma situação, um contexto real (a simulação baseada em sistemas multiagentes também envolve o uso de tecnologia de sistemas multiagentes seus conceitos, técnicas e ferramentas) [Sanchez e Lucas, 2002].

Sistemas que utilizam simulação baseadas em sistemas multiagentes possuem relacionamentos com outras áreas de conhecimento como: sistemas adaptativos complexos, gerenciamento e tomada de decisão e ciências sociais [Macal e North, 2005]. Estas áreas tem seu interesse no comportamento do indivíduo. Em essência a proposta desta abordagem de simulação está na representação de cada componente do sistema ou de cada agente, representando seus estados, necessidades, objetivos e decisões autônomas de acordo com seus interesses. Adicionando-se a isso a possibilidade de interações entre estes agentes de acordo com a representação interna de cada um. Esta abordagem pode ser dita como uma abordagem do tipo “*bottom-up*” em que o comportamento geral do sistema vai ser delineado pelas interações e comportamentos de seus agentes. Daí a interseção e interesse com as áreas de conhecimento supracitadas.

No Apêndice C – *Considerações sobre simulação* são apresentados alguns conceitos e considerações sobre a técnica de simulação computacional: características, vantagens e desvantagens. A simulação de eventos discretos e suas vantagens são bastante conhecidas, tanto academicamente quanto comercialmente. No entanto, para algumas situações esta técnica de simulação apresenta algumas dificuldades e problemas. As seguintes limitações podem ser citadas [Dubiel e Tsimhoni, 2005]:

- i) Movimentações de entidades no ambiente de simulação: dificuldade quanto a simulação de movimentos que não sejam predeterminados, ou seja, atribuir probabilidades para cada movimento que a entidade pode percorrer entre dois pontos, por exemplo. Este item é particularmente crítico quando se trata de simular movimentações humanas;

- ii) Capacidade das entidades de tomarem decisões em cada instante de tempo. Em um ambiente de simulação de eventos discretos existem pontos de decisão que devem ser especificados e uma lógica na tomada de decisão deve ser implementada para cada ponto. Para a tomada de decisão no tempo estes pontos devem ser distribuídos em um eixo temporal, o que pode aumentar muito o número de pontos, trazendo dificuldades para implementação e validação do modelo.
- iii) Incapacidade de que as entidades tomem decisões autônomas. As decisões das entidades são definidas previamente no modelo pelos usuários, o fluxo é pré-determinado. Entidades inteligentes não podem ser modeladas diretamente, o que novamente dificulta a implementação de entidades que representem comportamentos humanos.

Na simulação baseada em sistemas multiagentes aplicada nas ciências sociais, agentes representam pessoas ou grupos de pessoas e os relacionamentos entre grupos/pessoas são interações sociais [Macal e North, 2005]. Em essência, a prerrogativa é que, até determinado grau de abstração, estas interações entre os agentes podem ser definidas. Na perspectiva social esta premissa gera alguns questionamentos: a) Até que ponto o comportamento de seres humanos pode ser modelado (até que grau de abstração)? b) Quanto se sabe efetivamente sobre as interações sociais humanas para que as mesmas possam ser representadas? Estas e outras questões ainda permanecem em aberto, mas apesar disto importantes trabalhos vêm sendo desenvolvidos utilizando esta tecnologia, exemplos como: estudo sobre fatores culturais e sociais que levaram a extinção dos Anasazi nos Estados Unidos, a queda da civilização mesopotâmia, dentre outros [Kohler et al., 2005].

Esta dificuldade também foi encontrada quando da representação de agentes clientes (como representar a paciência de clientes quando estão aguardando atendimento na fila, por exemplo) e atendentes (quais preferências no trabalho e como isto impacta a produtividade, por exemplo) no modelo de simulação criado neste trabalho. Maiores detalhes serão apresentados nos capítulos 3 e 4.

Outros exemplos de estudos utilizando simulação baseada em agentes e ciência política, biológica e econômica podem ser encontrados em [Macal e North, 2005].

Alguns autores consideram a simulação baseada em sistemas multiagentes uma inovação na área de simulação capaz de endereçar questões até então não resolvidas. Outros acreditam que seja uma evolução natural dos paradigmas atualmente adotados como simulação de eventos discretos paralela, distribuída ou mesmo a simulação orientada a objetos [Davidsson, 2000].

Em [Sawhney et al., 2003] experimentos são apresentados e indicam que a abordagem de simulação multiagentes combinada com outras técnicas tradicionais de simulação adicionam grande flexibilidade na construção de modelos para sistemas considerados complexos. Em [Macal e North, 2005] é apresentada um resumo de áreas que estão utilizando simulação baseadas em sistemas multiagentes, vide tabela 2.2

Algumas questões foram propostas por Macal e North [Macal e North, 2005] e que deveriam ser respondidas durante a modelagem de uma simulação baseada em sistemas multiagentes:

- Na área de aplicação existem entidades que podem fazer uma representação natural com agentes?

- Existem decisões e comportamentos que podem ser definidos discretamente?
- Existem casos em que as entidades (agentes) da área em questão devem adaptar ou mudar seu comportamento?
- Existem casos em que os agentes têm relacionamentos dinâmicos com grupos ou outras entidades e estes relacionamentos são voláteis?
- Existem casos em que os agentes formam organizações e são capazes de aprender de forma coletiva?
- É importante que os agentes tenham um componente espacial para seus comportamentos e interações ?
- Para os agentes o passado não é, exclusivamente, um previsor do futuro (ou seja, condições do ambiente e interações podem influenciar nas decisões)?

As questões acima serão novamente citadas no capítulo 3 - *Modelo de Simulação de uma central de teletendimento* em que o modelo de simulação para centrais de teletendimento será apresentado. As respostas fornecidas às perguntas desta seção indicarão a adequação das centrais de teletendimento à simulação baseada em sistemas multiagentes.

Tabela 2.2: Resumo de áreas que utilizam simulação baseada em sistemas multiagentes [Macal e North, 2005]

Organização e Negócios	Biologia	Sociedade e Cultura	Infra-estrutura
- Indústria - Consumidores - Cadeia de suprimentos - Seguros	- Ecologia - Comportamento de agrupamento animal - Compartamento celular - Comportamento molecular	- Civilizações Antigas - Desobediência Social	- Transportes - Mercados de energia - Serviços públicos
Coletividade	Economia	Terrorismo	Militarismo
- Evacuação de áreas de risco - Movimentos Humanos	- Mercados financeiros artificiais - Redes comerciais	- Determinantes Sociais - Redes e organizações	- Comando - Controle

2.2.1 *Swarm* - um “framework” para simulação multiagente

O *Swarm* é uma plataforma de software livre para desenvolvimento de simulação baseadas em sistemas multiagentes. O projeto foi concebido em 1994 pelo Instituto Santa Fé, no Novo México. Seus objetivos principais são: facilitar a modelagem baseada em sistemas multiagentes para programadores inexperientes e estabelecer um padrão para desenvolvimento de simulação para simplificar o reuso e comparações entre modelos. A plataforma funciona sobre sistemas operacionais *Apple*, *Linux*, *Unix* e *Windows* (utilizando o *cygwin* um emulador *Linux* para a plataforma *Windows*). Possui também uma comunidade (SDG - “*Swarm Development Group*”) que disponibiliza modelos de exemplos, tutoriais, fóruns de discussões e organiza o processo de evolução da plataforma [Swarm, 2006].

Na figura 2.3 é apresentado um esquema geral simplificado de um modelo de simulação desenvolvido no *Swarm*. Este modelo está inserido em um ambiente, este ambiente por sua vez contém uma série de objetos (inanimados). Os agentes do modelo de simulação estão inseridos neste ambiente, suas interações são eventos (que estão distribuídos em um eixo temporal de execução formando uma “agenda” do agente) que estes agentes direcionam e executam de acordo com suas necessidades, interesses próprios e objetivos. Todas as

interações e comportamentos do sistema podem ser observados através de objetos específicos da plataforma. É possível também que o usuário do modelo interaja dinamicamente com a execução do modelo, fazendo alterações em tempo de execução. Estatísticas e indicadores de desempenho do sistema podem ser observados ou gerados durante ou após a finalização da execução do modelo.

As linguagens de programação utilizadas no desenvolvimento de modelos no *Swarm* são o *Objective C* (que possibilita uso de bibliotecas ou outros programas desenvolvidos em *C* ou *C++*) ou o *Java*. Os modelos se beneficiam da estrutura nativa destas linguagens quanto ao paradigma orientado a objetos. No *Swarm* também é possível criar hierarquias de simulações para que um modelo de simulação, por exemplo, possa conter outros modelos de simulação, ou seja, é possível representar agentes com elevado grau de complexidade. Alguns objetos são inerentes a plataforma e padronizam o desenvolvimento dos modelos de simulação, este é o caso para os objetos da classe “*Model*” que têm por intuito representar o modelo de simulação e seu controle de execução; “*Observer*” é responsável por apresentar estatísticas, indicadores ou mesmo apresentar o modelo de simulação de forma gráfica; “*Batch*” pode ser usado para executar, sem interações de usuário, o modelo de simulação e tem a capacidade de receber comando de entrada que pode ser passado para o modelo como variável de entrada.

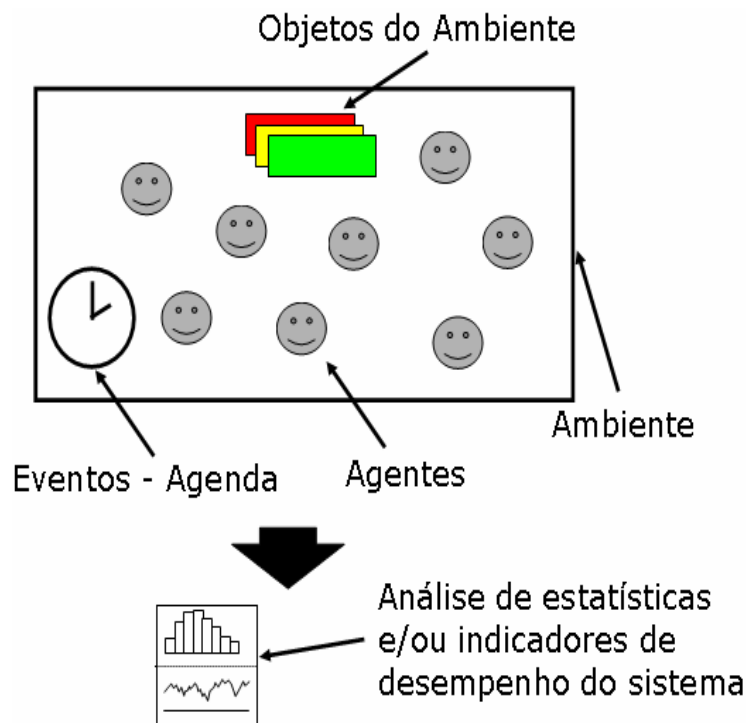


Figura 2.3: Esquema geral simplificado de modelo de simulação Swarm.

Além da abordagem multiagente, o software oferece outros recursos fundamentais para simulação: controle de estado atual do sistema (evidenciado pelo estado dos agentes), simulação de tempo, manipulação de eventos discretos, gerador de números pseudo-aleatórios, distribuições de probabilidade (*Poisson*, *gamma*, *lognormal*, *Bernoulli*, dentre outras), geração de arquivos em formato *.hdf*¹², criação e apresentação instantânea de vários tipos de gráficos, representação do ambiente de simulação em duas dimensões, apresentação do modelo durante o tempo de execução e outras bibliotecas. Maiores detalhes podem ser

12 HDF5 é um formato de armazenagem de dados do Centro Nacional para Aplicações de Supercomputadores, este formato pode ser carregado diretamente em softwares estatísticos como SAS e R.

encontrados em [Daniels, 1999] e [Johnson e Lancaster, 2006].

2.3 As centrais de teleatendimento

Uma central de teleatendimento, em essência, é um conjunto de recursos: hardware, software e pessoal para atendimento que está habilitado a prover serviços via telefone [Gans et al., 2006]. As centrais de teleatendimento são sistemas sócio-técnicos em que o comportamento dos clientes e dos atendentes são relacionados a indicadores que representam níveis de qualidade de serviço [Koole e Mandelbaum, 2002]. Outra conceituação é dada pela Resolução Normativa 57/2004 da Agência Nacional de Energia Elétrica – ANEEL que considera que uma central de teleatendimento é uma “unidade composta por estruturas física e de pessoal adequadas que tem por objetivo centralizar o recebimento de ligações telefônicas, distribuindo-as automaticamente aos atendentes, possibilitando o atendimento do solicitante”.

As centrais de teleatendimento têm ganhado cada vez mais espaço nas organizações como uma forma de contato ou relacionamento com os clientes. Este fato é evidenciado pelos itens destacados abaixo [Mandelbaum et al., 2001], [Sharp, 2003] e [<http://www.callcenter.inf.br>]:

- As Centrais de teleatendimento nos Estados Unidos empregavam cerca 1,55 milhões de pessoas em 1997, o que equivalia a 1,40% do setor privado;
- A AT&T registrou em um dia do ano de 1998 cerca de 260 milhões de chamadas, destas 40% eram destinadas a números do tipo 0800, o que pode indicar que boa parte destes 104 milhões de chamadas eram para centrais de teleatendimento;
- Estimativas indicam que cerca de 70% das interações com os clientes e empresas acontecem via centrais de teleatendimento;
- 700 bilhões de dólares em produtos e serviços foram vendidos através de centrais de teleatendimento em 1997;
- O número total de centrais de teleatendimento (internos ou externos as organizações) nos Estados Unidos em 2003 era estimada em cerca de 100.000;
- No mundo, as de centrais de teleatendimento renderam em 2005 US\$ 51,4 bilhões, e a previsão é que alcance US\$ 92 bilhões de receita em 2010.
- De acordo com relatório “*Brazil Call Center Services*” 2006 da IDC Brasil a receita proveniente das centrais no ano passado pelos países do continente americano totalizou US\$ 30,3 bilhões. Deste valor 5% foram referentes ao Brasil, quase o dobro do conquistado pelos demais países da América do Sul (que juntos somaram 3%). Na receita mundial o Brasil ficou com a fatia de 3,16%, alcançando o movimento de US\$ 1,6 bilhão em 2005.

Estes fatos apresentam a importância econômica das centrais. A necessidade de estabelecer e manipular um grande número de contatos com os clientes têm feito com que empresas busquem melhorar seus processos, serviços e tecnologia nas centrais de teleatendimento. Muitas empresas têm criado departamentos para lidar com esta temática: “*call centers*”, “*contact centers*” ou centro de interações com clientes [Sharp, 2003].

Por outro lado, o estudo das centrais de teleatendimento está concentrado em uma área de conhecimento denominada engenharia de serviços [Mandelbaum et al., 2001]. Segundo Mandelbaum a engenharia de serviços “busca desenvolver, com base científica, princípios e ferramentas (frequentemente culminando em softwares) que suportem e equilibrem qualidade

de serviço e eficiência sob perspectivas conflitantes de clientes, servidores, gerentes e, muito comumente, a sociedade”. A teoria de filas constitui-se como parte da base da engenharia de serviços para o desenvolvimento de tais princípios e ferramentas.

As centrais de teleatendimento podem ser modeladas como uma rede de filas [Whitt, 2002]. No entanto a complexidade inerente ao funcionamento das centrais de teleatendimento apresentam interessantes questões sob a perspectiva matemática [Mehrotra e Fama, 2003]:

- As centrais manipulam mais do que um tipo de chamada (o tipo de uma chamada pode ser definido pela natureza do serviço que está sendo solicitado. Por exemplo: uma compra, uma reclamação, um pedido dentre outros), ou seja, cada tipo de chamada pode ser tratado como uma fila com sua própria taxa de chegada e duração das chamadas;
- Através de equipamentos das centrais, lógicas/políticas sofisticadas de roteamento de chamadas podem ser avaliadas e escolhidas;
- Atendentes podem ser treinados para manipular uma chamada, vários tipos de chamadas ou todos os tipos de chamadas (todas estas particularidades devem ser consideradas no processo de roteamento das mesmas).

Nesta seção algumas das questões sobre o processo de tomada de decisão em centrais de teleatendimento e seus desafios serão apresentados. Os seguintes itens serão discutidos: o contexto das centrais de teleatendimento, seu processo na perspectiva interna e funcionamento; níveis, tomada de decisão e os desafios sob o qual os gerentes das centrais devem lidar no intuito de alcançar níveis de serviço e desempenho; trabalhos relacionados que remetem aos desafios gerenciais abrangendo modelos analíticos e modelos de simulação para suporte à decisão. Na última subseção buscou-se levantar o estado da arte sobre modelos de simulação em centrais de teleatendimento, após este ponto o modelo de simulação desenvolvido neste trabalho será apresentado.

2.3.1 Processo e funcionamento das centrais

As centrais de teleatendimento são utilizadas em vários setores da economia, em [Sharp, 2003] são apresentados estudos de casos das centrais no seguintes setores: comunicações, energia, serviços financeiros, governo, saúde, varejo, tecnologia e turismo. Além desta divisão por setor, as centrais podem também ser segmentadas por outras dimensões [Koole e Mandelbaum, 2002]: tamanho (número de atendentes, número de chamadas manipuladas por dia), distribuição geográfica (existem serviços de organizações que são prestados em vários lugares do mundo), funcionalidade quanto ao tipo de serviço prestado (de solicitação de segunda via de uma conta em uma concessionária de serviços públicos até um aconselhamento especializado em aplicações financeiras) dentre outros. Uma outra característica de uma central de teleatendimento é a capacidade dos atendentes realizarem chamadas (possibilitando que os atendentes sejam ativos gerando as chamadas para os clientes), por exemplo, no caso de uma central de vendas. Algumas centrais recebem chamadas, outras fazem chamadas e existem ainda centrais com as duas características, denominadas centrais mistas.

O advento das novas formas de comunicação providas pela Internet e outros meios de comunicação (*chat*, correio eletrônico, fax, mensagens por telefone celular dentre outros) tem possibilitado outras formas de interação entre o cliente e o provedor de serviços/produtos. É comum ouvir o termo “*Contact Center*”, centros de contato, para representar a extensão do

teleatendimento para outras mídias nas centrais [Whitt, 2002].

Neste trabalho o foco são as centrais de teleatendimento (envolvendo a telefonia como mídia exclusiva de acesso ao provedor de serviço/produto), com característica de chamadas entrantes.

Um exemplo de processo de funcionamento de uma central de teleatendimento com característica de chamadas entrantes esta representado na figura 2.4. Inicia-se com a necessidade de um cliente por um determinado produto/serviço, esta demanda gera uma chamada para a central. A partir desta chamada o cliente será atendido pelo ICT (Integração Computador Telefone) – que é um conjunto de equipamentos nos quais computadores e outros dispositivos de telefonia estão interligados. A URA – Unidade de Respostas Audíveis é um destes equipamentos, que é capaz de fornecer atendimento automático aos clientes para alguns tipos de serviços. Pode acontecer também do cliente não ser atendido, por exemplo, a chamada pode ser bloqueada pela central, visto que sua capacidade de atendimento pode ter sido excedida (considere uma central que detém 100 linhas para atendimento, pode acontecer que estas linhas estejam ocupadas no momento que o cliente originou a chamada). O bloqueio gera uma chamada não atendida. O processo continua caso o cliente necessite de atendimento específico (não disponibilizado pela URA), ou seja, será encaminhado para os atendentes, esta etapa de encaminhamento para os atendentes acontece em dois passos:

- i) Se os atendentes estiverem todos ocupados o cliente deverá aguardar em uma fila, até que alcance o instante de seu atendimento. Enquanto os clientes aguardam na fila, pode acontecer o que se denomina de abandono. Por questões pessoais o cliente decide abandonar a fila. O abandono gera uma chamada não atendida.
- ii) Caso haja algum atendente para realizar o atendimento ou algum atendente fique disponível e existam clientes na fila entra em ação um outro equipamento ICT, o DAC – Distribuidor Automático de Chamadas. É este equipamento que realiza o encaminhamento das chamadas dos clientes para os atendentes, ou seja, é ele que roteia as chamadas.

Uma vez que o cliente é encaminhado para um atendente, o atendimento será então consumado. Ressalta-se que toda toda essa dinâmica do processo de atendimento é acompanhada pelos gerentes das centrais, seu foco de interesse possui várias perspectivas quanto ao processo de tomada de decisão. Existem vários tipos de indicadores acompanhados pelos gestores (alguns inclusive com periodicidade horária), por exemplo: percentual de chamadas bloqueadas, percentual de chamadas abandonadas, número de chamadas não atendidas (bloqueadas ou derivadas de um abandono do cliente), nível de serviço prestado ao cliente, percentual de utilização dos atendentes, dentre outros. Este indicadores refletem o monitoramento da qualidade, produtividade e níveis de serviços da central.

Um outro fenômeno importante nas centrais é a religação, que acontece, por exemplo, em função de bloqueio da central (uma limitação de sua capacidade), de um abandono do cliente ou até mesmo um mau atendimento realizado por algum atendente e que gera uma insatisfação no cliente. Isto leva o cliente a ligar novamente para um novo serviço ou mesmo uma reclamação. A religação e o comportamento dos clientes sobre assunto podem impactar nos indicadores de desempenho das centrais de teleatendimento.

Figura 2.4: Um exemplo de processo de funcionamento de uma central de teleatendimento.

2.3.2 Indicadores de desempenho

O nível de serviço pode ser definido como “o grau de satisfação dos clientes com o serviço oferecido” [Koole, 2005]. Os indicadores das centrais para demonstração de nível de serviço consistem de diversos aspectos diferentes, alguns q

- TAb = tempo, em segundos, de espera do requisitante na fila antes de decidir abandonar a tentativa de acesso a central;
- TE = tempo, em segundos, entre o instante que o requisitante entra na fila e o instante em que o mesmo é atendido (chamadas abandonadas não consideradas neste item);

Com base nestas definições os indicadores de monitoração são apresentados (notação apresentada de acordo com a Resolução Normativa 57/2004 da ANEEL):

- Índice de nível de serviço básico:

$$\frac{\text{Total CA em até 30s}}{(\text{Total CA} + \text{Total CAb} + (\text{Total CO} \times K))} \times 100 \quad (1)$$

- Índice de Abandono: $\frac{\text{Total CAb}}{(\text{Total CRA})} \times 100 \quad (2)$

- Índice de chamadas ocupadas: $\frac{\text{Total CO}}{(\text{Total CO}f)} \times 100 \quad (3)$

- Tempo médio de abandono: $\frac{\sum_{k=0}^{CAb} TAb}{CAb} \quad (4)$

- Tempo médio de atendimento: $\frac{\sum_{k=0}^{CA} TA}{CA} \quad (5)$

- Tempo médio de espera: $\frac{\sum_{k=0}^{CE} TE}{CE} \quad (6)$

Outras definições de indicadores são apresentados em [Koole, 2005], assim como uma discussão sobre a utilização de médias para avaliação e monitoração das centrais. Koole apresenta a seguinte definição para o índice de nível de serviço que difere e pode ser comparada a Resolução Normativa 57/2004 da ANEEL (expressão adaptada de acordo com a notação adotada na Resolução Normativa 57/2004 da ANEEL):

$$\text{Índice de nível de serviço básico: } \frac{\text{Total CA em até } \left(\frac{\sum_{k=0}^{CE} TE}{CE}\right)}{\left(\text{Total CA} + \left(\text{Total CAb} > \frac{\sum_{k=0}^{CE} TE}{CE}\right)\right)} \times 100 \quad (7)$$

Em essência esta definição favorece o tempo médio de espera na central, contabilizando as chamadas e utilizando este tempo como um balizador para o nível de serviço. Outro indicador apresentado por Koole é um indicador com a visão interna da central e orientada para seu grau de eficiência. Este indicador é denominado de produtividade e está assim definido (expressão adaptada de acordo com a notação adotada na Resolução Normativa 57/2004 da ANEEL):

$$\text{Produtividade: } \frac{\text{Total tempo trabalhado}}{(\text{Total tempo disponibilizado para trabalho})} \times 100 \quad (8)$$

O numerador é definido como o tempo gasto em atendimento mais o tempo gasto após o atendimento. Já a definição do denominador é questionada por Koole, por exemplo: devem ser considerados tempos para descansos e/ou tempos para treinamentos no tempo total que o atendente disponibiliza para a central? Particularmente neste trabalho não se utilizou o termo produtividade, o entendimento da proposta de Koole foi tratada neste trabalho como percentual de utilização (ocupação) e não produtividade. No contexto deste trabalho o termo produtividade ganha outras conotações no sentido de valor agregado ou benefícios que certo grau de rendimento do atendente trás para a central, por exemplo: o rendimento de um atendente que atende 100 chamadas simples (1 minuto) e de clientes não prioritários para central não pode ser comparado ao rendimento de outro atendente que atende 10 chamadas (de 5 minutos) de clientes prioritários na central. A questão é que o tratamento do item produtividade pode seguir por um viés não trivial, quando se avalia o rendimento deste atendente para a central (leva-se em consideração número de chamadas atendidas, duração das chamadas, valor das chamadas para central ou mesmo outras variáveis). O modelo deste trabalho tratará percentual de utilização dos atendentes e não usará o termo produtividade.

Um resumo adaptado de [Saltzman e Mehrotra, 2004] sobre outras definições de indicadores, assim como outros indicadores pode ser visto na tabela 2.3.

Tabela 2.3: Resumo de outros indicadores utilizados nas centrais de teleatendimento e suas unidades de medidas, adaptado de [Saltzman e Mehrotra, 2004]

Indicadores	Unidades de Medida
Média de utilização Servidores	%
Média na fila	Clientes
Média atendimento	Clientes
Qualidade de Serviço – QoS	%
Tempo espera no sistema	Minutos ou segundos
Número de clientes atendidos	Clientes
Número de clientes abandono	Clientes
Número de clientes bloqueados	Clientes
Fração de abandono	%
Fração de bloqueio	%

Alguns trabalhos sobre a relação entre os indicadores das centrais podem ser encontrados em [Lopes et al., 2006], [L'Ecuyer e Avramidis, 2005], [Pichitlamken et al., 2003], [Koole, 2005], [Saltzman e Mehrotra, 2004], [Mazzuchi e Wallace, 2004]. Algumas relações (muitas delas empíricas) merecem menção e foram observadas nos trabalhos relatados:

- i) Tipicamente o percentual de abandono degrada o nível de serviço mas melhora o tempo médio de espera na fila, isto dependerá do comportamento de abandono dos clientes;
- ii) Para centrais de teleatendimentos bem dimensionadas, quanto ao número de atendentes (ou postos de atendimentos), o percentual de abandono torna-se algo pouco significativo para o nível de serviço;
- iii) Um grau maior de eficiência e qualidade de serviço é beneficiado pela idéia de “economia de escala”, ou seja, quanto maior o tamanho da central maior a eficiência (percentual de utilização dos atendentes) e qualidade (nível de serviço), este item, especificamente, é discutido em [Koole, 2005] assim como suas ressalvas e premissas – centrais “orientadas a qualidade e eficiência”;

- iv) Influências das variações de taxas de chegada das chamadas nos níveis de serviço;
- v) Combinações de várias habilidades ou competências de atendentes e seus relativos impactos nos indicadores.

2.3.3 Níveis de decisão e desafios gerenciais

As centrais de teleatendimento possuem um conjunto de recursos para prover seus serviços. Naturalmente, estes recursos são limitados. A distribuição e utilização dos recursos deve ser organizada pelos gestores de forma a encontrar um equilíbrio entre os indicadores de qualidade do serviço e as restrições da disponibilidade de recursos para manutenção ou expansão de capacidade. O foco da tomada de decisão está entre obter um retorno a partir de um determinado custo, satisfazendo os limites de níveis de serviço. Em [Koole, 2005] é apresentado o conceito de uma curva de eficiência que traduz esta relação entre custo e nível de serviço (aumentos de custos elevam o atingimento do nível serviço, por exemplo). Cada central possui sua própria curva de eficiência e esta curva vai variar em função de uma série de fatores como: tecnologia utilizada, infra-estrutura, processos adotados e serviços oferecidos. Ainda de acordo com Koole o processo de tomada de decisão que influencia diretamente estas variáveis, acontece em vários níveis:

- Estratégico - Tomadas pela alta gerência. Determinam qual o escopo da Central de teleatendimento na empresa, tipos de serviços prestados. Determina o orçamento a ser desembolsado na Central.
- Tático - Gerenciamento da Central. Busca responder questões como: utilização dos recursos, estrutura, organização (políticas de filas e roteamento das chamadas), contratação e treinamento de atendentes.
- Planejamento – São responsáveis pelo gerenciamento da força de trabalho (periodicidade semanal), programação, escalas e turnos dos atendentes.
- Diário - As variações e as situações de uma central de teleatendimento exige que decisões diárias sejam tomadas. Estas decisões são guiadas à luz da monitoração dos níveis de serviço e do percentual de utilização dos atendentes.
- Tempo real - Algumas decisões devem ser tomadas em tempo real. Diversos equipamentos de telecomunicação são capazes de tal feito, o DAC pode ser dado como um exemplo. Estas decisões envolvem algoritmos que implementam políticas complexas de operação.

Estes níveis representam um agrupamento das decisões (em função do foco da decisão) que devem ser tomadas nas centrais. Entre estes níveis de decisão encontram-se uma série de desafios com as quais os gestores devem lidar, todos eles relacionados a meta gerencial nas centrais: prover serviços com qualidade (aspecto quantitativo e qualitativo) sujeito a um orçamento pré-estabelecido para operacionalização da central. Uma destas dificuldades é referente as alterações abruptas nas demandas ou taxas de chegada das chamadas, sendo que que estas mudanças podem acontecer em um pequeno intervalo de tempo, muitas vezes um intervalo de horas. Assim as centrais necessitam manter um tempo de resposta pequeno e de acordo com a demanda [Betts et al., 2000]. Outra dificuldade é que a taxa de chegada e os tempos de atendimento podem variar em função do tipo de serviço, horários dentre outros.

O nível de serviço é degradado pelas chamadas não atendidas (vide figura 2.4) e uma das

razões que originam estas chamadas não atendidas é o comportamento de abandono dos clientes da central. Este comportamento pode variar, em [Zohar et al., 2002], por exemplo, é apresentado a diferença do comportamento de novos clientes e de clientes “experientes” o que configura outra dificuldade para os níveis de decisão. Existem vários fatores que estão associados com o ambiente interno e a rotina de trabalho nas centrais de teleatendimento, dentre os quais se pode citar [Derry et al., 2002]: número de interações com clientes, elevada carga de trabalho, trabalho repetitivo e rotineiro. Estes fatores levam ao absenteísmo nas centrais. Este absenteísmo deve ser considerado, pois pode influenciar os níveis de serviços.

historicamente deram resultados positivos). Por exemplo, solicitar a um grupo de atendentes que dobre o turno, inserir mensagem automática no sistema para avisar clientes sobre algum tipo de situação dentre outros. Esta forma de gerenciamento reativa não é adequada para as centrais, que devem buscar um gerenciamento mais proativo e que seja capaz de lidar com sua natureza e desafios. Neste intuito algumas centrais utilizam sistemas denominados: gerenciamento da força de trabalho. Estes sistemas auxiliam os gestores em seu processo de tomada de decisão, mas apenas 10 % das centrais utilizam este tipo de tecnologia [Sharp, 2003].

Modelos simbólicos estocásticos têm sido empregados no auxílio à tomada de decisão nas centrais de teleatendimento [Koole e Mandelbaum, 2002], [Gans et al., 2006], [Koole, 2005], [Whitt, 2002]. Este modelos são derivados da teoria de filas e consideram que as centrais de teleatendimento podem ser modeladas como uma rede de filas. Existem modelos analíticos e de simulação utilizados neste problema. Estes modelos também são o “coração” dos sistemas de gerenciamento da força de trabalho. Nesta seção alguns trabalhos a respeito destes modelos analíticos serão apresentados.

Uma notação utilizada para os modelos de fila é a notação de *Kendall*, [Hillier e Lieberman, 2001]:

$$A / B / s / k.$$

em que A representa a distribuição da taxa de chegada na fila. Já B representa a distribuição para o tempo de atendimento. Tanto A quanto B podem assumir as seguintes siglas: M -distribuição exponencial, markoviana; D -distribuição determinística, constante; E_k -distribuição de Erlang; G -distribuição genérica, na qual se conhece a média e variância, por exemplo. O terceiro parâmetro da notação de *Kendall*, s , representa o número total de servidores (ou no caso das centrais o número de atendentes). E k representa o tamanho máximo que fila poderá comportar (no caso das centrais o número máximo de clientes que poderão ficar esperando por um atendente).

Um dos modelos de fila mais utilizados nas centrais de teleatendimento é o modelo “*Erlang C*” proposto por Erlang [Erlang, 1917]. Este modelo apresenta as seguintes características:

- taxa de chegada com distribuição de Poisson;
- tempo de serviço com distribuição exponencial;
- múltiplos servidores em paralelo recebendo entradas advindas de uma única fila;
- servidores homogêneos, ou seja, são idênticos, mesma média de tempo de atendimento;
- atendimento na fila é do tipo “primeiro a entrar primeiro a sair”, e a fila possui tamanho infinito;
- não há abandonos na fila, ou seja, o requisitante do serviço espera na fila infinitamente até o seu atendimento.

Em resumo este modelo só admite um tipo de chamada, não considera bloqueios (número máximo de posições na fila) e desconsidera a questão de abandonos dos clientes durante a espera por atendimento na fila. O modelo é capaz de obter [Koole, 2005]:

- Probabilidade de atraso:

$$C(s, a) = \frac{a^s}{(s-1)!(s-a)} \left[\sum_{j=0}^{s-1} \frac{a^j}{j!} + \frac{a^s}{(s-1)!(s-a)} \right]^{-1}, \forall s > a \quad (9)$$

Onde:

λ é a taxa de chegada;

β é a tempo de atendimento,

a representa a carga do sistema, $a = \lambda \beta$, medida em *Erlangs*, e

s é o número de servidores.

- Média de espera para atendimento: $W = \frac{C(s, a)\beta}{(s-a)} \quad (10)$

- Nível de Serviço: $\delta = 1 - C(s, a)e^{-\frac{(s-a)\tau}{\beta}} \quad (11)$

Onde:

τ representa o tempo aceitável de espera na fila (o tempo de 30 segundos pode ser considerado como um tempo para que indique um bom nível de serviço, por exemplo).

- Percentual de Utilização: $v = \frac{\lambda\beta}{s} \quad (12)$

É comum encontrar na literatura a seguinte notação de $C(s, a) = \frac{a^s}{(s-1)!(s-a)} \left[\sum_{j=0}^{s-1} \frac{a^j}{j!} + \frac{a^s}{(s-1)!(s-a)} \right]^{-1}$

encontrar o número mínimo de atendentes a ser estipulado no funcionamento da central [L'Ecuycr e Avramidis, 2005].

- Tipos de chamadas: o tempo de atendimento das chamadas segue uma distribuição exponencial. Existem trabalhos que ressaltam que este fato não se estabeleceu na realidade [Kooile, 2005]. Uma das principais motivações para isto é que estas centrais forneciam mais do que um tipo de serviço e estes tipos apresentam médias de tempo atendimento diferentes.
- Variação temporal: a figura 2.5 representa um gráfico com a variação do número de chamadas por hora do dia em uma central de teleatendimento, estes dados foram obtidos a partir de uma visita realizada em uma central durante as investigações feitas neste trabalho. A figura reflete os valores de uma quarta-feira de novembro de 2006 [Valladares, 2007]. A taxa de chegada das chamadas é dependente da dimensão tempo e pode sofrer alterações em função de eventos que são capazes de alterar a demanda de chamadas [L'Ecuyer e Avramidis, 2005]. As curvas de frequência de chamadas são diferentes entre os dias, meses e anos.

2.3.5 Simulação em centrais de teleatendimento

As limitações dos modelos do tipo “Erlang” tem motivado uma série de trabalhos e pesquisas sobre o assunto. As incertezas e complexi

simulação de centrais de teleatendimento, assim como as vantagens, potencialidades e capacidade de resolução de problemas dos gestores utilizando técnicas de simulação para problemas de decisão existentes nas centrais. A figura 2.6 apresenta um diagrama adaptado que sintetiza os diversos aspectos que devem ser levados em consideração em modelos de simulação de centrais de teleatendimento.

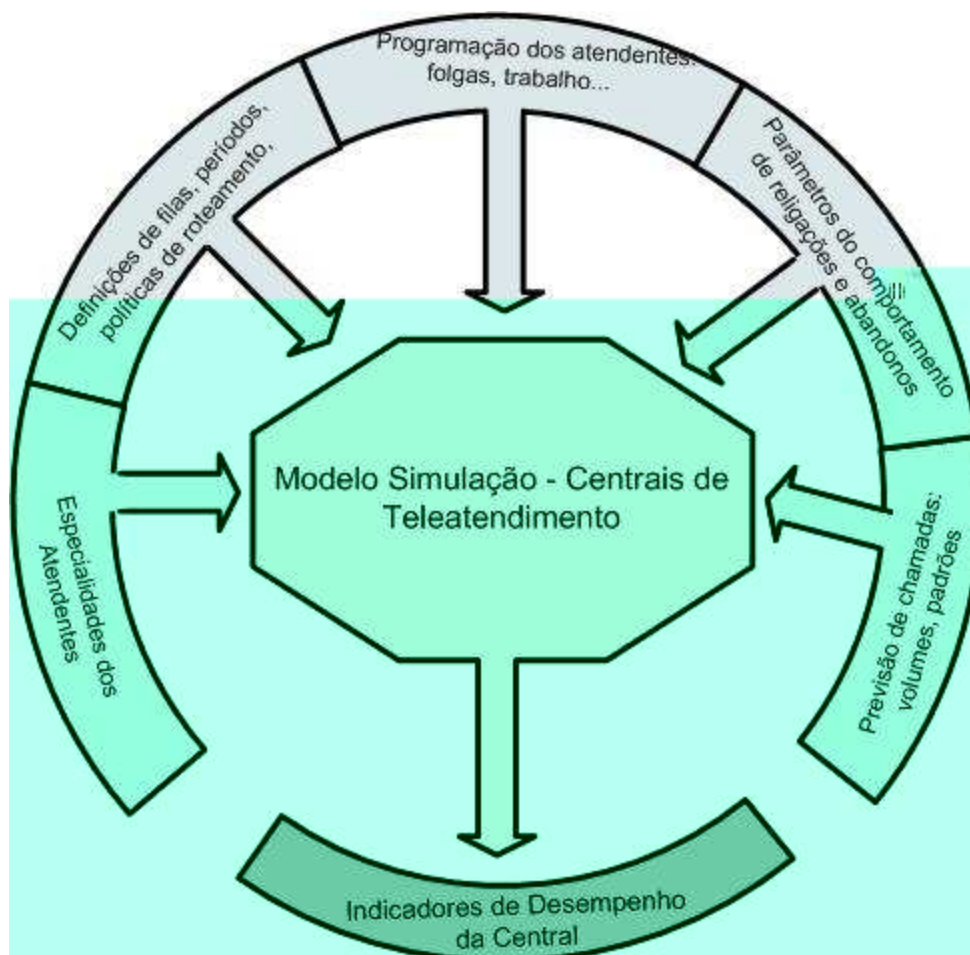


Figura 2.6 Esquema de simulação em centrais de teleatendimento, adaptado [Mehrotra e Fama, 2003].

Em [Mazzuchi e Wallace, 2004] e, posteriormente, em [Wallace e Saltzman, 2005] há a descrição de um modelo de simulação de eventos discretos desenvolvido na linguagem C. Este modelo é capaz de simular o roteamento baseado em habilidades. No último artigo, a ferramenta é comparada com um outro modelo desenvolvido na ferramenta de simulação ARENA, vantagens e desvantagens das duas abordagens são comentadas.

Em [Saltzman e Mehrotra, 2004] há um modelo de simulação utilizando o ARENA, no entanto toda a interface é feita utilizando-se planilhas eletrônicas. O intuito é apresentar uma interface simples para que as análises tipo “e se” possam ser facilmente utilizadas pelos gestores sem conhecimento detalhado sobre o modelo criado. O modelo também é utilizado para responder de forma empírica a uma série de situações propostas.

Um estudo de caso sobre uma central de teleatendimento de uma companhia distribuidora de gás é analisado em [Takakuwa e Okada, 2005]. O modelo de simulação neste trabalho foi utilizado em conjunto com técnicas de busca para encontrar o número ótimo de atendentes considerando restrições de agendas e habilidades. Inicialmente um ponto viável de solução foi encontrado utilizando programação inteira para posterior utilização de simulação e técnicas de

busca. Os autores citam que a solução proposta foi realmente utilizada na prática, demonstrando sua eficiência.

3 Modelo de Simulação de uma central de teleatendimento

O modelo de simulação proposto neste trabalho representa as centrais de teleatendimento utilizando um paradigma baseado em sistemas multiagentes. Este paradigma aplicado a central representa atendentes, clientes, equipamentos e gestores como agentes e objetos que estão inseridos no ambiente da simulação. Desta forma, o modelo representa de forma única cada atendente, cada cliente, cada equipamento, cada gestor e suas interações. Os agentes tomam suas decisões, interagem com outros agentes e objetos, expressam seus interesses particulares e necessidades específicas de acordo com seus objetivos. Estes comportamentos e interações foram modelados em um nível granular: cada agente. O universo de agentes e objetos com suas características configuram-se como os componentes do modelo para simular o funcionamento da central. Estes componentes e suas propriedades existem na simulação com a possibilidade de representar situações reais que acontecem na operação e apresentar o comportamento geral da operação e funcionamento da central de teleatendimento. Esta dinâmica é demonstrada fornecendo indicadores sobre o desempenho do sistema de funcionamento da central.

Esta forma de modelagem apresenta desafios que foram encontrados quanto a modelagem dos agentes; por exemplo, no que se refere a simular comportamento humano (“paciência” dos clientes, preferências de atendentes, dentre outros). Esta abordagem de modelagem granular mínima faz com que o funcionamento geral do sistema convirja para um comportamento que se aproxima da realidade. Apesar das dificuldades que foram encontradas o modelo apresentou resultados consistentes: quando comparado com a realidade e quando comparado com modelos analíticos “clássicos”.

O objetivo do modelo de simulação proposto é defender a viabilidade da abordagem baseada em sistemas multiagentes para simulação das centrais de teleatendimento; averiguar, pela utilização do modelo e de forma empírica (experimentando diferentes cenários partindo das clássicas perguntas “e se”) a correlação entre indicadores de desempenho da central; e mostrar que esta abordagem oferece soluções para problemas que os modelos “clássicos” não abordam.

A base motivadora do modelo de simulação está concentrada nas seguintes considerações:

- I) Conceituação apresentada em [Koole e Mandelbaum, 2002] se referindo as centrais como sistemas sócios-técnicos em que o comportamento dos clientes e atendentes (que são modelados como agentes) são relacionados a indicadores que representam níveis de qualidade de serviço;
- II) Em [Whitt, 2002] é ressaltado que a simulação utilizada para as centrais é uma abordagem atrativa para prever e controlar os indicadores de desempenho, mas que possui alguns desafios para ser mais efetiva. Estes desafios são: analisar eficientemente sistemas complexos (como as centrais), auxiliar o entendimento estatístico dos resultados dos experimentos e melhoria para suporte aos gestores na realização dos experimentos (análises do tipo “e se”);
- III) Alguns autores (ver *subseção 2.3.4 - Modelos para auxílio a tomada de decisão “clássicos”* e *subseção 2.3.5 - Simulação em centrais de teleatendimento*) ressaltam a

dificuldade para representar o comportamento humano nos modelos¹³.

A aplicação do modelo de simulação baseado em sistemas multiagentes aplicada nas centrais de teleatendimento possui como subsídio as seguintes respostas para as questões apresentadas em [Macal e North, 2005] quando da intenção de utilização desta abordagem em alguma área de aplicação (ver seção 2.2 - *Modelos de simulação baseados em sistemas multiagentes*):

- Na área de aplicação existem entidades que podem fazer uma representação natural com agentes?

Para definir a central de teleatendimento [Koole e Mandelbaum, 2002] o termo “socio-técnicos” é utilizado e o comportamento dos clientes e atendentes são destacados e relacionados a indicadores que representam níveis de qualidade de serviço. Atendentes, clientes e gerentes das centrais podem ser modelados como agentes autônomos vistos seus interesses, necessidades, objetivos e formas de atuação distintos no ambiente de simulação. Além destes agentes acredita-se que pela complexidade de operação e capacidade autônoma o DAC, o equipamento que roteia chamadas dos clientes para os atendentes nas centrais, também pode ser visto como um agente, visto que o mesmo poderá alterar seu comportamento de acordo com as metas de indicadores de desempenho estabelecidos na central durante sua operação.

- Existem decisões e comportamentos que podem ser definidos discretamente?

Sim, todos os comportamentos e decisões dos agentes modelados para as centrais estão orientados pelo eixo temporal e as decisões e comportamentos também estão relacionados a um universo discreto de opções (a decisão de um agente cliente sobre abandonar ou não a fila, por exemplo).

- Existem casos em que as entidades (agentes) da área em questão devem adaptar ou mudar seu comportamento?

Em [Mandelbaum et al., 2001] é apresentado estudo sobre uma base de dados com cerca de 450.000 chamadas que representa a quantidade de chamadas das centrais durante o ano de 1999, esta base de dados é real, pertencente a um instituição financeira do Estado de Israel. Este trabalho recebeu autorização do autor para que a base pudesse ser utilizada [Mandelbaum, 2006]. Uma análise estatística de dados foi realizada nesta base e foi constatado elevado percentual de clientes que aumentavam sua paciência de espera na fila levando em consideração um determinado dia, ou seja, em um determinado dia se o cliente abandona a fila e volta a ligar no mesmo dia este cliente torna-se mais “paciente” neste dia. O que já não se pode afirmar em casos que o cliente liga em dias diferentes (este assunto será novamente discutido no capítulo 4 na subseção 4.4.1 – *Simulação do abandono e o comportamento de espera dos clientes*). Este caso é apenas um exemplo da adaptação de comportamento dos agentes clientes.

- Existem casos em que os agentes têm relacionamentos dinâmicos com grupos ou outras entidades e estes relacionamentos são voláteis?

13 Anteriormente neste trabalho foi apresentado algumas dificuldades que as ciências sociais encontram na representação de comportamento humano em modelos de simulação, no entanto interessantes resultados tem sido obtidos com a utilização de técnicas de simulação baseadas em sistemas multiagentes.

Conforme descrito na *subseção 2.3.1 - Processo e funcionamento das centrais*, o processo da central é dinâmico, assim como o relacionamento entre atendentes e clientes durante o atendimento da demanda. Outra característica é a variabilidade de demandas e comportamentos dos agentes como: abandonos, absenteísmo, eficiência, reLigações, rotatividade de pessoal dentre outras.

- Existem casos em que os agentes formam organizações e são capazes de aprender de forma coletiva?

A resposta para esta questão está além do escopo deste trabalho. No entanto acredita-se que haja indícios sobre efeitos de atuação de formas coletivas relacionados aos seguintes eventos na central: absenteísmo dos atendentes (evento externo coletivo/grupos, que leva a uma ausência), aumento na taxa de demanda (anúncio de uma grande promoção em uma mídia de comunicação, falta de luz em uma região ou mesmo troca de informações coletivas sobre a qualidade do serviço da central) dentre outros. Acredita-se que estudos de caso e pesquisas de campo devam ser feitas para esclarecer este item.

- É importante que os agentes tenham um componente espacial para seus comportamentos e interações?

No caso da simulação e agentes na central, este requisito é substituído por um componente temporal, que armazena uma “espécie” de memória, consciência do tempo. O agente atendente, por exemplo, detém o conhecimento do total trabalhado por unidade de tempo (dia, por exemplo, que poderia ajudar a representar seu grau de “cansaço” naquele dia) e o tempo total (o que pode representar “experiência” ou eficiência). Assim como o agente cliente possui a informação sobre o tempo que o mesmo aguarda atendimento na fila.

- Para os agentes o passado não é, exclusivamente, um previsor do futuro (ou seja, condições do ambiente e interações podem influenciar nas decisões)?

A adaptação do comportamento de paciência de clientes em um mesmo dia pode ser citado como um caso. Por exemplo, o passado representado por uma média de tempo de espera pode ser ignorada em uma situação de emergência do cliente, ou seja, apesar de um histórico com pouca paciência, este pode ser obrigado a ter “paciência” no caso de uma força maior.

Dotados desta argumentação e subsídios utilizou-se o modelo de simulação baseado em sistemas multiagentes como paradigma para simular as redes de filas e o *Swarm* como um “framework” de simulação multiagentes a fim de prover informações gerenciais e fornecer suporte para os níveis de decisão gerencial (planejamento e diário) de centrais de teleatendimento.

O modelo possui vários parâmetros de entrada que alteram o ambiente e influenciam o comportamento dos agentes. Esta dinâmica é apresentada com o uso de monitoração dos indicadores de nível de serviço para suportar o processo de tomada de decisão nos níveis táticos e de planejamento. Uma abordagem semelhante não foi encontrada na literatura com foco em simulação (já na abordagem de implementação em sistemas multiagentes alguns trabalhos foram encontrados [Brazier et al., 1999], [Enoki et al.,2001]).

Este capítulo apresenta o modelo de simulação. Inicialmente é feita a descrição do modelo. Na seção seguinte são apresentados mais detalhadamente alguns agentes contidos no modelo,

para tal uma representação em diagramas de classes do modelo será mostrada. O foco é detalhar aspectos relativos à implementação de alguns agentes, suas possibilidades e capacidades de extensões futuras. A última seção apresenta os parâmetros de entrada e os indicadores de desempenho gerados.

3.1 Descrição e apresentação

A figura 3.1 apresenta um esboço conceitual do modelo de simulação que é composto por agentes, interações e objetos. Esta figura foi o ponto de partida para a implementação do modelo. Ela foi adaptada a partir da descrição de funcionamento geral das centrais de teletendimento, que foi discutida utilizando a figura 2.4.

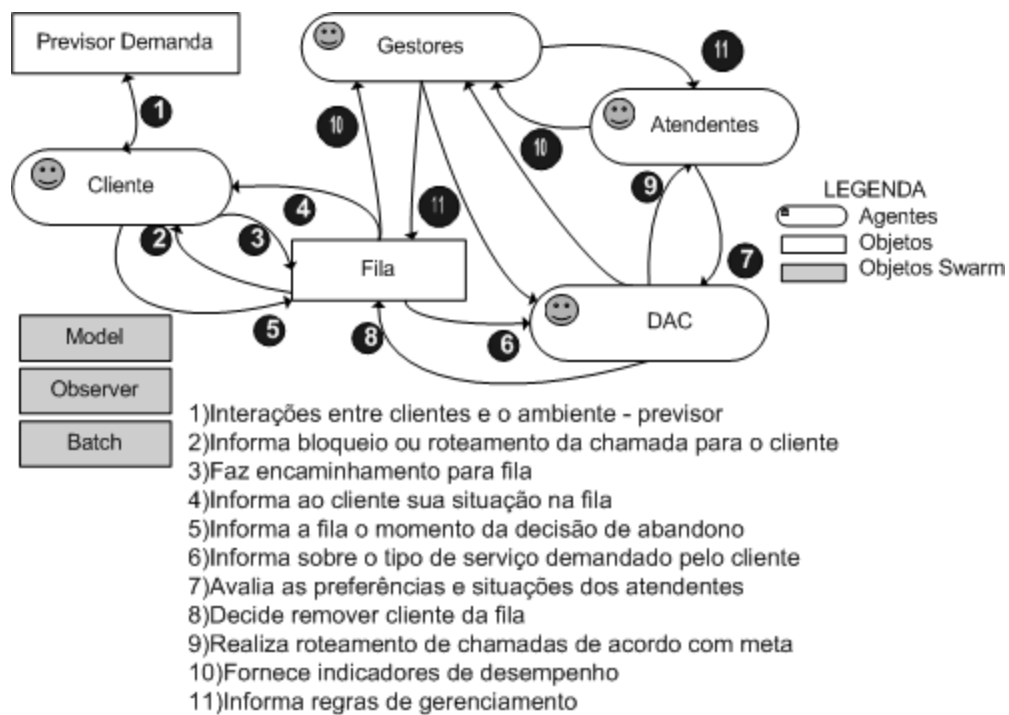


Figura 3.1 Descrição conceitual do modelo com os agentes, interações e objetos da simulação para as centrais de teletendimento

Inicialmente um conjunto de agentes do tipo cliente são representados e criados no ambiente de simulação, em princípio estes clientes não possuem nenhum histórico de atendimento (no entanto, seria possível que a criação destes agentes acontecesse com o histórico de interações que o mesmo já realizou com a central). Com os clientes criados os mesmos devem decidir sobre o instante em que vão fazer chamadas para a central de teletendimento. Este papel de gerar chamadas é feito entre o agente cliente e o objeto previsor de demandas (vide a etapa A da figura 3.2). Uma vez realizada a chamada a mesma será encaminhada para a central. Os agentes atendentes, DAC, gestor e o objeto fila também existem, desde o início do modelo de simulação. Da mesma forma que o agente cliente, estes agentes são criados sem histórico. Com o início do processo de atendimento pela central pode acontecer do cliente ser informado de um bloqueio, ou ser encaminhado direto para o atendente via DAC ou ser informado que deverá aguardar na fila (vide a etapa B da figura 3.2). No caso em que o cliente aguarda na fila pode acontecer do cliente decidir abandonar a fila. Se isto não acontecer e o cliente aguardar na fila, este objeto pode passar algumas informações (vide a etapa C da figura 3.2) para o agente cliente. Esta troca de informação

pode ocasionar, por exemplo, uma reavaliação de situação e tomada de decisão, por parte do agente cliente. Enquanto existem clientes na fila ou há demanda chegando para a central o objeto fila repassa estas informações para o DAC que conhece e verifica a situação dos atendentes (vide a etapa D da figura 3.2). Se for propício este agente solicita ao objeto fila um cliente para roteamento da chamada (vide a etapa E da figura 3.2).

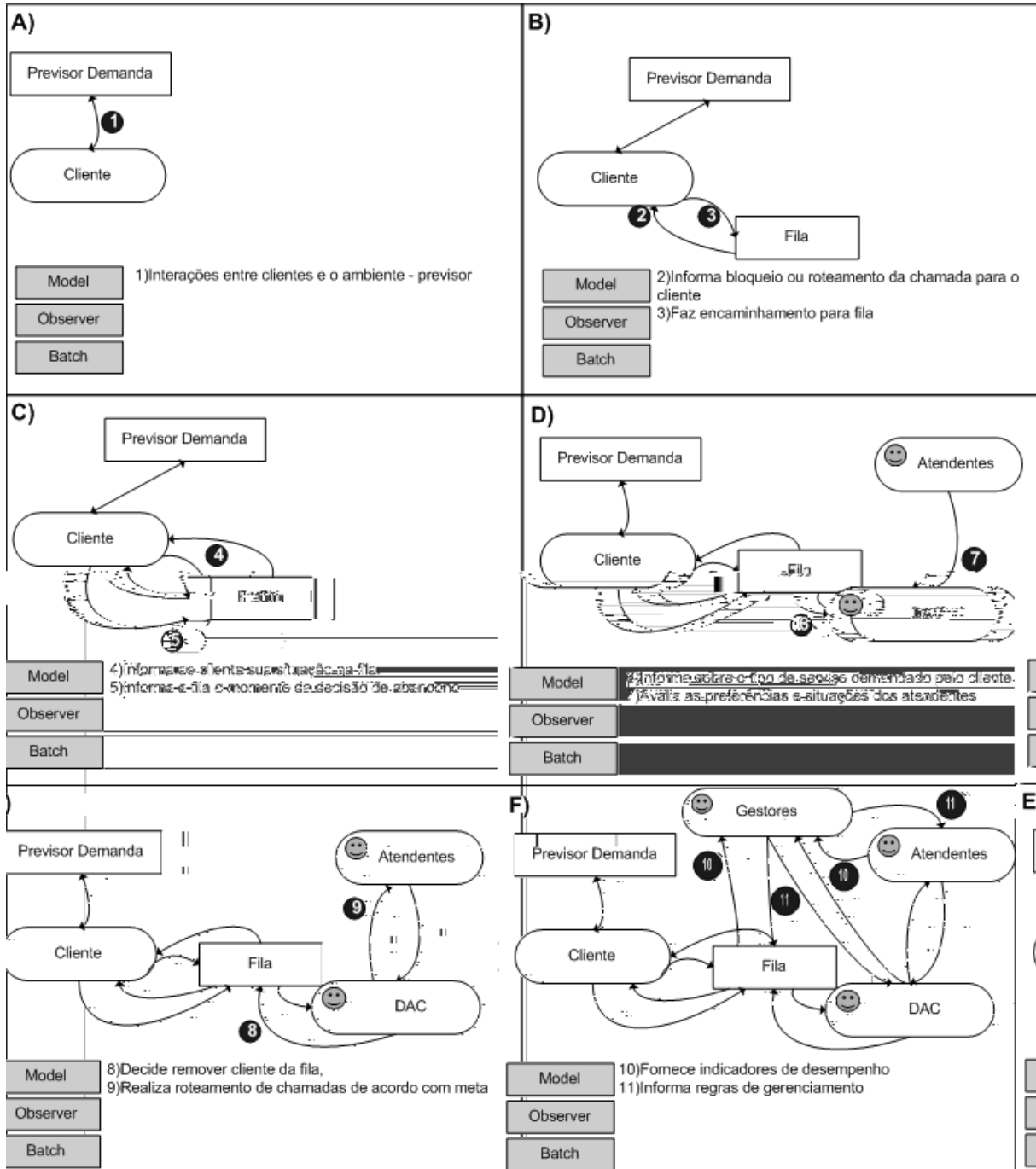


Figura 3.2 Descrição do modelo de simulação em etapas. As etapas de A até F representam a evolução gradual na simulação das centrais de teleatendimento.

Esta dinâmica de funcionamento no modelo envolvendo agentes e objetos é monitorada pelo agente gestor. Isto é feito em tempo “real”. Este agente avalia a situação do sistema geral da central de teleatendimento e pode, de acordo com a meta (aumentar o nível de serviço

ou reduzir o custo com ligações pagas pela central, por exemplo) informar preferências ou regras de funcionamento para os demais objetos e agentes (vide a etapa *F* da figura 3.2). A figura 3.3 apresenta os estados sob a perspectiva das chamadas no modelo de simulação, objetos e agentes estão envolvidos no tratamento da chamada e seus estados.

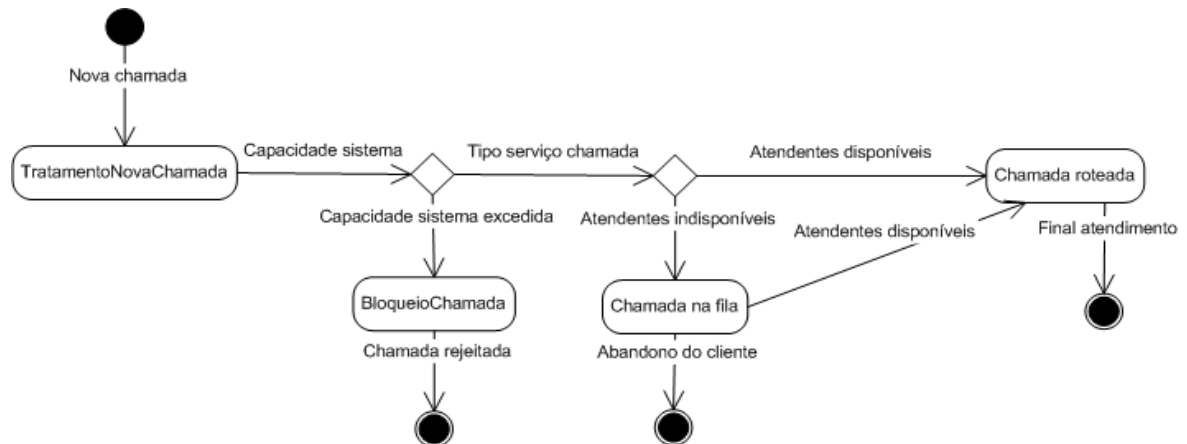


Figura 3.3 Diagrama de estados sob a perspectiva das chamadas

Os objetos representam o ambiente de simulação e interagem com os agentes. O objeto preditor de demanda interage com os agentes clientes no intuito de prover informações para a geração de chamadas. Este objeto está preparado para fornecer a taxa de chegada (com distribuição de Poisson, por exemplo), ou mesmo outros tipos de distribuição. Outra possibilidade é que este objeto possa fornecer aos clientes dados das demandas a partir de um arquivo, por exemplo, a lista de chamadas de uma central real. Os objetos *Model*, *Observer* e *Batch* são objetos do “framework” e cada um deles possui uma funcionalidade específica no modelo de simulação: *Model* – controla todos os eventos e comunicações do ambiente, é o modelo de simulação em si, manipula e controla o tempo total da simulação. *Observer*¹⁴ e *Batch* – são dois objetos distintos que tem por objetivo, respectivamente: apresentar visualmente a execução do modelo e descrever resultados; e processar (sem visualizações) e descrever os resultados finais.

Outro objeto que desempenha papel fundamental no modelo é a fila, ou na verdade, a tele-fila. É o ponto inicial de contato, “ponte” dos clientes com os atendentes para efetivação do serviço (esta interação ainda passa pelo DAC). A fila é constituída de posições que são ocupadas pelos clientes. Este objeto possui vários tipos de tratamentos que podem ser dados aos clientes, como por exemplo, referente a sua política de organização: primeiro cliente a chegar primeiro cliente a ser atendido ou uma política de organização por prioridades dos clientes. No modelo de simulação o agente gestor recebe informações do objeto fila sobre indicadores de controle e monitoração do desempenho da central.

Além dos objetos há alguns tipos de agentes, parte autônômica, no modelo de simulação. São quatro os tipos no modelo: agentes clientes, atendentes, DAC e gestor.

Agente Cliente

O tipo de agente cliente foi idealizado para representar os clientes que geram chamadas para as centrais e seus comportamentos quanto à interação com a central. Este agente interage com um outro objeto denominado preditor demanda que é responsável por motivar o ambiente para os momentos da demanda. Os clientes tratam sua experiência na fila como uma

¹⁴ Neste trabalho este objeto não foi explorado, apenas algumas visualizações padrões do próprio “framework” estão disponíveis. Para execução de experimentos foi utilizado o objeto Batch.

parte do serviço provido pela central e fazem julgamentos próprios sobre os serviços prestados pela central. Da mesma forma que a central tem indicadores, cada cliente pode ter sua própria definição sobre o grau de qualidade dos serviços da central. Esta informação assim como outras podem ser utilizadas para a tomada de decisão de um determinado agente. Por exemplo, um cliente deve ou não realizar chamadas, se ocorrido um bloqueio ou abandono, faz ou não uma religação; ou mesmo a decisão do momento de abandono. Todas estas decisões são distribuídas entre os vários agentes clientes. Veja figura 3.4 que representa estados sob a perspectiva do agente cliente.

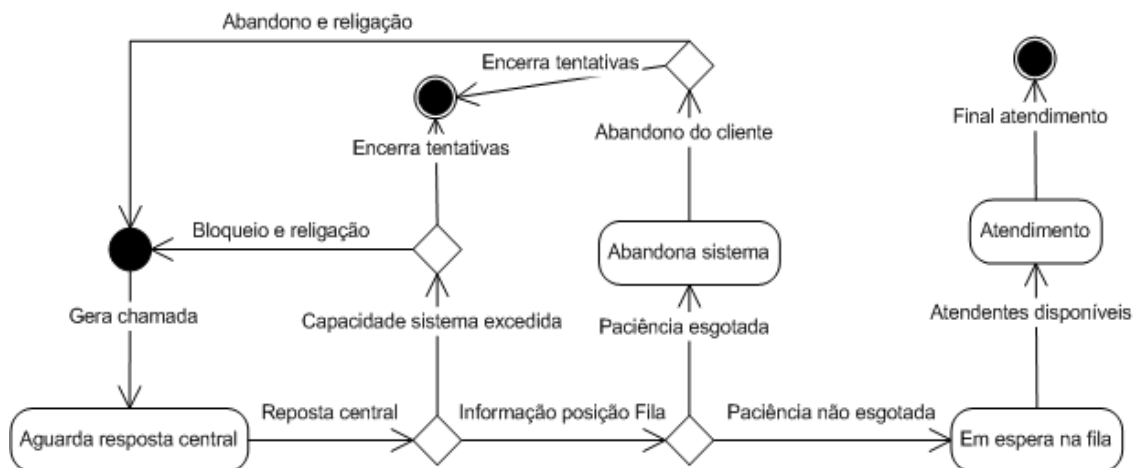


Figura 3.4 Diagrama de estados sob a perspectiva do agente cliente

Agente Atendente

O tipo de agente atendente realiza os atendimentos e é responsável por receber uma chamada encaminhada pelo tipo de agente DAC e, posteriormente, comunicar ao mesmo o final de um atendimento. Possui o conhecimento sobre sua eficiência e utilização: quantidade de chamadas e tempo trabalhado por unidade de tempo (um dia, por exemplo) ou tempo total desde sua admissão na central. Existem opções particulares de cada atendente: escala de trabalho, habilidades para cada um dos tipos de serviços, preferências para cada tipo de chamada, dentre outros.

Agente DAC

O DAC, distribuidor automático de chamadas, é responsável por encaminhar as chamadas dos clientes para os atendentes, ou seja, ele recebe demandas dos clientes que podem ou não estar na fila. Os atendentes interagem com o distribuidor para comunicar o final de um atendimento. É neste agente também que há memória sobre as habilidades, preferências e eficiência dos atendentes para cada tipo de serviço. Dotado desta memória a respeito das habilidades e do conhecimento sobre atendentes ocupados e não ocupados em determinado momento, este agente faz a atribuição de clientes e atendentes, baseado em uma série de critérios de acordo com as metas da central, por exemplo, existem várias políticas de roteamento adotadas pelo DAC. O agente DAC também fornece informações sobre os indicadores para o agente gestor.

Figura 3.5 Descrição conceitual do modelo implementado neste trabalho com os agentes, interações e objetos da simulação para as centrais de teleatendimento

Agente Gestor

O tipo de agente gestor troca mensagens com todos os objetos e agentes. Através destas interações este agente detém as informações relativas ao desempenho de todo o sistema. A fila

característica é a noção da dimensão temporal, ou seja, existe informação interna, por exemplo, quando o mesmo realiza chamadas em um mesmo dia ("timeUnit"). Outro item é sobre o comportamento de abandono e decisões relacionadas a este tema: "patience" e "increasePatienceBehavior". Este tipo de agente possui dois momentos distintos no modelo de simulação quando ele está relacionando-se com a central e quando não está. Quando o agente está se relacionando com a central, o mesmo encontra-se em constante avaliação de sua situação, por exemplo, a decisão sobre quanto tempo que irá esperar na fila. Este item sobre o comportamento de espera na fila é alvo de um experimento e será descrito na *subseçsumcememés*)-1.87

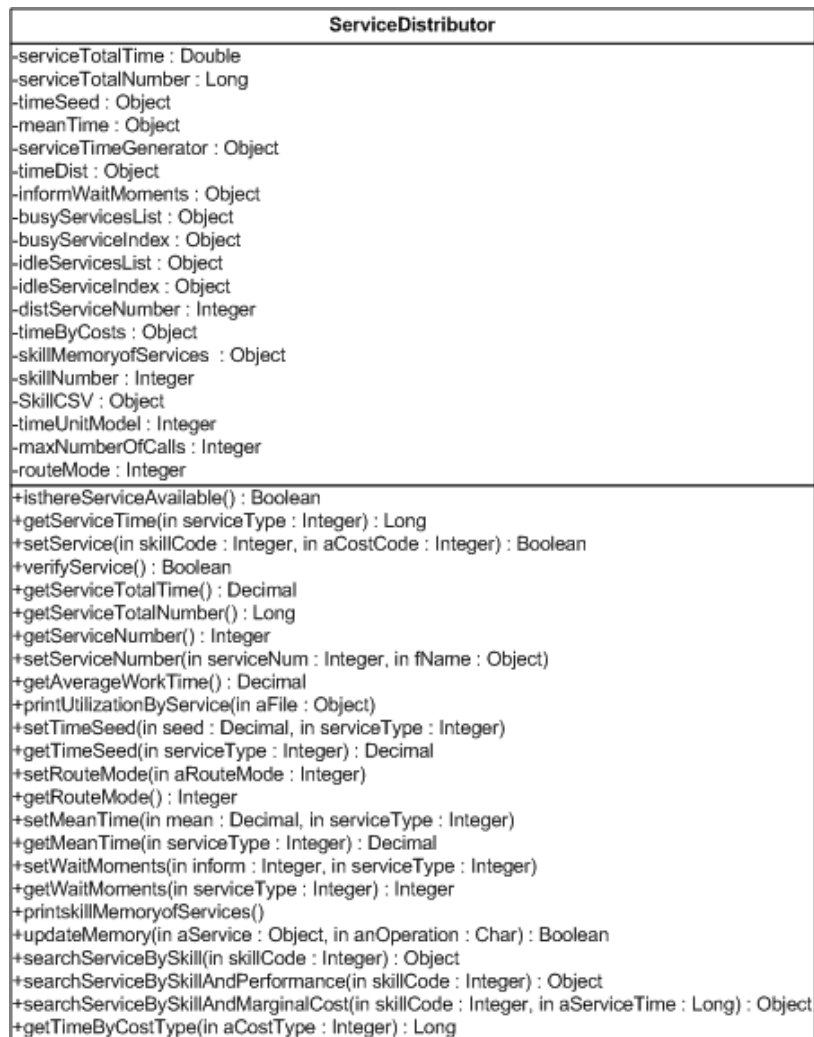


Figura 3.7 Diagrama de classe para o tipo de agente DAC- "ServiceDistributor"

3.3 Descrição de parâmetros de entrada e de saída do modelo

O modelo de simulação permite diferentes tipos de configurações que visam abranger diferentes possibilidades e flexibilizar a adaptação do modelo para os mais variados tipos de casos. Existem parâmetros que alteram questões gerais relacionadas ao ambiente de simulação (tempo de simulação, tipos e número de atendentes, tipos de chamadas, tamanho máximo da fila) e outros que são distribuídos que poderão influenciar o comportamento dos agentes (comportamento de abandono do cliente, habilidade e eficiência de cada atendente, políticas de roteamento do DAC).

3.3.1 Parâmetros de entrada

Os parâmetros a seguir apresentam as opções de configuração que influenciam a forma como o modelo irá executar:

- Clientes: número de agentes do tipo clientes inseridos no modelo;
- Tempo de simulação: tempo em segundos para execução do modelo. Este tempo

será convertido para um número de chamadas de acordo com a taxa de ocorrência de cada serviço. Exemplo: tempo de simulação pela taxa de ocorrência por segundo => $60 * 1 = 60$ Chamadas

- Unidade de tempo da simulação: corresponde a equivalência de um ciclo na simulação, pode equivaler a uma unidade básica como um dia de trabalho - deve ser expresso em segundos assim como o parâmetro tempo de simulação;
- Tipos de serviços: número de tipos de serviços ou tipo de chamadas que serão considerados no modelo de simulação;
- Arquivo com informações para cada tipo de serviço: nome de um arquivo que deverá ter o número de linhas igual ao especificado no parâmetro tipos de serviços. Este arquivo deve conter os seguintes dados para cada tipo de serviço:
 - Intervalo de chegada das chamadas, exemplo: 1 minuto. Pode ser especificado por unidade de tempo de simulação, ou seja, durante a simulação este valor será alterado por unidade de tempo (horário, por exemplo);
 - Taxa de ocorrência pelo intervalo de chegada (20 chamadas). Também pode ser alterado por unidade de tempo da simulação, da mesma forma que o intervalo de chegada das chamadas;
 - Média do tempo de atendimento/serviço, expresso em segundos. Pode ser alterado por unidade de tempo da simulação;
 - Média da paciência, expresso em segundos ou -1 para paciência infinita;
 - Informação de tempo de espera, ou seja, se o cliente será informado para aquele determinado serviço o tempo médio de espera na fila (valor 1 ou 0);
- Fator de bloqueio da fila: equivale ao tamanho máximo que a fila pode chegar. O valor -1 significa que a fila pode ter tamanho "infinito";
- Fator indicativo de um bom serviço: fator que indica o nível de serviço. Na prática equivale a tempo em segundos que o cliente pode esperar na fila para ser considerado como um bom serviço. Exemplo: 30 segundos;
- Número de atendentes: número máximo de atendentes que podem trabalhar na central de teletendimento;
- Arquivo descritivo dos atendentes: nome de um arquivo que deverá ter o número de linhas igual ao especificado no parâmetro número de atendentes, uma linha para cada atendente, este arquivo deve listar o grau de eficiência e as habilidades para cada tipo de serviço que cada atendente possui;
- Comportamento de abandono: informa se os clientes poderão ter um comportamento de paciência crescente durante um mesmo dia =1 ou não =0;
- Tipos de custos: informa o número de custos diferentes de chamadas existentes na simulação, valores mais baixos possuem prioridade maior, ou seja, custos com indicadores 1 e 2 são mais caros e assim podem ser tratados em políticas de atendimento e roteamento, por exemplo. Este item é importante para o processo

de decisão sobre políticas de roteamento em centrais do tipo 0800 – o cliente não paga a chamada;

- Política de roteamento: informa o tipo de roteamento que será feito no modelo de simulação. Nesta versão os seguintes métodos de roteamentos estão disponíveis¹⁵:
 - Roteamento baseado em habilidades;
 - Roteamento baseado em habilidades e eficiência;
 - Roteamento baseado em habilidades e custo marginal (considera a quantidade de chamadas e a duração das mesmas).
- Escala de atendentes por faixa horária: informa a distribuição do número de atendentes durante o tempo de simulação, este número deve ser expresso por unidade de tempo da simulação, ou seja, para cada unidade de tempo pode ser especificada um número de atendentes.

3.3.2 Saídas do modelo

Os resultados do modelo são indicadores e foram considerados dado sua importância no processo de tomada de decisão.

- Média de tempo de espera na fila: tempo médio que o

Brasil, números 0800, por exemplo - o cliente não tem custo com as ligações), o valor por tipo de chamadas é expresso em segundos. Esta abordagem é importante no processo de tomada de decisão destes tipos de centrais, dado que os custos com as ligações é uma das maiores despesas que a central possui [Valladares, 2007].

4 Resultados Experimentais

Neste capítulo é apresentado uma série de experimentos realizados com o modelo de simulação desenvolvido. O objetivo é evidenciar a viabilidade da abordagem e as suas capacidades adicionais. Os seguintes experimentos serão discutidos: comparação de resultados com modelos clássicos do tipo “Erlang”, comparação com uma outra abordagem que utiliza cadeias de Markov para tratar tipos de serviços nas centrais de teleatendimento; comparação de resultados da simulação com dados reais de uma central de teleatendimento. Uma vez demonstrada a confiabilidade do modelo, um outro conjunto de experimentos irá apresentar capacidades derivadas de sua flexibilidade, estes itens tratam de alguns temas como: a paciência dos clientes enquanto esperam por atendimento, políticas de atendimento e roteamento. Estudos de caso também serão apresentados para exemplificar a extensão do modelo e os conceitos apresentados.

4.1 Os Modelos de fila clássicos e o modelo de simulação

O primeiro conjunto de experimentos realiza um teste comparativo dos resultados gerados pelo modelo de simulação com aqueles calculados a partir de modelos analíticos do tipo “Erlang”, para esta comparação foram utilizados alguns indicadores apresentados na tabela 4.1. Neste experimento adotou-se, conforme pode ser observado na tabela 4.2, alguns modelos e seus parâmetros. Estes modelos são interessantes por conterem variadas características e foram propostos em [Mazzuchi e Wallace, 2004] e também utilizados em [Lopes et al., 2006].

Tabela 4.1 Indicadores utilizados para experimentos comparativo entre resultados do modelo de simulação e modelos analíticos do tipo “Erlang”

Indicadores	Definição
	Probabilidade de Bloqueio.
W_i	Média de espera para atendimento para o tipo de chamada i .
i	Nível de Serviço, ou seja, percentual de chamadas do tipo i atendidas em até unidades de tempo
	Utilização dos atendentes, definido como percentual entre o tempo trabalhado e o tempo total

Revisando o que foi apresentado na subseção 2.3.4 – Modelos “clássicos” para auxílio a tomada de decisão, os parâmetros da tabela 4.2 indicam os seguintes itens: λ é a taxa de chegada; β é a taxa de atendimento; s é o número de servidores; e τ representa o tempo aceitável de espera na fila (usado no cálculo do nível de serviço). Os modelos possuem particularidades específicas: o modelo $M/M/90/0$ não possui fila (não há possibilidade do cliente esperar), os modelos $M/M/1/\infty$ e $M/M/2/\infty$ permitem uma fila de tamanho infinito e com taxa de chegadas pequenas, já o ultimo modelo, $M/M/108/10$, suporta uma fila de até 10 chamadas de clientes, e, talvez, seja o modelo que mais se assemelha ao caso de uma central de teleatendimento real.

Para cada modelo descrito na tabela 4.2 foram realizados seis séries de experimentos considerando 10.000, 20.000, 40.000, 100.000, 500.000 e 1.000.000 de chamadas. Para cada série de experimentos o modelo de simulação foi executado dez vezes e para cada execução do modelo (com os mesmos parâmetros) houve a variação da semente para geração dos

valores pseudo-aleatórios das taxas de chegada (distribuição Poisson) e tempos dos serviços (distribuição exponencial). Os resultados obtidos das médias de cada série de experimentos está sintetizado na tabela 4.3. Além da média, para cada conjunto de experimentos foi calculado o desvio padrão, os resultados podem ser observados na tabela 4.4.

Tabela 4.2 Parâmetros descritivos dos modelos utilizados no experimentos de comparação com modelo analíticos do tipo “Erlang”

Modelo	Parâmetros			
	λ	β	s	τ
M/M/90/0	10 chamadas/min	10 min	90	30 s
M/M/1/∞	0,09 chamadas/min	10 min	1	30 s
M/M/2/∞	0,186 chamadas/min	10 min	2	30 s
M/M/108/10	10 chamadas/min	10 min	108	30 s

Tabela 4.3 Resultados com as médias dos indicadores na execução do modelo de simulação com 10.000, 20.000, 40.000, 100.000, 500.000 e 1.000.000 de chamadas.

Modelo	Indicador	Resultado Analítico	Resultados da Simulação – Swarm					
			Média					
			10.000	20.000	40.000	100.000	500.000	1.000.000
M/M/90/0	B - (%)	14,61	15,58	15,06	14,67	14,86	14,75	14,77
	W_i - (s)	NSA	NSA	NSA	NSA	NSA	NSA	NSA
	δ_i - (%)	NF	84,41	84,94	85,32	85,13	85,25	85,22
	v - (%)	92,00	91,21	92,82	93,26	94,46	94,72	94,77
M/M/1/∞	B - (%)	NSA	NSA	NSA	NSA	NSA	NSA	NSA
	W_i - (s)	90,00	99,72	94,30	94,53	92,82	92,95	91,87
	δ_i - (%)	10,40	8,93	9,61	9,79	10,04	10,18	10,16
	v - (%)	92,00	91,52	90,76	90,42	90,39	90,15	90,14
M/M/2/∞	B - (%)	NSA	NSA	NSA	NSA	NSA	NSA	NSA
	W_i - (s)	63,30	102,30	77,65	74,64	62,10	66,98	66,88
	δ_i - (%)	11,10	8,45	10,21	10,04	10,78	10,75	10,60
	v - (%)	92,00	94,79	93,68	93,31	93,20	93,13	93,19
M/M/108/10	B - (%)	1,30	1,80	1,54	1,48	1,37	1,35	1,35
	W_i - (s)	0,09	0,10	0,09	0,09	0,09	0,09	0,09
	δ_i - (%)	91,90	89,22	90,09	90,41	90,62	90,87	90,91
	v - (%)	91,00	89,66	88,44	90,38	91,06	91,29	91,41

NSA = Não se Aplica

NF = Não Fornecido

Para todos os modelos houve melhora dos resultados à medida que número de chamadas era aumentado. O que pode ser observado avaliando a tabela 4.3 com os valores da média de cada conjunto de experimentos. Outra observação é que a partir de 100.000 chamadas as diferenças entre as médias fica marginal. Este item também é evidenciado na tabela 4.4, observando-se os desvios padrões. Apesar de haver para os casos M/M/1/∞ e M/M/2/∞ um valor do desvio padrão alto para o indicador de tempo médio de espera na fila, os outros

modelos, $M/M/108/10$ e $M/M/90/0$, não se comportaram da mesma forma (estes modelos não tem tamanho da fila infinita). O valor de chamadas 100.000 parece indicar ser o número de chamadas para realização de experimentos com o modelo de simulação, no entanto ressalta-se

foram gerados seguindo a distribuição exponencial (utilizada para fornecer o tempo de atendimento no modelo) com um valor médio de seiscentos segundos. Da esquerda para a direita e de cima para baixo estão os gráficos com respectivamente *10, 100, 1.000, 10.000, 100.000 e 1.000.000* valores gerados. É visualmente perceptível que a medida que o número de valores gerados é aumentado a curva dos gráficos se assemelha mais e mais com a curva exponencial. Acredita-se que esta característica seja a principal causa da diferença entre os

representam 4.000 minutos ou 67 horas de "realidade" simulada, ~0,09 chamadas/min correspondem a 444.444 minutos ou 7407 horas ou 309 dias e ~0,186 chamadas/minuto correspondem 3.584 horas ou 150 dias; o modelo de simulação foi construído para trabalhar com a granularidade de representação do tempo da simulação em segundos. Na máquina em que foram realizados os experimentos (AMD 3.0-64, 512MB) e com um exemplo em que a taxa de chegada é ~10 chamadas/min cada segundo de execução do programa *Swarm* é equivalente a 10⁵ segundos de "realidade" simulada.

Tabela 4.5 Diferenças percentuais entre os resultados dos indicadores obtidos com a execução do modelo de simulação com 100.000 de chamadas e os modelos analíticos.

Modelo	Diferenças - %			
	B	W _i	δ _i	v
M/M/90/0	1,71	NSA	NF	2,67
M/M/1/∞	NSA	3,13	3,46	1,75
M/M/2/∞	NSA	1,90	2,88	1,30
M/M/108/10	5,38	0,00	1,39	0,07

NSA = Não se Aplica

NF = Não Fornecido

4.2 O modelo de Stanford e Grassman e o modelo de simulação

Outro conjunto de experimentos realizados verificou a capacidade do modelo para tratar situações em que existam vários tipos de serviços que são prestados pela central de teleatendimento. Cada um destes tipos de serviços possui características próprias (taxa de chegada, taxa de atendimento, número de atendentes dentre outros). No entanto, este tipos de serviço compartilham recursos únicos na central: acesso telefônico, uma mesma fila para os clientes e atendentes. O roteamento das chamadas dos clientes para os atendentes é realizado com a restrição de busca de um atendente que está preparado (foi treinado, por exemplo) para prestar o serviço.

Os modelos analíticos do tipo “Erlang” não tratam mais do que um tipo de serviço. Uma abordagem para este caso pode ser encontrada em [Stanford e Grassmann, 1998] em que é descrito um caso de uma central de teleatendimento com 20 atendentes, sendo 12 especialistas (realizam atendimento para o tipo de serviço 1, por exemplo) e os demais generalistas (realizam atendimento do tipo de serviço 1 e também do tipo de serviço 2, por exemplo). A tabela 4.6 apresenta detalhes dos parâmetros do modelo.

Tabela 4.6 Parâmetros descritivos de um modelo utilizado no experimento representando uma central com dois tipos de serviços diferentes [Stanford e Grassmann, 1998]

Tipos de Atendentes	Parâmetros			
	λ	β	s	τ
1 - Especialistas	0,384 chamadas/seg	25 s	12	10 s
2 - Generalistas	0,256 chamadas/seg	25 s	8	10 s

Para solução deste problema Stanford e Grassmann utilizaram cadeias de *markov* contínuas com uma matriz geométrica. Estes autores também assumiram as seguintes proposições:

- i) quantidade de chamadas são as mesmas para ambos os tipos, sendo assim o tempo do experimento varia de acordo com taxa de chegada das chamadas;
- ii) como o tempo da simulação varia, o cálculo do percentual de utilização do

modelo “Erlang C”, apresentado na expressão (12), foi alterado para tratar os dois tipos de atendentes da seguinte forma:

$$v = \frac{\lambda\beta}{s_1 + s_2} \tag{14}$$

Neste contexto foi realizado um experimento para comparar os resultados do trabalho de Stanford e Grassmann com os resultados do modelo de simulação. A primeira questão a ser destacada neste experimento é a alteração da premissa i) adotada por Stanford e Grassmann descrita anteriormente. Esta premissa assume que um mesmo número de chamadas é feita para os dois tipos de serviço o que faz com que o tempo de simulação seja diferente para os serviços no modelo de simulação, ou seja, no caso da simulação de 100.000 chamadas para serviço do tipo 1 com taxa de chegada 0,384 chamadas por segundo o tempo da simulação seria ~260.416 segundos; já para o tipo de serviço 2 com taxa de chegada 0,256 chamadas por segundo o tempo de simulação seria maior, ~390.625 segundos. Seriam simulados 4,52 dias, no entanto a partir do dia 3 não haveria mais chamadas do tipo 1, visto que sua taxa de chegada é maior do que o do tipo 2. Esta situação não reflete a realidade das centrais, a situação de que a partir de determinado momento não ocorrem mais chamadas de um tipo de serviço é irreal. O modelo de simulação deste trabalho busca a representação das centrais, desta forma esta premissa proposta por Stanford e Grassmann foi alterada. O modelo de simulação equaliza o tempo de simulação e faz um ajuste no número de chamadas, ou seja, para simulação de 4 dias, por exemplo, o número de chamadas para o tipo 1 será ~149.962 e para o tipo 2 ~99.975.

Tabela 4.7 Resumo comparativo de resultados entre o modelo de simulação e o trabalho de [Stanford e Grassmann, 1998] para o caso de uma central com dois tipos de serviços diferentes

Modelo	Indicador	Resultado Stanford e Grassmann (1998)	Resultados da Simulação – Swarm		Diferença - (%)
			Média	Desvio Padrão	
M ₂ /M/20/∞	B - (%)	NSA	NSA	NSA	NSA
	W ₁ - (s)	1,43	1,50	0,05	4,89
	W ₂ - (s)	9,39	10,18	0,18	8,41
	δ ₁ - (%)	NF	95,32	0,17	NF
	δ ₂ - (%)	NF	67,70	0,47	NF
	v - (%)	80,00	81,60	0,12	2,00

NSA = Não se Aplica

NF = Não Fornecido

Foram realizados dez experimentos (AMD 3.0-64, 512MB) em que manteve-se os parâmetros gerais e houve a variação da semente para geração nos números pseudo-aleatórios. Foi estipulado que houvesse a simulação de pelo menos 100.000 chamadas para o tipo 2, que apresenta menor taxa de chegada das chamadas. Para este caso cada experimento levou cerca de 15 segundos, ou seja, para cada 1 segundo simulado no Swarm havia a equivalência de ~226.041 segundos reais, considerando a taxa de chegada de 0,256 chamadas por segundo.

A tabela 4.7 apresenta um resumo comparativo com os resultados obtidos com o modelo de simulação versus o modelo de Stanford e Grassmann. O desvio padrão do conjunto de

experimentos do modelo de simulação manteve-se baixo. Apesar da alteração descrita anteriormente as diferenças percentuais entre os resultados do modelo de simulação e o modelo de Stanford e Grassmann são menores que 8,50%.

Em função da premissa adotada pelo modelo de Stanford e Grassmann, que não reflete a realidade das centrais, acredita-se que os resultados obtidos com o modelo de simulação sejam mais consistentes.

4.3 Modelo de simulação – comparação com dados reais de uma central de teleatendimento

Outro experimento realizado neste trabalho buscou comparar os dados e indicadores reais de uma central de teleatendimento com os resultados que modelo de simulação baseado em sistemas multiagentes pode gerar quando configurado para representar tal central. Durante as investigações neste trabalho uma aproximação com uma empresa prestadora de serviços públicos de Minas Gerais foi feita. Esta empresa atende todo o estado mineiro e clientes em outros estados. Possui cerca de 6.000.000 de clientes. Além da central de teleatendimentos a empresa possui agências fisicamente instaladas em todo estado. A central e estas agências são os pontos de contato dos clientes com a organização. Uma reunião foi realizada com os gestores da central para apresentar o presente trabalho, a partir desta reunião algumas correspondências eletrônicas foram trocadas para que o experimento descrito nesta seção pudesse ser viável.

A referida central de teleatendimento recebe em média, por mês, cerca de 1.600.000 chamadas, possui 344 PAs - Postos de Atendimentos (que são ocupados por vários atendentes em turnos e escalas diferentes), sua fila possui uma capacidade de 540 posições (em função da capacidade do tronco utilizado), um percentual de abandono que gira em torno de 4% e o indicador para cálculo de nível de serviço – τ , que representa o tempo aceitável de espera na fila para os clientes, de 30 segundos [Valladares, 2007]. O indicador mais importante para esta central é - Nível de Serviço, percentual de chamadas atendidas em até unidades de tempo, visto que este indicador é alvo de auditorias por parte de um órgão regulador do setor, além do comprometimento da central em ter excelência no atendimento de seus clientes.

Para realização do experimento, o gestor forneceu alguns dados sobre um dia típico da central. A tabela 4.8 apresenta estes dados que são assim caracterizados:

Segue um demonstrativo do volume de chamadas ao longo de uma quarta-feira (mês 11/2006) dentro dos padrões normais (chamamos de dia típico). Nessa data, tivemos um pequeno aumento de chamadas no horário noturno (pós 20h), mas que não chegou a interferir negativamente no desempenho do dia. Vale salientar que esse volume de chamadas (quase 60.000 nesse dia) diz respeito a chamadas recebidas. Dessas, parte é atendida nas PAs, parte é atendida pela URA e parte é abandonada. O TMA é o tempo médio de atendimento. [Valladares, 2007].

Uma outra informação fornecida sobre este dia, apresentado na tabela 4.8, é que o indicador de nível de serviço, τ , foi de 87,70 % [Valladares, 2007].

Com estes dados a meta do experimento é representar a central de teleatendimento para o referido dia apresentado. O intuito é a comparação do indicador fornecido, o τ . A primeira observação sobre os dados, vide tabela 4.8, é sobre a variação do número de chamadas recebidas em função do horário. Um gráfico sobre este item é mostrado na figura 2.5. Outra observação é a variação do número de postos de atendimento, uma estratégia de escalas

(número de atendentes) por horário é adotada para lidar com esta questão, por exemplo, de 00:00 até 05:00 há o menor número de postos de atendimento, 12 postos de atendimento, e de 09:00 até as 16:00 o maior número, 344 postos de atendimento. Esta variação dos dois ápices, em termos percentuais, representa um crescimento/decrescimento de $\sim 2.866,67\%$. O tempo médio de atendimento é apresentado de forma horária, a média destes valores horários para o dia é $\sim 220,33$ segundos com um desvio padrão de $\sim 38,35$.

Utilizando modelos do tipo “Erlang” o indicador para o caso desta central é de $\sim 99,99\%$, que fica distante do valor real ($12,29\%$), as motivações para esta diferença foram discutidas na subseção 2.3.4 – Modelos “clássicos” para auxílio a tomada de decisão. Este cálculo foi feito utilizando a ferramenta disponibilizada em [Koole, 2007]. Para este cálculo a taxa de chegada adotada foi a média das taxas das faixas horárias: $0,55$ chamadas por segundo; o número de atendentes foi a média do número de PAs das faixas horárias: 166 e o tempo de atendimento também foi a média tempo de atendimento das faixas horárias: $\sim 220,33$ segundos.

Tabela 4.8 Dados reais com número de chamadas, número de atendentes e tempo médio de atendimento de uma central de teleatendimento, valores correspondentes a um dia “típico” de funcionamento [Valladares, 2007].

Horário	Dados Repassados			
	Chamadas Recebidas	Chamadas atendidas – PAs	Nº PAs	Tempo médio atendimento – (s)
00:00	141	106	12	206
01:00	60	46	12	182
02:00	35	28	12	209
03:00	17	14	12	195
04:00	46	34	12	183
05:00	62	50	12	177
06:00	479	373	40	161
07:00	1559	1086	70	202
08:00	3363	2720	213	228
09:00	4580	3830	344	262
10:00	4640	3838	344	273
11:00	4314	3550	344	273
12:00	3988	3369	344	253
13:00	4432	3837	344	247
14:00	4599	3935	344	256
15:00	4439	3864	344	254
16:00	3975	3393	344	253
17:00	3486	2981	230	248
18:00	3012	2451	180	260
19:00	2819	2345	170	217
20:00	2798	1803	100	214
21:00	1780	1389	70	145
22:00	875	703	50	161
23:00	340	259	30	229

Na realização do experimento com o modelo de simulação uma premissa foi adotada. Esta premissa é relativa ao percentual de chamadas atendidas pela URA, o percentual é um dos motivos da diferença entre as colunas chamadas recebidas e chamadas atendidas PAs da tabela 4.8. Este tipo de atendimento e interação entre cliente e central não é tratado pelo modelo de simulação, uma vez que só envolve o agente cliente, não há interações com outros

agentes do modelo. Desta forma para calcular o número de chamadas totais a serem simuladas desconsideraram-se as chamadas atendidas pela URA. Fez-se uso da informação sobre o percentual de abandono: $\sim 4,00\%$ que foi decrescido da diferença percentual entre as colunas “Chamadas Recebidas” e “Chamadas atendidas - PAs”. Este cálculo indicou para cada faixa horária o percentual de chamadas atendidas pela URA, a média geral no dia ficou $\sim 15,84\%$ com desvio padrão de $0,05$. Desta forma o número anterior de chamadas no dia que era de 55.839 chamadas passou para 48.238 no modelo de simulação.

O modelo de simulação foi então configurado com os dados apresentados: um dia de simulação e o número de chamadas, taxa de chegada, tempo de atendimento e número de atendentes seguindo as faixas horárias. A mesma metodologia adotada no conjunto de experimentos para comparação com modelos clássicos (do tipo “Erlang” e o modelo de Stanford e Grassman) foi utilizada, ou seja, o modelo foi executado dez vezes e para cada execução houve a variação apenas da semente para geração dos números pseudo-aleatórios. Cada experimento levou cerca de ~ 58 segundos, ou seja, 1 dia (carga de trabalho de 24 horas) foi simulado em aproximadamente 1 minuto.

O valor obtido no conjunto de experimento para o indicador nível de serviço médio no dia - foi de $81,97\%$ com um desvio padrão de $0,04$. Se comparado com o número real apurado na central $87,70\%$, há uma diferença de $5,97\%$.

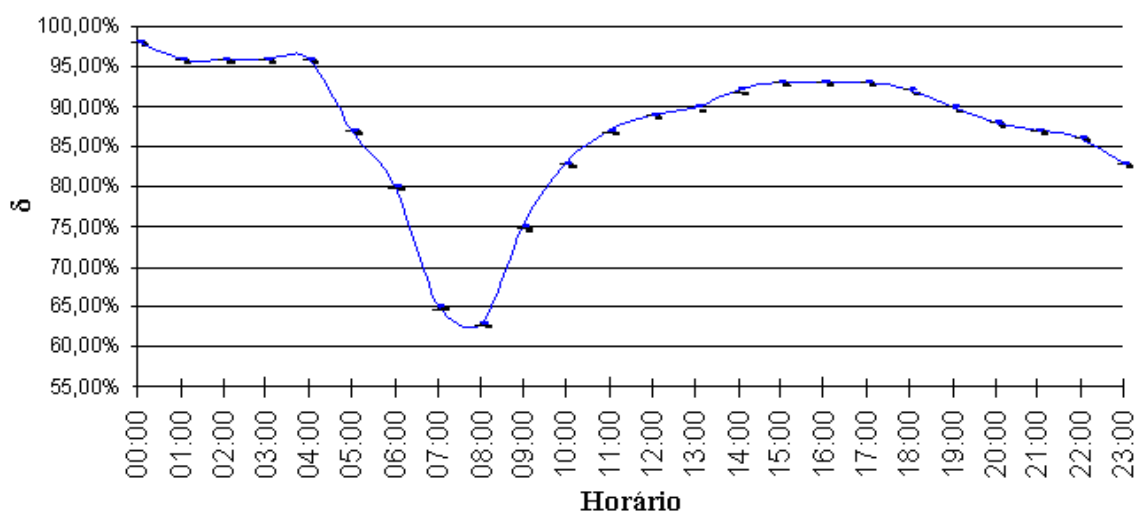


Figura 4.2 Gráfico com o indicador nível de serviço - δ estratificado por faixa horária e obtido através do modelo de simulação no experimento de comparação do indicador simulado com o valor real.

Esta diferença levou a um segundo experimento. Durante a visita realizada na central vários assuntos foram abordados e tratados: rotatividade de pessoal, absenteísmo, comportamento de abandono, religações dos clientes, produtividade e eficiência dos atendentes. Dentre estes itens a eficiência dos atendentes é um fator que pode ser utilizado no experimento da simulação de um dia. Este fator também foi ressaltado pelo gestor da central, ou seja, existe um grupo de atendentes mais experientes, para estes atendentes algumas vezes as chamadas mais complexas e emergenciais são encaminhadas. Como no modelo de simulação cada atendente é um agente do modelo, foi feita uma alteração em cerca de 40 agentes do tipo atendente. Esta alteração fez um acréscimo de 30% de eficiência para estes atendentes, ou seja, para qualquer chamada estes atendentes realizam o atendimento mais rapidamente. Com esta alteração o indicador nível de serviço - que era $81,97\%$ foi para $86,32\%$. Comparado com o número real apurado na central a diferença é de $1,38\%$. A figura 4.2 apresenta o valor

do indicador nível de serviço - por faixa horária. Neste gráfico a média de 86,32% obtida pelo modelo de simulação está estratificada e é apresentada por faixa horária com seus desvios padrões, que são muito pequenos e quase imperceptíveis no gráfico.

A curva na figura 4.2 ainda não foi comparada com a realidade da central, mas em principio a média é bastante próxima. Outras alternativas também devem ser avaliadas e testadas no modelo de simulação. Para resposta de quais variáveis realmente influem no processo é necessário conhecimento especialista, neste caso representado pelo conhecimento dos gestores da central. Neste sentido o modelo atua como suporte ao entendimento da dinâmica de funcionamento da central.

4.4 Capacidades adicionais do modelo de simulação

Nesta seção alguns experimentos são demonstrados focando a flexibilidade do modelo em lidar com questões relacionadas a complexidade do funcionamento e da dinâmica existente nas centrais de teleatendimento. Os experimentos realizados baseiam-se na premissa de que o modelo de simulação apresenta resultados confiáveis quando comparados com modelos analíticos (diferença percentual dos indicadores, erro 8,50 %, para o pior caso) e em situações reais (diferença percentual dos indicadores, erro 1,38 %).

4.4.1 Simulação do abandono e o comportamento de espera dos clientes

Para explorar a flexibilidade do modelo de simulação um aspecto relevante foi considerado para estender o modelo. Este aspecto é o abandono dos clientes enquanto estão aguardando por atendimento na fila. O modelo “Erlang A” - $M/M/s + M$ considera a questão do abandono [Koole e Mandelbaum, 2002]. Este modelo atribui ao comportamento dos clientes uma distribuição exponencial de abandono (o modelo deste trabalho também implementa a abordagem utilizada no modelo “Erlang A”). No entanto em [Zohar et al., 2002] e [Mandaulbaum et al., 2006] uma antítese é apresentada baseada em dados reais. Em síntese o fator determinante para o comportamento do abandono é a paciência dos clientes, ou seja, um comportamento humano. O modelo de simulação foi utilizado para simular um comportamento específico de alguns clientes, o impacto na consideração deste comportamento nos indicadores da central também será apresentado e discutido.

4.4.1.1 Análise do abandono na perspectiva de uma base de dados real

Uma revisão sobre as tentativas existentes para modelar o comportamento de abandono dos clientes nas centrais é apresentada em [Zohar et al., 2002]. No referido texto o processo de abandono é conceituado como predominantemente psicológico e que é disparado quando sentimentos negativos por parte do cliente são acumulados durante a espera, e ultrapassam determinado limite pessoal. A modelagem da “paciência” ainda é um desafio a ser explorado. No caso das centrais de teleatendimento uma série de fatores podem ser citados como capazes de influenciar a paciência do cliente: percepção do serviço da central pelo cliente (boa qualidade, má qualidade, histórico de atendimentos), percepção do tempo de espera, urgência ou importância do serviço ou produto requerido na chamada, quem é o pagador da chamada (centrais do tipo 0800 ou o próprio cliente) dentre outros.

Os modelos clássicos de fila, como “Erlang A”, consideram que não há relação da paciência do cliente com o desempenho do sistema, ou seja, a distribuição exponencial é utilizada independentemente, por exemplo, de haver um grande congestionamento na central.

Em [Zohar et al., 2002] uma antítese para este item é demonstrada através de um caso real e o comportamento de clientes experientes e novatos são apresentados com diferenças. Um modelo para estes comportamentos também é apresentado e experimentos são realizados. Sobre o assunto de abandono em centrais de teleatendimento importantes trabalhos têm sido publicados pelo professor Mandelbaum, algumas observações são citadas em [Mandaulbaum et al., 2006]:

- Não necessariamente a paciência segue uma distribuição exponencial, ela varia significativamente de acordo com: o tipo de serviço, a urgência/utilidade do serviço para o cliente, e se são ou não fornecidas informações para o cliente enquanto o mesmo espera (previsão do tempo de atendimento, por exemplo);
- O tempo de espera dos clientes que não abandonam é exponencial;
- Clientes experientes adaptam seu comportamento de acordo com o desempenho global do sistema, enquanto que novos clientes são menos sensíveis a este desempenho.

Nas investigações feitas sobre abandono o ponto inicial foi uma base de dados disponibilizada pelo professor Mandelbaum [Mandaulbaum et al., 2001], esta base esta disponível no site do professor. O autor deste trabalho obteve autorização para utilização da mesma [Mandaulbaum, 2006]. Vide *Anexo I – Mensagens eletrônicas enviadas e recebidas*.

Um tratamento para esta base de dados foi realizado e a mesma foi migrada para um banco de dados relacional, desta forma consultas poderiam ser realizadas. A base apresenta dados detalhados para cada uma das chamadas realizadas: instante em que o cliente entrou na URA, instante em que recebeu atendimento, instante em que abandonou, identificador do cliente, identificador do atendente dentre outros. No *Apêndice B – Base de dados e programas para obtenção de resultados* é apresentada a tabela das chamadas e os programas utilizados para obter informações e características do ambiente da central de teleatendimento, seus clientes e atendentes.

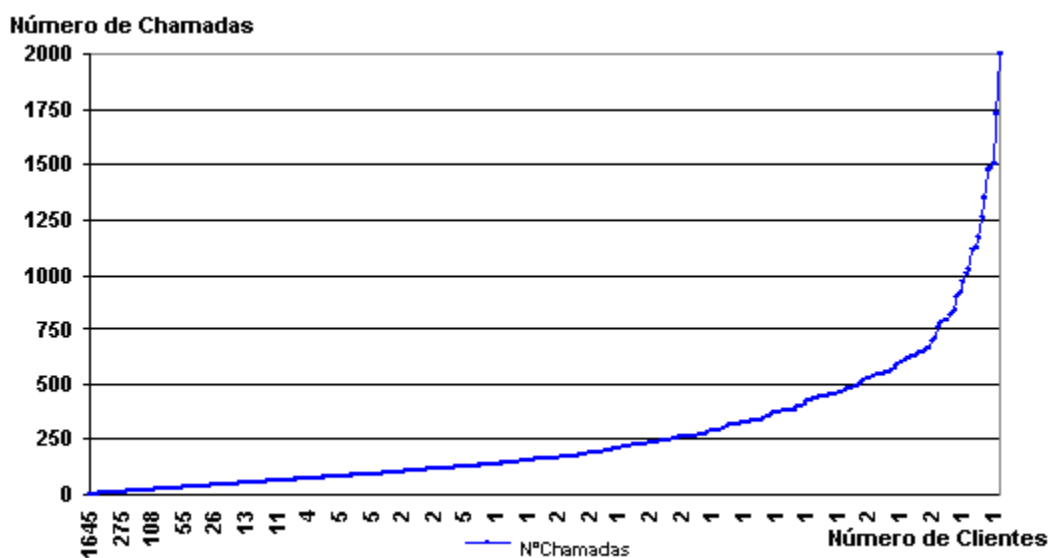


Figura 4.3 Gráfico com o número de chamadas pelo número de clientes que efetuaram-as. Informação derivada da base de dados disponibilizada em [Mandaulbaum, 2006].

A base de dados apresenta um total de 444.448 chamadas, que é todo o movimento de chamadas no ano de 1999 de uma instituição financeira sediada no Estado de Israel. Na base

existem 12.907 clientes. O comportamento destes clientes quanto ao número de chamadas é apresentado na figura 4.3. Esta figura mostra o número de clientes pelo número de chamadas, por exemplo, em um lado do gráfico estão 1.645 clientes que fizeram 2 chamadas durante o ano, já em outro há 1 cliente que realizou 1996 chamadas.

Dos 12.907 clientes cerca de 10.196 ligaram para central mais de uma vez no ano, destes 7.442 clientes abandonaram suas chamadas enquanto esperavam na fila por atendimento. A tabela 4.9 apresenta os dez maiores grupos de clientes em percentual do total relativo ao percentual de abandono. Após o décimo grupo todos os percentuais do total são menores que 3,00% e apresentam percentuais de abandono diversos. Dos clientes que abandonaram 4.976 clientes abandonaram mais de uma vez, um total de 34.035 chamadas, ou seja, existe uma série histórica que pode ser avaliada para estes clientes.

Tabela 4.9 Os dez maiores grupos de clientes por percentual do total. Contém o número de clientes e o percentual de abandono para cada um dos grupos apresentados. Informação derivada da base de dados disponibilizada em [Mandaulbaum, 2006]

Grupos	Percentual do Total - (%)	Nº de Clientes	Percentual de abandono - (%)
1	12,31	916	50,00
2	9,10	677	33,00
3	6,80	506	25,00
4	5,60	417	20,00
5	4,14	308	16,00
6	3,55	264	14,00
7	3,06	228	40,00
8	2,71	202	12,00
9	2,49	185	28,00
10	2,43	181	66,00

Identificado este conjunto de dados foi feita uma análise estatística dos dados avaliados; média e desvio padrão dos tempos de espera na fila até o instante de abandono foram avaliados. Vários tipos de agrupamentos de dados foram também criados para análise: média e desvio por cliente, mês, dia, tipo de serviço, por hora do dia dentre outros. Estas análises foram feitas com o intuito de perceber nas chamadas dos clientes com reincidência de abandono, como foi a "paciência" para cada chamada. A busca era por um padrão. A paciência foi reduzida ou aumentada com o tempo? A premissa é que só é possível inferir algo sobre a paciência do cliente quando acompanha-se a evolução do relacionamento deste cliente com a central.

A massa de dados de abandono foi estudada e avaliada e não conseguiu-se perceber relações nos dados. Não foi possível neste estudo determinar o comportamento da paciência ou inferir relações de variáveis (talvez em um trabalho futuro técnicas de mineração de dados possam ser aplicadas para explorar este item).

Apesar disto um padrão ou "comportamento" foi observado e pode ser evidenciado na base de dados: comportamento crescente de paciência em um mesmo dia e tipo de serviço. Este comportamento foi observado sobre um conjunto de clientes que abandonaram o atendimento em um dia e voltam a ligar naquele mesmo dia, quando isto acontece este grupo sempre apresenta um aumento de paciência, mesmo se eles abandonam novamente. Sua paciência ou o tempo de espera na fila em uma nova tentativa de chamada para a central é crescente. Isto é válido sempre para um mesmo dia em chamadas para um mesmo tipo de serviço.

Das 34.035 chamadas de clientes que já abandonaram mais de uma vez cerca de 2.182 chamadas não seguiram o padrão de comportamento crescente de paciência em um mesmo dia e tipo de serviço. Considerando somente o conjunto de chamadas que aconteceram em um mesmo dia tem-se cerca de 12.788 chamadas, deste total cerca de 82,94% das chamadas dos clientes que acontecerem no mesmo dia e tipo de serviço seguiram o padrão de comportamento. Sobre a perspectiva dos clientes, em 1.023 clientes foi observado o caso de não seguir a regra, sob o total de 4.976 clientes que já abandonaram mais de uma vez cerca de 79,45% dos clientes seguiram este comportamento. Em suma a regra acontece em ~83,00% das chamadas com abandono em um mesmo dia e tipo de serviço, e do total de clientes que já abandonaram mais de uma vez cerca de ~80,00% tiveram este comportamento.

Acredita-se que isto se deve em função da urgência do serviço demandado, se o cliente liga várias vezes no dia é porque deve haver um forte motivador indicando necessidade do serviço ou produto, logo ele se torna mais e mais paciente naquele dia, período de sua demanda.

4.4.1.2 Comportamento crescente para o tempo de espera na fila e suas implicações nos indicadores da central de teleatendimento

Uma vez evidenciado este fato duas questões foram levantadas no decorrer do trabalho:

- i) Seria possível implementar este comportamento nos agentes clientes do modelo de simulação?
- ii) O que acontece com os indicadores da central com este comportamento dos clientes? Que tipos de impactos este comportamento de cada cliente pode gerar nos indicadores da central?

A resposta para a primeira questão está descrita na *seção 3.2 – Detalhamento de alguns agentes*. Foi descrito, por exemplo, que o agente cliente possui memória sobre todos os relacionamentos que ele fez com a central de teleatendimento. Outra capacidade deste agente é ter consciência do tempo, ou seja, em suas decisões ele leva em conta a dimensão tempo, conforme requerido para realização do experimento que representa um comportamento de paciência crescente durante um mesmo dia.

Dada a flexibilidade do modelo que suporta a primeira indagação, um experimento foi elaborado para responder a segunda questão. A descrição das seguintes características é válida para o experimento realizado:

- 1 dia de simulação, este dia possui 16 horas de trabalho, ou seja, de 07:00 às 23:00, assim como acontece na base de dados avaliada;
- 4.000 clientes foi o número de clientes que existem para interagir com a central no dia que será simulado;
- A taxa média de chegada das chamadas é 0,80 chamadas por segundo;
- O tempo médio de atendimento é de 60 segundos;
- O tempo aceitável e considerado como bom atendimento - τ é de 30 segundos;
- A fila da central possui uma limitação de 100 posições;
- O número de atendentes é 50.

Capacidade adicional do modelo de simulação:

- A média da paciência inicial para cada cliente é de 30 segundos. A cada nova chamada para a central no mesmo dia esta paciência aumenta em 30,00% - taxa crescente da paciência;
- Dos 4.000 clientes, cerca de 80,00% apresentam o comportamento de taxa crescente de paciência em um mesmo dia e tipo de serviço;

Uma questão sobre este experimento foi levantada: como se comporta a taxa de crescimento da paciência? Na descrição acima foi estipulado um valor de 30,00%, por exemplo. Mas qual é a taxa existente na base de dados? Qual a forma da curva desta variável? Neste intuito um novo trabalho de análise da base de dados foi feito. O resultado é a figura 4.4 que apresenta as chamadas que tiveram um comportamento de paciência crescente após uma chamada anterior no mesmo dia, assim o valor em percentual da diferença entre o tempo de espera anterior e o tempo de espera em que a paciência aumentou foi calculado. Em um eixo o percentual de crescimento em relação ao tempo de espera anterior em um mesmo dia e em outro as chamadas que tivera

548(a)-57.6318()-70.7156(m)-di-76-44- dia da de crescim

dados.

A tabela 4.10 apresenta os resultados dos indicadores obtidos com os experimentos. O indicador A significa o percentual de abandono dos clientes. Nestes experimentos adotou-se a mesma metodologia dos experimentos realizados até então (o indicador é uma média de um conjunto dez execuções do modelo de simulação em que a semente do gerador de número pseudo-aleatórios é o único parâmetro alterado). Algumas observações sobre os indicadores na tabela 4.10: a medida que o percentual de abandono - A diminui ou aumenta o tempo médio de espera na fila - W_i também diminui ou aumenta, isto ocasiona oscilações no indicador de nível de serviço δ_j . Já o percentual de ocupação - v praticamente não se altera.

Este experimento corrobora com uma série de trabalhos, apresentados em [Mandaulbaum, 2005], que ressaltam a importância em considerar a questão da paciência do cliente quando na modelagem das centrais de teleatendimento. O experimento adiciona ainda a questão de incluir na modelagem das centrais um comportamento dos clientes (a partir de fatos evidenciados na base de dados). Este comportamento dos clientes, taxa crescente de paciência

teleatendimento lidam com vários tipos de serviço na prática. A política de roteamento de chamadas baseada em habilidades é uma política conhecida e aplicada neste contexto [L'Ecuyer e Avyramidis, 2005], [Mazzuchi e Wallace, 2004], [Koole, 2005], [Wallace e Whitt, 2006] e [Wallace e Saltzman, 2005]. Nesta subseção será apresentado um experimento aplicando a política de roteamento baseado em habilidades no modelo de simulação.

4.4.2.1 Definições sobre política de roteamento baseado em habilidades

Os modelos analíticos para as centrais de teleatendimento consideram apenas um tipo de serviço prestado pela central, vide *subseção 2.3.4 Modelos “clássicos” para auxílio a tomada de decisão*. Isto não reflete a realidade. As centrais prestam vários tipos de serviços com características diferentes, como: atendentes, taxa de chegada de chamadas, média de tempo de atendimento. Modelos considerando mais que um tipo de chamada são recentes [Whitt, 2002]. A capacidade de tratar vários tipos de chamadas apresenta novas questões ao processo de tomada de decisão:

- i) Quantos tipos de atendentes por tipo de chamada d

avalia-se um intermediário ou em ultima instância um generalista. Caso não haja nenhum atendente, o cliente será encaminhado para fila respeitando-se o fator de bloqueio da mesma. Veja o pseudo-algoritmo da política de roteamento baseada em habilidades apresentado na figura 4.5.

No modelo de simulação a atribuição de rotear as chamadas é responsabilidade do agente do tipo DAC. Este agente interage com a fila avaliando a posição e situação de clientes. Este agente também conhece as escalas de trabalho e características de cada um dos atendentes, como por exemplo, suas habilidades e capacidades (nível 1,2 3...) para atendimento de cada tipo de serviço. Com estas informações este agente realiza o roteamento baseado em habilidades.

```

nova_chamada = novo (evento)
  Se (todos atendentes estão ocupados)
    então cliente é colocado na fila ou bloqueado
  senão
    tipo_serviço = nova_chamada.tipo_serviço
    Se (existem especialistas disponíveis para este tipo_serviço)
      então roteia_chamada_para_especialista (nova_chamada)
    senão
      Se (existem intermediários disponíveis para este tipo_serviço)
        então roteia_chamada_para_intermediários (nova_chamada)
      senão
        Se (existem generalistas disponíveis para este tipo_serviço)
          então roteia_chamada_para_generalistas (nova_chamada)
        senão cliente é colocado na fila ou bloqueado
    
```

Figura 4.5 Pseudo-algoritmo para roteamento baseado em habilidades, adaptado de [Mazzuchi eWallace, 2004].

4.4.2.2 Estudo de caso

A resolução normativa Nº. 57 da Agência Nacional de Energia Elétrica (ANEEL) estabelece para as concessionárias de energia metas nos níveis de serviços das centrais de teleatendimento. Mais especificamente alguns tipos de centrais deverão melhorar seu índice de nível de serviço de 80,00% em 2005 para 95,00% em 2008 para que cumpram a norma da ANEEL. Isto demonstra a demanda por eficiência.

Sob esta ótica uma central fictícia foi representada. Considere três tipos de chamadas: **A-** Comercial para informações e reclamações, **B-** Comercial para solicitação de serviços e **C-** Emergencial (casos de fios partidos, falta de energia dentre outros). A tabela 4.11 apresenta informações sobre o cenário atual da central. Observe que o número de atendentes é maior para o tipo de chamada sobre solicitação de serviços, visto que há uma maior demanda para este tipo de serviço. Da mesma forma existe uma preocupação por parte da empresa em manter os níveis de serviço para as chamadas emergenciais altos; mais altos do que o exigido pelo órgão regulamentador. O fator de bloqueio desta central é de 10 posições na fila.

Tabela 4.11 Parâmetros para representar o estudo de caso de um central de teleatendimento fictícia utilizada no estudo de caso com a abordagem de política de roteamento baseada em habilidades.

Tipos de chamadas	Parâmetros			
	λ	β	s	τ
A - Comercial para informações e reclamações	6 chamadas/ min	5 min	38	30 s
B - Comercial para solicitação de serviços	10 chamadas/ min	5 min	60	30 s
C - Emergencial	0.20 chamadas/ min	5 min	2	30 s

No atual contexto os atendentes do tipo A têm recebido muitas reclamações sobre a indisponibilidade dos atendimentos emergenciais. Utilizando o modelo de simulação, avaliou-se que os níveis de serviço estão da seguinte forma: $\delta_A = 94,99\%$, $\delta_B = 95,51\%$ e $\delta_C = 70,34\%$. Sendo assim, o serviço emergencial está com $\sim 25,00\%$ abaixo da meta de $95,00\%$. Uma possível solução para este problema seria treinar os atendentes para que os mesmos pudessem atender outros tipos de chamadas e usar uma política de roteamento das chamadas baseada em habilidades. Outra restrição é manter a taxa de ocupação dos atendentes em torno de $80,00\%$, e desta forma o problema seria resolvido sem contratações de mão de obra adicional. Para esta avaliação foram elaborados algumas alternativas de distribuição de habilidades envolvendo todos os atendentes da central - cenários. Estes cenários representados na figura 4.6 foram elaboradas visando avaliar a migração dos atendentes de especialistas (situação atual) até a condição onde todos são generalistas (os cenários incrementam o número de atendentes seguindo esta característica). Os cenários de 2 a 6 foram escolhidos sem critério específico, são sugestões escolhidas para o estudo de caso. Já o cenário 1 representa a atual situação. Com estes cenários de distribuição o modelo de simulação foi utilizado para verificar os impactos (análises do tipo “e se”) nos indicadores: W_i e δ_i .

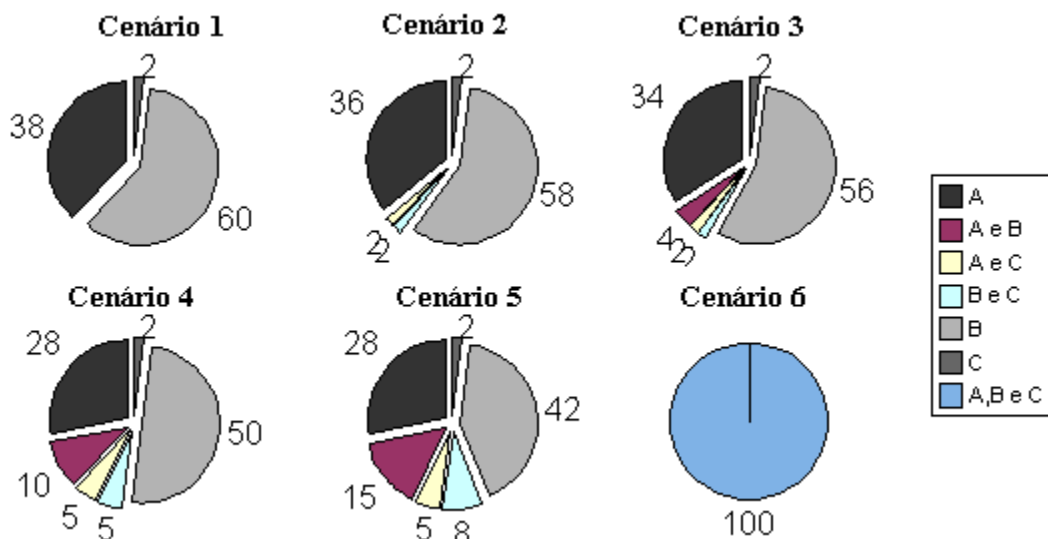


Figura 4.6 Cenários de simulação com a distribuição de atendentes e habilidades por tipo de serviço prestado pela central fictícia do estudo de caso.

A primeira análise é quanto ao tempo médio de espera na fila, apresentada na figura 4.7. O cenário 1 apresenta um tempo médio muito alto. Já o cenário 2, com o treinamento de quatro atendentes (dois do tipo A e dois do tipo B), a média sofre uma queda significativa, abaixo de 10 segundos. A partir do cenário 4 os tempos de espera de todos os serviços ficam menores que 2 segundos.

Outra análise é sobre os níveis de serviço, apresentada na figura 4.8. Da mesma forma que o tempo médio na fila há uma mudança significativa do cenário 1 para o cenário 2, ou seja, o tipo de serviço emergencial de $70,34\%$ vai para $92,76\%$. No entanto este valor ainda não é satisfatório para a meta da central, $95,00\%$. Somente no cenário 4, em que foram treinados mutuamente 20 atendentes é que o nível de serviço atinge $98,28\%$. Outro ponto é observado no cenário 6 em que os níveis de serviço ficam aproximadamente iguais com crescimento marginal se comparado aos cenários 4 e 5. Neste cenário todos os agentes são generalistas, este fenômeno é denominado de “resource pooling” [Wallace e Whitt, 2006].

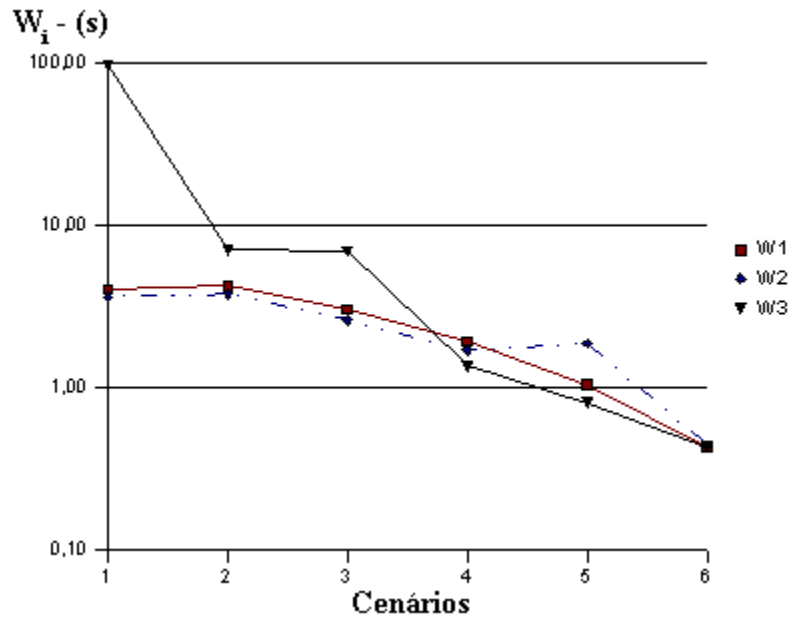


Figura 4.7 Resultados para o indicador do tempo médio de espera na fila - W_i em segundos, para cada um dos cenários do estudo de caso utilizando a política de roteamento baseada em habilidades.

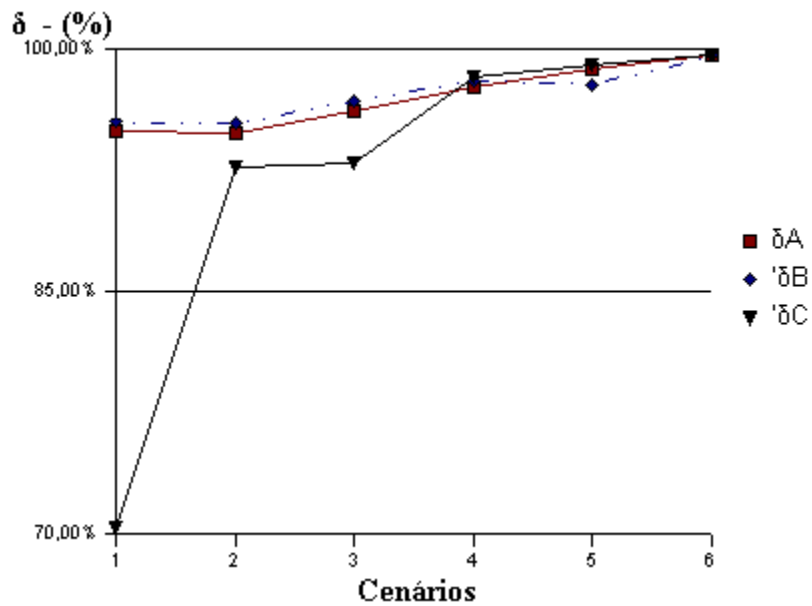


Figura 4.8 Resultados para o indicador de nível de serviço - δ_i , para cada um dos cenários do estudo de caso utilizando a política de roteamento baseada em habilidades.

O cenário 4 parece ser a solução mais adequada dentro dos demais avaliados, uma vez que ele atende a todas as restrições dos indicadores e apresenta menor custo (despesas com treinamentos e salários) em função do número de atendentes. Os resultados obtidos indicam que existe um problema de otimização a ser resolvido. Quantos atendentes e com que tipo de habilidade devem existir para atender as restrições de nível de serviço. Em trabalhos futuros, métodos clássicos de otimização podem ser aplicados aos resultados da simulação orientando para a melhor decisão, uma mescla de técnicas para solução efetiva do problema.

4.4.3 Filas com prioridades e políticas de roteamento baseada em custo marginal e desempenho dos atendentes

Muitas centrais de teleatendimento oferecem ligações grátis (telefones do tipo 0800) para seus clientes. O que é observado no Brasil, por exemplo, para as concessionárias de serviços públicos (saúde, energia elétrica, água e saneamento básico, dentre outros). Cerca de 60,00 a 70,00 % do custo das centrais é gasto em pagamentos de salários e benefícios para os atendentes [Koole e Mandelbaum, 2002]. No entanto, nas centrais que oferecem ligações

que este experimento pudesse ser comparado com outras abordagens e outros tipos de modelos em futuros trabalhos.

Tabela 4.12 Valores por minuto para cada tipo de chamada em um central fictícia idealizada para avaliação de fila com prioridades e políticas de roteamento de chamadas.

Prioridade	Valor por Minuto	Descrição
1	R\$ 0,40	Chamada longa distância de telefone móvel
2	R\$ 0,30	Chamada longa distância de telefone fixo
3	R\$ 0,20	Chamada local de telefone móvel
4	R\$ 0,10	Chamada local de telefone fixo

Nesta central existem quatro tipos de chamadas classificadas quanto ao custo da ligação. Na tabela 4.12 são representados os tipos de ligação. As chamadas com maior custo, neste experimento, possuem uma prioridade maior na central. A distribuição da ocorrência das chamadas no experimento foi adotada como constante e igualmente dividida para cada um dos tipos de chamadas existentes, ou seja, para cada chamada existe a probabilidade uniforme e constante de 25,00% de que a chamada possa ser de um dos quatro tipos de custo. Em um caso real esta distribuição deve ser avaliada de acordo com a série histórica de chamadas.

O modelo de simulação foi executado para o caso base (sem políticas de atendimento e roteamento) para que os resultados e indicadores pudessem ser avaliados. Na tabela 4.13 estão os resultados do experimento (a mesma metodologia de realização dos experimentos anteriores foi adotada). Existe uma coluna que representa o valor em Reais das chamadas de acordo com seu tipo de custo e tempo de duração (envolve o tempo de espera na fila e o tempo do atendimento). Outra análise apresentada é quanto a utilização percentual dos atendentes, este conjunto de dados é apresentado dada característica heterogênea dos atendentes quanto ao desempenho no tempo de atendimento das chamadas. Além da média e desvio padrão do percentual de utilização dos atendentes – , o máximo e o mínimo também estão na tabela 4.13 apresentando o valor do atendente que teve maior e o que teve a menor utilização no conjunto de atendentes do caso base.

Tabela 4.13 Resultados obtidos com o modelo de simulação refletindo o funcionamento da central de teleadendimento, incluindo os valores de custo das ligações e análise do percentual de ocupação dos atendentes.

Indicadores da Central				Análise sobre percentual de ocupação dos atendentes			
Custo Total	Tempo Espera	Tempo Atendimento	Tempo Total	Média	Desvio Padrão	Máximo	Mínimo
R\$ 100,00	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%	100,00%

4.4.3.2 Políticas de atendimento e roteamento de chamadas para reduzir o custo com ligações

Apresentado o caso base, três alternativas para reduzir o custo foram propostas: estratégia de atendimento utilizando filas com prioridades, roteamento de chamadas baseado em desempenho e roteamento de chamadas baseado no custo marginal.

Fila com prioridades

Política: chamada mais cara é atendida primeiro. O modelo de simulação até então implementa uma fila com organização do tipo primeiro cliente a entrar primeiro cliente a sair.

Esta alternativa altera diretamente o indicador de tempo médio de espera na fila – W . A idéia por de trás desta alternativa é reduzir o tempo na fila dos clientes que fizeram chamadas com a prioridade do tipo 1, 2,3... respectivamente.

A alternativa de prioridades apresenta um risco que é o caso de clientes com chamadas de baixa prioridade não serem atendidos ou sempre serem “ultrapassados” na fila por clientes com chamadas de maior prioridade. Este tipo de risco faria com que os clientes de chamadas de baixa prioridade tivessem um tempo médio de espera na fila muito superior que os demais. Para resolver este tipo de problema uma condição de contorno foi adotada. Na fila sempre que o tempo de espera de qualquer cliente com qualquer tipo de chamada ultrapassa em até três vezes o tempo médio de espera na fila calculado pela central, esta chamada recebe prioridade máxima para que seja atendida o mais rápido possível. Em princípio o valor de três vezes foi estipulado por se imaginar um valor razoável de espera na fila. Posteriormente outros valores podem ser avaliados, assim como sua influência nos indicadores da central.

Roteamento baseado em desempenho - RBD

Esta alternativa visa explorar a questão advinda da característica heterogênea do grupo de atendentes. No grupo de atendentes do experimento existem atendentes que são mais eficientes quanto ao tempo médio de atendimento. Alguns são 40,00 % e outros 25,00 % mais rápidos, existem também aqueles que não possuem grau de desempenho no tempo de atendimento.

A idéia desta abordagem é simples, sempre que houver a chegada de uma chamada, esta será atribuída para o atendente com maior grau de desempenho. A figura 4.9 apresenta o pseudo-algoritmo de roteamento.

```

Enquanto houver nova_chamada
{
    atendente_selecionado = 0;
    max_desempenho = -1;
    Para cada Atendente livre m faça
    {
        se m.desempenho > max_desempenho
        então max_desempenho = m.desempenho;
            atendente_selecionado = m;
    }
    Atribua(nova_chamada, atendente_selecionado);
}

```

Figura 4.9 Pseudo-algoritmo RBD – roteamento baseado em desempenho aplicado no estudo de caso para reduzir o custo total com ligações na central de teleatendimento.

Esta abordagem apresenta um risco, ela aumenta a utilização dos atendentes com maior grau de desempenho. Os atendentes com maior desempenho serão, em média, mais utilizados que os outros (trabalham mais). Durante as investigações neste trabalho não se encontrou nenhuma referência ou citação de caso prático em que houvesse diferenças de salário ou benefícios para os atendentes com maior percentual de utilização. Este fator deve ser considerado em longo prazo, dado a alta rotatividade de pessoal nas centrais. A rotatividade das centrais (mundial) esta em torno de 22%, sendo que em alguns ramos de atuação este número chega a 50% [Sharp, 2003]. Acredita-se que esta carga de trabalho sem retorno para o trabalhador possa ser uma das causas deste nível de rotatividade.

Outro fator de longo prazo que acredita-se acontecer em função da concentração de

trabalho nos atendentes com maior desempenho é o inverso deste fato. Atendentes com menor desempenho tem menor utilização. Desta forma, na operação da central, a existência de oportunidades para que os atendentes com menor desempenho desenvolvam suas habilidades e se tornem mais eficientes não são propiciados por esta política. Pelo contrário a política foca em redução de custos e sempre privilegia a utilização de atendentes com maior grau de desempenho. Isto pode apresentar resultado em um curto prazo, mas considerando um horizonte de tempo maior a adoção desta política pode oferecer outros tipos de riscos.

Roteamento baseado em custo marginal - RBCM

A política RBCM é baseada na atribuição de chamadas para os atendentes utilizando o conceito de custo marginal. A meta desta alternativa é fornecer uma política que considere a natureza heterogênea dos atendentes quanto ao seu desempenho no tempo de atendimento das chamadas, mas que também considere no roteamento critérios para “balancear” a carga de trabalho.

O custo marginal é o aumento de custo ocasionado pela produção de uma unidade extra de produto [Dixit e Pindyck, 1994]. Em outras palavras, indica quanto custa cada aumento unitário de produção. A política aqui apresentada foi adaptada a partir do trabalho de [Barak et al., 2000] em que o conceito de custo marginal é apresentado e utilizado na alocação de tarefas em um “cluster” de computadores. Vide *Apêndice D - Algoritmo de alocação de tarefas em “cluster” de computadores baseado no custo marginal* para maiores detalhes. A adaptação foi feita para tratar as diferentes características dos atendentes e alocar as chamadas para os atendentes com o menor custo marginal no momento do roteamento da chamada.

O conceito do custo marginal adotado neste trabalho leva em consideração a variação do esforço de um atendente ao atender mais uma chamada. Seu foco é uma busca por um equilíbrio geral da carga de trabalho de todos os atendentes buscando minimizar a carga máxima de atendentes com maior desempenho. O algoritmo proposto em [Barak et al., 2000] é denominado de *ASSIGN-U*, para maiores detalhes veja o *Apêndice D – Algoritmo de alocação de tarefas em “cluster” de computadores baseado no custo marginal*. Este trabalho propõe uma adaptação do algoritmo *ASSIGN-U*. Para o RBCM uma chamada é atribuída para o atendente com o menor custo marginal. Considere

- $a = \text{constante}, 1 < a < 2;$
- $l_{ij} = \text{carga de trabalho do atendente } i \text{ antes de receber a chamada } j;$
- $p_{ij} = \text{carga de trabalho que a chamada } j \text{ adicionará para a atendente } i;$

O algoritmo executa em tempo real atribuindo a chamada j para o atendente i com o menor custo marginal:

$$H_{ij} = a^{l_{ij} + p_{ij}} - a^{l_{ij}} \quad (14)$$

Um desafio para utilização do algoritmo *ASSIGN-U* nas centrais de teleatendimento é sobre como mensurar a carga de trabalho e desempenho dos atendentes. Como as centrais avaliam o desempenho dos atendentes e qual é o valor agregado de cada chamada atendida? O desempenho de atendentes ou grupos de atendentes deve ser avaliado em conjunto com as informações de clientes atendidos [Sharp, 2003]. O trabalho de Sharp indica que é possível calcular o retorno de cada atendente ou grupo de atendentes medindo o tempo de atendimento das chamadas e o número de chamadas considerando-as sob o ponto de vista de quão

valorosas são estas chamadas. Por exemplo, um atendente que atende poucas chamadas de clientes importantes (grandes consumidores, por exemplo), provavelmente, agrega mais valor a central do que outro que atende mais chamadas com um curto tempo de atendimento para clientes que não trazem resultados para a organização. Seguindo esta afirmação, duas foram as variáveis selecionadas para representação da carga de trabalho:

- Tempo de atendimento;
- Quantidade de chamadas atendidas;

Neste sentido a carga de trabalho ficaria “balanceada” para casos, por exemplo, em que atendentes que tivessem atendido poucas chamadas com alto tempo de atendimento e para os casos de atendentes que atenderam muitas chamadas advindas de casos mais simples (pouco tempo de atendimento). Esta premissa foi adotada neste trabalho. Este item é um ponto que deve ser avaliado, pois a representação da carga de trabalho também pode envolver outros itens como: motivação do atendente, grau de cansaço, experiência entre outros

Assumido este fato, para implementação do algoritmo de roteamento baseado em custo marginal - RBCM considere um conjunto com n atendentes com as seguintes características conhecidas:

- r_{ci} - capacidade em quantidades de segundos de atendimento do atendente i ;
- r_{mi} - capacidade quanto ao número de chamadas do atendente i .

Considere também que existe para cada chamada os seguintes parâmetros:

- t_j = quantidade segundos necessários para o atendimento da chamada j . Valor em princípio não conhecido;
- m_j = identificador da chamada j , valor conhecido pois cada chamada representa uma unidade;

Sendo k um conjunto de chamadas atendidas pelo atendente i até instante t ; o total de chamadas e o tempo total de atendimento são, respectivamente, definidas por:

$$\bullet \quad l_{mit} = |k| \tag{15}$$

$$\bullet \quad l_{cit} = \sum_{j=0}^{|k|} t_j \tag{16}$$

Outra característica da proposta de Barak et al. é a capacidade de tratar características heterogêneas quando se considera a carga de trabalho, por exemplo, no caso de computadores são consideradas memória e capacidade de processamento. Para o caso dos atendentes um tratamento análogo deve ser aplicado para possibilitar a inserção na expressão (14) das duas características: quantidade de chamadas e o tempo de atendimento das chamadas (r_{mi} e r_{ci}).

Para o algoritmo ASSIGN-U a função deve referenciar a carga de uma dada característica que pode ser expressa pela: quantidade utilizada dividida por sua capacidade (utilização máxima). Sendo assim a expressão (14) poderia ser reescrita com seguinte função objetivo:

$$\text{Min}_i H_{ij} = \left(a^{\frac{l_{mit}+1}{r_{mi}}} - a^{\frac{l_{mit}}{r_{mi}}} \right) + \left(a^{\frac{l_{cit}+t_j}{r_{ci}}} - a^{\frac{l_{cit}}{r_{ci}}} \right) \tag{17}$$

Para a característica tempo no atendimento deve-se conhecer sua capacidade r_{ci} , neste sentido adotou-se o caso em que o atendente esteve *100,00%* do tempo de sua escala de trabalho realizando atendimentos (considera-se que todos os atendentes trabalham em escalas com o mesmo número de segundos), isto representa sua capacidade plena. Quanto a questão da capacidade para o número de chamadas atendidas - r_{mi} é algo que é tratado pelo próprio algoritmo que armazena o maior número de chamadas multiplicando-o por dois, esta avaliação do maior número de chamadas acontece nos momentos de roteamento das chamadas, ou seja, a cada atribuição este valor é revisto.

A figura 4.10 apresenta o pseudo-algoritmo. O tratamento descrito para os casos das características heterogêneas é apresentado. Da mesma forma também é mostrado a abordagem que utiliza o custo marginal. Em princípio o algoritmo é o mesmo apresentado em [Barak et al., 2000], no entanto, as adaptações discutidas foram inseridas para o caso do processo de roteamento de chamadas para os atendentes nas centrais de teleatendimento.

apresenta o maior desvio padrão quanto ao percentual de utilização dos atendentes. Quando comparado o valor do percentual do atendente com maior utilização e o de menor utilização existe uma diferença de $\sim 11,00\%$. Os atendentes com maior desempenho trabalham, em média, mais que os outros. Durante a realização deste trabalho não se encontrou nenhuma referência ou citação de caso prático em que houvesse diferenças de salário ou benefícios para os atendentes com maior percentual de utilização ou desempenho. Conforme esperado esta abordagem direciona a utilização dos atendentes com maior desempenho quanto ao tempo de atendimento.

Tabela 4.14 Resultados obtidos com o modelo de simulação utilizando as alternativas propostas inclui os valores de custo das ligações, indicadores e análise do percentual de ocupação dos atendentes.

Indicadores da Central							Análise sobre percentual de ocupação dos atendentes			
Alternativa utilizada	Custo Total Chamadas - (R\$)	Economia - (%)	B - (%)	W_i - (s)	δ_i - (%)	A - (%)	v - (%)	Desv. Pad. de v - (%)	Maior v - (%)	Menor v - (%)
Atual	12.764,50	0,00	0,00	7,05	94,14	0,00	92,75%	1,34%	94,59%	90,18%
Fila com prioridades	12.627,71	1,08	0,00	7,19	94,62	0,00	92,80%	1,32%	94,62%	90,28%
RBD e Fila com prioridades	12.252,59	4,18	0,00	6,15	96,58	0,00	91,16%	4,87%	96,40%	85,24%
RBCM e Fila com prioridades	12.314,00	3,66	0,00	5,84	97,72	0,00	92,31%	1,21%	94,49%	89,51%

A abordagem de custo marginal apresenta o maior nível de serviço dentre as demais. O desvio padrão do percentual de utilização dos atendentes e tempo médio de espera na fila também são os menores. Quanto ao indicador custo total com as chamadas a diferença entre esta abordagem e o roteamento baseado em desempenho ficou com uma diferença diária de R\$61,41, ainda sim gerando um economia diária de $\sim R\$ 451,00$.

Apesar da alternativa de filas com prioridades ter apresentado a menor economia na redução de custo com as ligações sua potencialidade não deve ser ignorada. Para o caso da central fictícia simulada o tempo médio de espera na fila é pequeno e observe que esta alternativa pouco alterou o indicador W_i (na verdade este indicador até aumentou). Mas para o caso de centrais em que o tempo médio de espera na fila for maior, acredita-se que esta alternativa deva ser mais significativa do que as demais, visto que ele irá reduzir o tempo de espera na fila para os clientes com maior custo de chamadas. Outros experimentos neste sentido deverão ser realizados.

5 Conclusões e comentários

Nesta dissertação a abordagem de simulação baseada em sistemas multiagente foi adotada para representar as centrais de teleatendimento. Os agentes, suas interações e a representação de seus interesses, domínios de conhecimento e objetivos foram explorados. A viabilidade da adoção deste paradigma foi verificada através de uma série de experimentos que buscaram atestar a eficácia do modelo. Além disto, situações e proposições observadas no dia a dia das centrais foram incorporadas e tratadas pelo modelo, demonstrando sua flexibilidade quanto a aderência na complexidade operacional existente nas centrais.

Como pode ser observado o modelo possui uma representação que segue uma abordagem do tipo “*bottom-up*” em que o comportamento de atendentes e clientes, implementados como agentes, influenciam o comportamento geral do sistema. Tal abordagem em modelos de simulação não foi encontrada na literatura. Os resultados aqui apresentados solidificam esta proposta de representação como uma alternativa para a área de engenharia de serviços que busca o estudo de questões como qualidade de serviço e eficiência nas centrais.

Os resultados expostos foram derivados, inicialmente, de comparações:

- Comparação com modelos analíticos do tipo “*Erlang*”, os resultados foram apresentados nas tabelas 4.3, 4.4 e 4.5 e indicam, por exemplo, uma diferença percentual para o indicador de nível de serviço - de 1,39% para o modelo *M/M/108/10* que é o que mais se assemelha ao ambiente real das centrais;
- Comparação com um caso real em que um dia de simulação foi realizado. Uma diferença percentual para o indicador de nível de serviço - de 1,38% foi obtida.

Após estes resultados comparativos o modelo de simulação foi estendido para demonstrar sua capacidade em lidar com questões reais das centrais de teleatendimento.

O comportamento de clientes relativo a taxa de paciência crescente em um mesmo dia e tipo de serviço foi “minerado” de uma base de dados, vide *subseção 4.4.1 - Simulação do abandono e o comportamento de espera dos clientes*. Para a abordagem baseada em agentes este comportamento é implementado de forma natural (comportamento particular real de um grupo de clientes, no modelo se transformou em um comportamento particular simulado de um grupo de agentes), o que para outras abordagens pode representar um desafio. A inserção deste comportamento nos agentes do tipo clientes foi realizado com intuito de avaliar seu impacto nos indicadores de desempenho da central de teleatendimento. Foi avaliado um estudo de caso que, dependendo da configuração operacional da central, a desconsideração deste comportamento pode ocasionar erros de ~7,00% no indicador de nível de serviço – . Este aspecto é importante e deve ser considerado na modelagem das centrais.

Um estudo de caso, *subseção 4.4.2 Roteamento baseado em habilidades*, apresentou as capacidades do modelo em suportar o processo de tomada de decisão. Pela utilização do modelo, e de forma empírica, foi possível experimentar diferentes cenários partindo das clássicas perguntas “e se”. A correlação entre indicadores de desempenho da central foi demonstrada alterando-se, por exemplo, as habilidades dos atendentes. A resolução de um problema em que desejava-se aumentar o nível de serviço da central sem a contratação de mão de obra adicional (atendentes) foi tratada usando uma política conhecida na literatura (roteamento baseado em habilidades).

Finalmente, na *subseção 4.4.3 Filas com prioridades e políticas de roteamento baseada em custo marginal e desempenho dos atendentes*, um experimento foi apresentado para resolver outro tipo de problema das centrais. Um aspecto com os quais os gestores das centrais devem lidar no dia a dia é o custo. Neste experimento este foi o foco. Estratégias de atendimento e políticas de roteamento foram propostas visando reduzir um dos custos das centrais de teleatendimento, o custo com as ligações. Mais uma vez o modelo apresentou sua flexibilidade e foi possível, por exemplo, adaptar um algoritmo, que é utilizado para alocação de tarefas em “*cluster*” de computadores (utiliza o conceito econômico de custo marginal), para que este fosse utilizado no roteamento de chamadas. O modelo apresentou um resultado que significaria uma redução ~4,20% nos custos diários com chamadas pagas pela central.

5.1 Trabalhos futuros

Considerando a flexibilidade do modelo de simulação baseado em sistemas multiagentes há um grande potencial de trabalhos futuros.

Um primeiro ponto sugerido é a simulação de centrais que fazem chamadas para os clientes ou mesmo centrais que são mistas, ou seja, centrais que fazem e recebem chamadas (o modelo de simulação apresentado neste trabalho considera centrais com características de chamadas entrantes).

O agente cliente descrito no modelo oferece várias possibilidades de trabalhos futuros. Importantes desafios estão relacionados aos comportamentos de paciência e ligações [Koole, 2005]. Conforme apresentado na *subseção 4.4.1 Simulação do abandono e o comportamento de espera dos clientes* o modelo ofereceu suporte para simular um comportamento específico de um grupo de clientes. Este tema é multidisciplinar e deve envolver a colaboração de outras áreas de conhecimento na busca de sua compreensão e explicação. Acredita-se que o modelo apresenta oportunidades para este tipo de trabalho.

Considerar questões sobre o agente do tipo atendente. Um aspecto a se explorar é sobre a rotatividade de pessoal, um problema real para as centrais [Sharp, 2003]. Este é um problema de longo prazo, algumas proposições de estudo são:

- Simular as paradas e intervalos programados para os atendentes. O modelo não seria orientado para a qualidade e eficácia, mas sim orientado para a realidade;
- Simular o aumento do percentual de utilização de acordo com a experiência ou diminuição da variância do tempo de atendimento com a experiência;
- Simular a rotatividade nas centrais de teleatendimento e estratégias para o problema de tomada de decisão sobre contratação de mão de obra.

Outro ponto é melhorar e desenvolver os algoritmos de alocação das chamadas para os atendentes no agente do tipo DAC. Uma idéia seria na linha da aprendizagem de máquina, por exemplo: um conjunto de políticas de roteamento estaria disponível e este agente teria a capacidade de aprender com base na experiência quando uma ou outra política é recomendada visto uma situação na central, como seria o caso de um pico de demanda de clientes. Assim, o DAC, receberia algumas diretrizes e metas estabelecidas pelos gestores da central e agiria adaptando suas funções para prover os resultados predeterminados.

Os resultados obtidos no estudo de caso da *seção 4.4.2 - Roteamento baseado em habilidades* indicam que existe um problema de otimização a ser resolvido. Quantos atendentes com que tipo de habilidade devem existir para atender as restrições de nível de

serviço. Métodos clássicos de otimização podem ser aplicados aos resultados da simulação orientando para a melhor decisão.

5.2 Comentários

A implementação com o *Swarm* se deu de forma natural. O paradigma de agentes e sua respectiva implementação no *Swarm* adapta-se ao conceito do ambiente sócio-técnico das centrais em que clientes e atendentes (modelados como agentes) interagem. Acredita-se que a introdução de novas características e ou comportamentos nos agentes, tais como religações, absenteísmo dentre outros também possam ser tratados sem dificuldades de implementação.

Ainda quanto a implementações, outro ponto é a utilização do *Swarm*, no que se refere a sua capacidade de apresentação da simulação em tempo real. O *Swarm* possibilita a apresentação de vários tipos de gráficos e ambientes em espaço bidimensional. Esta capacidade fortalece a potencialidade do modelo quanto à configuração de cenários e apresentação de resultados para análises do tipo “e se”.

Uma questão foi apresentada no *capítulo 3 - Modelo de Simulação de uma central de teleatendimento*, mas sua resposta esta além do escopo deste trabalho. A questão é:

- Existem casos em que os agentes formam organizações e são capazes de aprender de forma coletiva?

Acredita-se que haja indícios sobre efeitos de atuação de formas coletivas relacionados a: absenteísmo dos atendentes, aumento na taxa de demanda dentre outros. Este é um tipo de trabalho que, provavelmente, terá propriedades multidisciplinares e deverá envolver estudos de caso e pesquisas de campo para seu esclarecimento.

Na *subseção 3.4.1 - Simulação do abandono e o comportamento de espera dos clientes* foi descrita uma atividade de análise de dados em que o comportamento de taxa crescente de paciência em um mesmo dia e para um mesmo tipo de serviço foi descoberto e apresentado. Técnicas de mineração de dados e descoberta de conhecimento devem ser aplicadas nesta base. A descoberta de outros comportamentos ou padrões são contribuições em direção ao entendimento da dinâmica das centrais e podem ser simuladas no modelo para que seu entendimento seja compreendido sob a perspectiva de impacto nos indicadores das centrais de teleatendimento.

A proposta de roteamento baseada em custo marginal poderia ganhar um forte argumento. De acordo com os experimentos realizados na *seção 4.4.3 - Filas com prioridades e políticas de roteamento baseada em custo marginal e desempenho dos atendentes* esta política de roteamento apresenta o menor desvio padrão e menor diferença no percentual de utilização dos atendentes. Acredita-se que esta política utilizada no longo prazo diminuiria a rotatividade de pessoal e também o custo de operação das centrais (custo com demissões e desligamentos) o que justificaria a longo prazo sua utilização no lugar de outros tipos de políticas focadas em custo, como por exemplo, orientadas a desempenho dos atendentes.

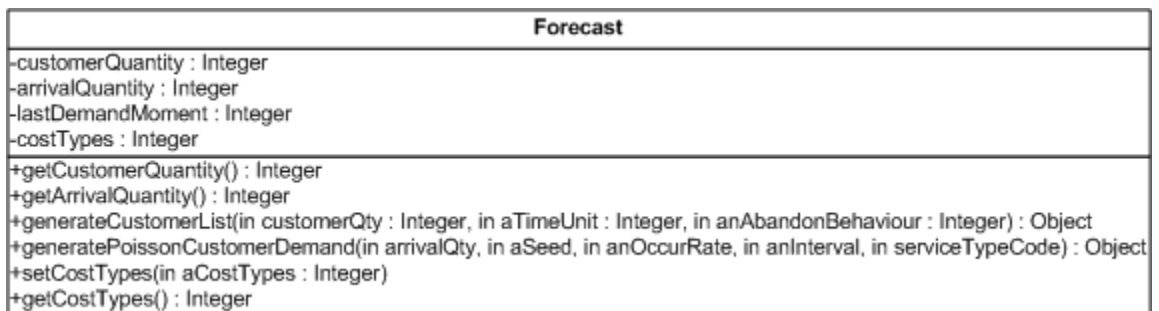
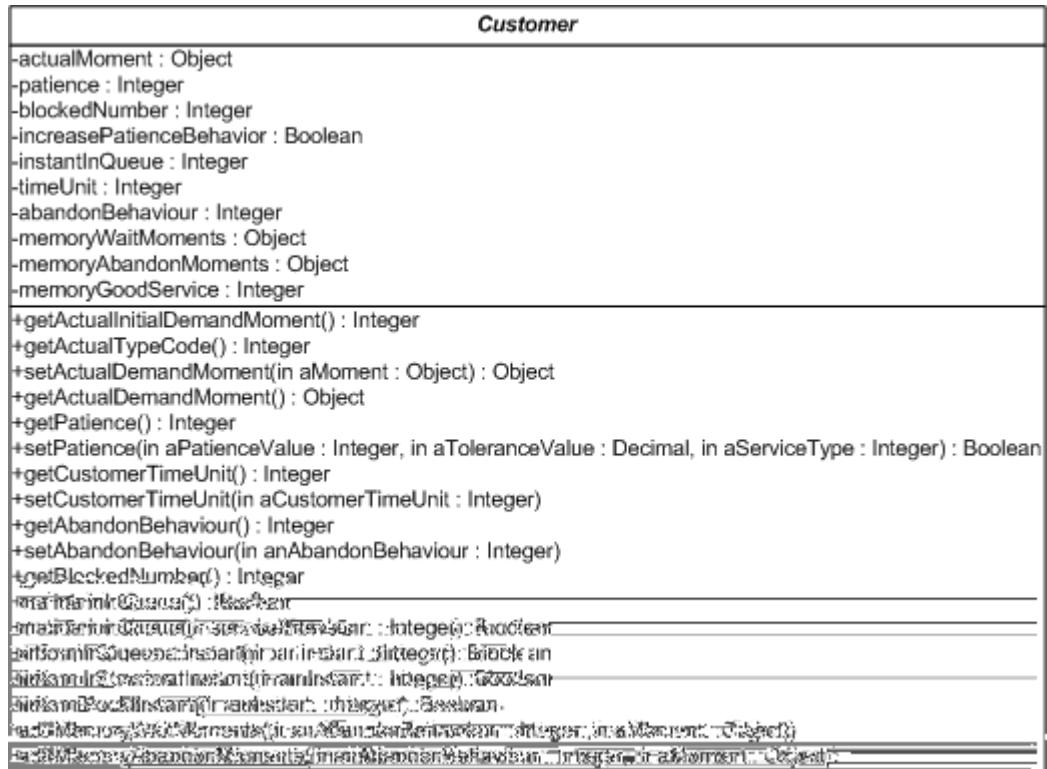
Ainda sobre o tema da *seção 4.4.3 - Filas com prioridades e políticas de roteamento baseada em custo marginal e desempenho dos atendentes* mais experimentos devem ser realizados. Estes experimentos devem avaliar outras configurações de diferentes centrais, permitindo assim entender, demonstrar e justificar a utilização de cada uma das estratégias apresentadas na referida seção.

O tópico políticas de roteamento é um item relevante nas centrais e pode ser tratado não só

com a proposição de novas políticas de roteamento de chamadas. Alguns trabalhos comentam sobre o roteamento de chamadas em centrais de teleatendimento virtuais em que suas unidades físicas são transcontinentais [Mandeulbaum, 2005]. Estratégias de roteamento para estes e outros casos podem ser explorados no modelo de simulação.

Outro ponto é sobre a implementação do agente gestor, que não foi implementado neste trabalho. Este ponto, em particular, já está sendo tratado pela equipe que também ajudou a desenvolver este trabalho. Técnicas como coordenadas descendentes e busca tabu já estão sendo avaliadas na implementação deste agente.

Apêndice A – Diagrama de classes do modelo



Queue
-queueList : Object -inputQuantity : Integer -outputQuantity : Integer -askedServiceQuantity : Object -actualLength : Integer -goodWaitTimeFactor : Integer -goodWaitTimeAccomplishedNumber : Object -blockingFactor : Integer -blockingNumber : Object -totalBlockingNumber : Integer -abandonNumber : Object -totalAbandonNumber : Integer -serviceTypeNumber : Integer -timeByCosts : Object -waitTime : Object -serviceDist : Object -timeUnitModel : Integer
+addQueueLastSBR(in aCustomer : Object) +removeQueueFirstSBR() +printQueueOrderByCosts() +searchInQueueServiceCodeType(in codeType : Integer) : Boolean +negotiateWithCustomersInQueue() +setServiceDistributor(in anObject) +getServiceDistributor() : Object +getQueueList() : Object +getWaitTime(in serviceType : Integer) : Decimal +getServiceTotalTime() : Decimal +getblockingFactor() : Integer +setblockingFactor(in anObject, in anFactor : Integer) +getblockingNumber(in anObject, in serviceType : Integer) : Integer +getblockingNumber() : Integer +getGoodWaitTimeFactor() : Integer +setGoodWaitTimeFactor(in anFactor : Integer) +getActualLength() : Integer +changeCustomerPositionOnQueue() +getAskedServiceQuantity(in serviceType) : Integer +getGoodServiceAccomplishedNumber(in serviceType) : Integer +getServiceAccomplishedNumber() : Integer +getAverageServicesWorkTime() : Decimal +getTimeByCostType(in aCostType : Integer) : Long

Service
-actualServiceTime : Long -serviceFinishedEstimatedTime : Decimal -workTime : Long -workTimeByPeriod : Long -numberOfCalls : Integer -numberOfCallsByPeriod : Integer -timeWithoutInterval : Integer -skillNumber : Integer -actualskillNumber : Integer -skillArray : Object -performance : Decimal
+isAvailable() : Boolean +getActualServiceTime(in serviceTime : Long, in currentTime : Long) : Long +getServiceFinishedEstimatedTime() : Decimal +getPerformance() : Decimal +getWorkTime(in serviceTime : Long, in currentTime : Long) : Long +getWorkTimeByPeriod(in serviceTime : Long, in currentTime : Long) : Long +getNumberOfCalls() : Integer +getNumberOfCallsByPeriod() : Integer +getTimeWithoutInterval() : Integer +getSkill(in skillLevel : Integer) : Integer +addSkill(in skillCode : Integer)

ServiceDistributor
<pre> -serviceTotalTime : Double -serviceTotalNumber : Long -timeSeed : Object -meanTime : Object -serviceTimeGenerator : Object -timeDist : Object -informWaitMoments : Object -busyServicesList : Object -busyServiceIndex : Object -idleServicesList : Object -idleServiceIndex : Object -distServiceNumber : Integer -timeByCosts : Object -skillMemoryofServices : Object -skillNumber : Integer -SkillCSV : Object -timeUnitModel : Integer -maxNumberOfCalls : Integer -routeMode : Integer </pre>
<pre> +isthereServiceAvailable() : Boolean +getServiceTime(in serviceType : Integer) : Long +setService(in skillCode : Integer, in aCostCode : Integer) : Boolean +verifyService() : Boolean +getServiceTotalTime() : Decimal +getServiceTotalNumber() : Long +getServiceNumber() : Integer +setServiceNumber(in serviceNum : Integer, in fName : Object) +getAverageWorkTime() : Decimal +printUtilizationByService(in aFile : Object) +setTimeSeed(in seed : Decimal, in serviceType : Integer) +getTimeSeed(in serviceType : Integer) : Decimal +setRouteMode(in aRouteMode : Integer) +getRouteMode() : Integer +setMeanTime(in mean : Decimal, in serviceType : Integer) +getMeanTime(in serviceType : Integer) : Decimal +setWaitMoments(in inform : Integer, in serviceType : Integer) +getWaitMoments(in serviceType : Integer) : Integer +printskillMemoryofServices() +updateMemory(in aService : Object, in anOperation : Char) : Boolean +searchServiceBySkill(in skillCode : Integer) : Object +searchServiceBySkillAndPerformance(in skillCode : Integer) : Object +searchServiceBySkillAndMarginalCost(in skillCode : Integer, in aServiceTime : Long) : Object +getTimeByCostType(in aCostType : Integer) : Long </pre>

Apêndice B – Base de dados e programas utilizados para obtenção de resultados

Descritivo da base de dados

```
CREATE TABLE [calls] (
    [vru+line] [char] (6) NOT NULL ,
    [call_id] [char] (5) NOT NULL ,
    [customer_id] [varchar] (12) NULL ,
    [priority] [char] (1) NULL ,
    [type] [char] (2) NULL ,
    [date] [smalldatetime] NULL ,
    [vru_entry] [smalldatetime] NULL ,
    [vru_exit] [smalldatetime] NULL ,
    [vru_time] [int] NULL ,
    [q_start] [smalldatetime] NULL ,
    [q_exit] [smalldatetime] NULL ,
    [q_time] [int] NULL ,
    [outcome] [varchar] (7) NULL ,
    [ser_start] [smalldatetime] NULL ,
    [ser_exit] [smalldatetime] NULL ,
    [ser_time] [int] NULL ,
    [server] [varchar] (255) NULL ,
    CONSTRAINT [PK_calls] PRIMARY KEY
    (
        [vru+line],
        [call_id]
    )
)
```

Amostra da base de dados

vru+line	call_id	customer_id	priority	type	date	vru_entry
AAO101	33116	9664491	2	PS	1999-01-01 00:00:00	1999-01-01 00:00:00
AAO101	33117	0	0	PS	1999-01-01 00:00:00	1999-01-01 00:34:00
AAO101	33118	27997683	2	PS	1999-01-01 00:00:00	1999-01-01 06:55:00
AAO101	33119	0	0	PS	1999-01-01 00:00:00	1999-01-01 07:41:00
AAO101	33120	0	0	PS	1999-01-01 00:00:00	1999-01-01 08:03:00
AAO101	33121	0	0	PS	1999-01-01 00:00:00	1999-01-01 08:18:00
AAO101	33122	0	0	PS	1999-01-01 00:00:00	1999-01-01 08:28:00

vru_exit	vru_time	q_start	q_exit	q_time	outcome
1999-01-01 00:00:00	5	1999-01-01 00:00:00	1999-01-01 00:03:00	153	HANG
1999-01-01 00:34:00	11	1999-01-01 00:00:00	1999-01-01 00:00:00	0	HANG
1999-01-01 06:55:00	6	1999-01-01 06:55:00	1999-01-01 06:55:00	17	AGENT
1999-01-01 07:41:00	10	1999-01-01 00:00:00	1999-01-01 00:00:00	0	AGENT
1999-01-01 08:03:00	10	1999-01-01 00:00:00	1999-01-01 00:00:00	0	AGENT
1999-01-01 08:18:00	9	1999-01-01 00:00:00	1999-01-01 00:00:00	0	AGENT
1999-01-01 08:28:00	10	1999-01-01 00:00:00	1999-01-01 00:00:00	0	AGENT

ser_start	ser_exit	ser_time	server
1999-01-01 00:00:00	1999-01-01 00:00:00	0	NO_SERVER
1999-01-01 00:00:00	1999-01-01 00:00:00	0	NO_SERVER
1999-01-01 06:55:00	1999-01-01 06:56:00	54	MICHAL
1999-01-01 07:41:00	1999-01-01 07:44:00	208	BASCH
1999-01-01 08:03:00	1999-01-01 08:05:00	107	MICHAL
1999-01-01 08:18:00	1999-01-01 08:23:00	275	KAZAV
1999-01-01 08:28:00	1999-01-01 08:30:00	102	KAZAV

Programas para obtenção de resultados e padrões – SQL (*Structure Query Language*)

```

/*Do conjunto total de chamadas quantos clientes
fizeram mais de uma chamada para a central
de teleatendimento? */
SELECT customer_id, COUNT(*) as CHAMADAS
FROM dbo.calls
WHERE customer_id <> '0'
GROUP BY customer_id
HAVING COUNT(*) >1
ORDER BY 2 DESC

select CHAMADAS, COUNT(*) AS CLIENTES
FROM (SELECT customer_id, COUNT(*) as CHAMADAS
FROM dbo.calls
WHERE customer_id <> '0'
GROUP BY customer_id
HAVING COUNT(*) >1) AS T
GROUP BY CHAMADAS
ORDER BY 1,2 DESC

/* Do conjunto total de chamadas quantos clientes
fizeram mais de uma chamada para a central de
teleatendimento? */
SELECT customer_id, COUNT(*) as CHAMADAS
INTO Cliente_maior_Chamadas
FROM dbo.calls
WHERE customer_id <> '0'
GROUP BY customer_id
HAVING COUNT(*) >1

select customer_id, count(*) AS abandono
into cliente_abandono
FROM dbo.calls
where customer_id in (SELECT customer_id
FROM Cliente_maior_Chamadas)
and outcome='HANG'
group by customer_id

select perc_abandono, count(*) as clientes
from (SELECT a.customer_id, CHAMADAS,
abandono, ((abandono*100)/CHAMADAS) as perc_abandono
FROM dbo.Cliente_maior_Chamadas as c, cliente_abandono as a
where c.customer_id=a.customer_id ) as t
group by perc_abandono
order by 2 desc

/*Dos clientes com algum evento de abandono,
quantos abandonaram mais de uma vez? */

select count(*) from cliente_abandono
where abandono > 1

/*Se eles abandonaram mais de uma vez, eles
abandonaram enquanto solicitavam o mesmo serviço? */

```

```

select a.customer_id, a.abandono,c.type,
count(*) from cliente_abandono a, calls c
where c.customer_id= a.customer_id
and c.outcome='hang'
and a.abandono > 1
group by a.customer_id, a.abandono,c.type
--having count(*) > 1
order by 1

/*Dos clientes com reincidência de abandono como
foi a "paciência" para cada chamada? Não há padrão,
ou a paciência foi reduzida ou aumentada com o
tempo?*/

select a.customer_id, a.abandono as abandono_total,
c.type, month(c.[date]) as mes,day(c.[date]) as dia,
q_time, count(*) as abandono_service
into abandon_servicetype
from cliente_abandono a, calls c
where c.customer_id= a.customer_id
and c.outcome='hang'
and a.abandono > 1
group by a.customer_id, a.abandono,
c.type,month(c.[date]),day(c.[date]),q_time
order by 1

/*****
--Programa que avalia o crescimento crescente de paciência dos clientes
*****/

USE SIMUCT
GO

-- Declaração de variáveis para armazenar os valores retornados pelo FETCH.
DECLARE
@customer_id varchar(12),
@customer_id_ANT varchar(12),
@abandono_total int,
@type char(2),
@type_ANT char(2),
@mes int,
@mes_ANT int,
@dia int,
@dia_ANT int,
@q_time int,
@q_time_ANT int,
@abandono_service int,
@cont int,
@contAX int

DECLARE ab_cursor CURSOR FOR
SELECT * FROM abandon_servicetype
ORDER BY 1,3,4,5

OPEN ab_cursor

```


Apêndice C – Considerações sobre simulação

Alguns modelos são utilizados no processo de tomada de decisão nas centrais de teleatendimento, essencialmente modelos simbólicos analíticos. Estes modelos apresentam algumas limitações dada a complexidade operacional inerente das centrais. A simulação é uma técnica aderente para solução destas limitações [L'Ecuyer e Avyramidis, 2005]. Este apêndice apresenta alguns conceitos, vantagens, desvantagens e tipos de simulação.

Introdução aos modelos de simulação

A palavra modelo é um termo utilizado em várias áreas de conhecimento: estatística, computação, engenharias, dentre outros. Com uma visão oriunda da Pesquisa Operacional o termo modelo é empregado em abordagens para resolução de problemas na tomada de decisão. A figura C.1 apresenta um esquema com macros passos que pode ser utilizado na criação de modelos [Goldbarg e Luna, 1999].

Existem vários tipos de modelos, eles podem ser classificados em [Perros, 2003] :

- i) Análogos (exemplos: sistema de fluídos utilizado para representar o tráfego urbano)
- ii) Icônicos (exemplos: mapas, protótipos de carros dentre outros) e
- iii) Simbólicos (exemplos: equações matemáticas ou softwares).

Um modelo de simulação é usado, na vida real, para estudar sistemas ainda não existentes ou sistemas em que não há viabilidade de configuração ou alteração de suas características (exemplo: configuração de máquinas e equipamentos em uma fábrica, nova forma de atendimento em um banco dentre outros). Normalmente o foco do estudo da simulação está em quantificar indicadores de desempenho do sistema e avaliar o grau de influência de suas variáveis de entrada nestes indicadores. Tipicamente, modelos de simulação são simbólicos e estocásticos e via de regra desempenham papel importante para o processo de tomada de decisão em seus vários níveis (estratégico e tático). Modelos de simulação podem também ser classificados em: a) contínuos (depende de variáveis que assumem valores em um domínio contínuos tais como o conjunto de números reais) ou discretos (depende de variáveis que assumem valores pertencentes a um domínio de valores finitos, números inteiros) e b) estáticos (o modelo não varia ao longo do tempo) ou dinâmicos (o modelo é capaz de representar sua evolução ao longo do tempo).

Valor agregado dos modelos de simulação

Os modelos de simulação são capazes de trazer valor agregado para os usuários que fazem uso destes para tomada de decisão. Os seguintes itens podem ser citados neste sentido [Chung, 2004]:

- Entendimento claro e detalhado sobre a operação do sistema: alguns sistemas são em essência complexos, ou seja, seu estudo e compressão é muito difícil sem o entendimento de seus componentes ou subsistemas e suas interações. Como exemplo imagine um sistema de transporte logístico multimodal (rodoviário, ferroviário, marítimo e aeronáutico). Para estes tipos de sistema um modelo dinâmico é essencial para entendimento claro de seu funcionamento.
- Criação de métodos de operação e/ou políticas de utilização de recursos para melhorar

o desempenho do sistema: mesmo quando já exista um entendimento do sistema, ainda sim pode-se adicionar melhorias ao mesmo. Algumas das oportunidades de melhorias estão relacionadas a alteração nas formas de operação ou mesmo nas políticas de utilização de recursos. Como exemplo em um sistema logístico multimodal avalia-se a programação, controle e utilização dos recursos envolvidos.

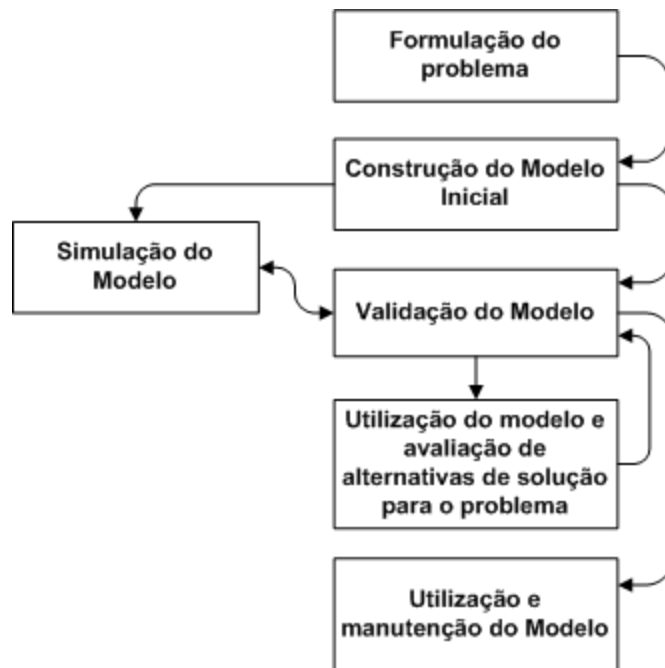


Figura C.1: Esquema com macros passos utilizado na criação de modelos para processos de tomada de decisão, adaptado de [Goldbarg e Luna, 1999].

- Teste de novos conceitos e seus impactos antes da execução destas na prática: quando o sistema ainda não está em execução ou mesmo implementado a simulação é uma ferramenta que auxilia na avaliação de impactos no caso de sua adoção ou mesma na proposição de novas formas ou conceitos. Novamente no sistema logístico pode se implementar um conceito de maximar o uso de linhas ferroviárias, no entanto qual o impacto desta alteração no tempo de entrega para os clientes, por exemplo.
- Obtenção de informação sem causar impactos ao funcionamento do sistema real: ainda no exemplo citado no item acima é possível obter informações da referida alteração com nenhuma afetação no sistema vigente.
- Facilidade de uso e construção: modelos de simulação normalmente possuem requisitos analíticos menores do que outros tipos de técnicas utilizadas na construção de modelos simbólicos. Com o advento da simulação computacional e com ferramentas específicas para construção de modelos a necessidade de profundos conhecimentos analíticos têm facilitado a disseminação desta técnica entre usuários. Um breve resumo sobre a história da evolução da simulação pode ser encontrada em [Sawhney et al., 2003].
- Realização de experimentos em tempo reduzido: normalmente os modelos de simulação são implementados computacionalmente. Avaliações que poderiam levar muito tempo na prática com o modelo de simulação podem ser realizados no computador mais rapidamente, reduzindo assim o tempo de experimentação.

- Facilidade de demonstração do modelo: muitas ferramentas de simulação possuem módulos gráficos com representação em duas ou em até mesmo três dimensões (uma apresentação resumida das ferramentas: *SIMPAK*, *ARENA* e *AutoStat/Automod* pode ser encontrado em [Chung, 2004]). Estas funcionalidades simplificam a interação do usuário com o modelo e fazem com que os resultados possam ser facilmente demonstrados.

Desvantagens

Conforme citado anteriormente existem outras técnicas utilizadas na criação de modelos simbólicos. Algumas desvantagens na utilização de modelos de simulação [Chung, 2004]:

- Não produz resultados assertivos quando as variáveis de entrada contém erros: o estudo das variáveis de entrada do modelo e a obtenção destes dados é um passo importante na construção do modelo; e em algumas vezes quando há a sua inacurácia pode acontecer um retrocesso para etapa anterior (formulação do modelo – veja figura C.1 do apêndice). Se não houver qualidade e confiabilidade nos dados de entrada, não haverá qualidade e confiabilidade nas respostas fornecidas pelo modelo de simulação.
- Não fornece respostas simples para problemas complexos: de acordo com Chung é provável que para problemas complexos respostas complexas sejam geradas. Neste caso vale ressaltar que a simulação é um instrumento de auxílio para a tomada de decisão provendo um melhor entendimento de todo sistema, suas variáveis e o grau de influência das mesmas nos indicadores de desempenho, ou seja, para problemas complexos o conhecimento dos especialistas ou a fusão da simulação com outras técnicas são necessárias para efetiva resolução do problema.
- Não possui a capacidade de resolver problemas: novamente destaca-se a necessidade do conhecimento para resolução de problemas complexos. A utilização da simulação é análoga a um “*Cyborg*”, ou seja, os especialistas e usuários poderão contar com instrumental capaz de dar informações que antes os mesmos não detinham. Destaca-se aqui outra vez a utilização da simulação com outras técnicas clássicas para solução de problemas.

Apêndice D – Algoritmo de alocação de tarefas em “cluster” de computadores baseado no custo marginal

O trabalho de [Barak et al., 2000] considera um “cluster” com n máquinas (estações de trabalho) com as seguintes características:

- r_{ci} - velocidade da CPU na máquina i
- r_{mi} - quantidade de memória na máquina i

Considere também que existe uma seqüência de tarefas que possuem os seguintes parâmetros:

- a_j = indicador de tempo da chegada da tarefa j
- t_j = quantidade de CPU requerida pela tarefa j
- m_j = quantidade de memória requerida pela tarefa j

Sendo $J(t,i)$ um conjunto de tarefas na máquina i em um dado instante t ; a carga da CPU e quantidade de memória podem ser, respectivamente, definidas por:

- $l_{cti} = |J(t,i)|$
- $l_{mti} = \sum_{j \in J(t,i)} m(j)$ (18)

O tempo para completar uma dada tarefa, $c(j)$, deverá atender a seguinte equação:

$$\int_{a_j}^{c_j} \frac{r_{ci}}{l_{cti}} = t_j \quad (19)$$

O “slowdown” ou a lentidão, atraso da tarefa será:

$$\frac{c_j - a_j}{t_j} \quad (20)$$

Por exemplo, se o tempo para completar uma tarefa $c(j)$ é de 90ms e o instante da chegada $a(j)$ com 0ms e a CPU requerida como 30ms. Então o atraso será da ordem de 3.

O trabalho de Barak et al. se resume em um método para atribuição de tarefas inspirado em modelos econômicos para minimizar o “slowdown” de todas as tarefas. Em essência há uma busca por um equilíbrio geral da carga de trabalho das máquinas buscando minimizar a carga máxima das máquinas do “cluster”. A figura 4.10 reflete esta idéia. São apresentados dois gráficos (sem a utilização da proposta de custo marginal e outros com a utilização da proposta) nos eixos horizontais têm-se o número x de máquinas (estações de trabalho) e no outro eixo uma quantidade y de carga de trabalho.

O algoritmo proposto em [Barak et al., 2000] é denominado de ASSIGN-U que atribui cada tarefa para a máquina que possui o menor custo marginal. Considere:

- $a = \text{constante}, 1 < a < 2$;
- l_{ij} - carga da máquina i antes de receber a tarefa j ;
- p_{ij} - carga que a tarefa j adicionará para a máquina i ;

O algoritmo executa em tempo real atribuindo a tarefa j para a máquina i que minimiza o custo marginal:

(21)

Figura D.1 Comparação utilizando exemplo de um “cluster” fictício sem (gráfico 1) e com (gráfico 2) a

outras abordagens e é aplicado na prática nos softwares para suporte a “clusters” *Mosix* e *Open Mosix*.

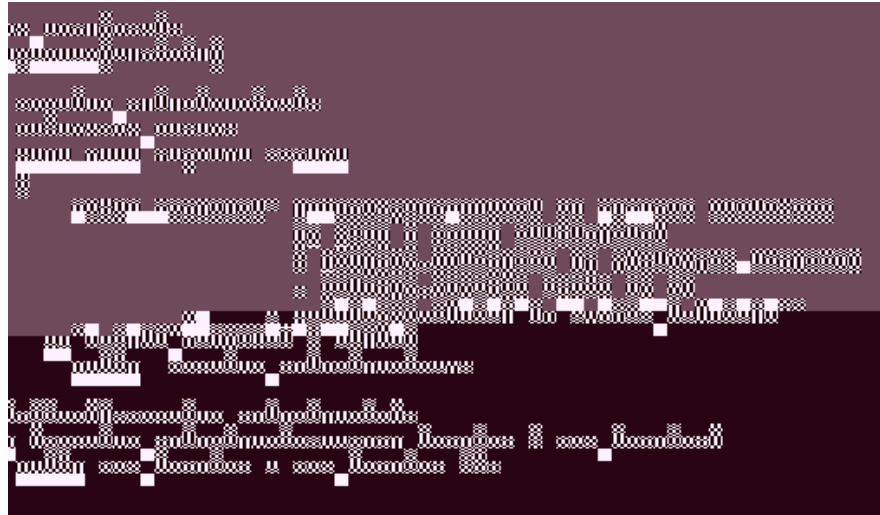


Figura D.2 Algoritmo ASSIGN-U para alocação de recursos baseado no conceito de custo marginal, adaptado da proposta apresentada em [Barak et al., 2000].

Anexo I – Mensagens eletrônicas enviadas e recebidas

Date: Thu Oct 26 16:57:59 2006
From: Avi Mandelbaum <avim@ie.technion.ac.il>
 [[Add to Address Book](#) | [Block Address](#) | [Report as Spam](#)]
To: <clm@excite.com>

ank Call-Center Data

rk and data have been of value to you. Please
 ay you see fit - I would also be happy to
 done

eague Professor in Brazil, also interested
 vice-systems. We have been talking about a
 l, so perhaps to two of us shall even meet

=====
 2 4 829-5688
 }-4504
 iling Address:
 l Engineering and Management
 NION, Haifa 32000, ISRAEL
 =====

On Thu, 26 Oct 2006, wrote:

> Dear Professor MandeulBaum, I am a mastering student in the department
 > of Computer Science at Catholic University of Minas Gerais, Belo
 > Horizonte, Brazil. I have been read your articles about Call center and
 > service engineering. Particularly I have given attention to one
 > "Empirical Analysis of a Call Center". In our thesis we are developing
 > an artificial intelligence approach to treat some decision problems in
 > call center. In this way your work have been very valued. We would like
 > to use the "Anonymous Bank" Call-Center Data in some experiments.
 > Available at:
 > <http://iew3.technion.ac.il/~serveng/course2006spring/callcenterdata/index.html>
 > This data will help us in your models.

Regards,Claudio Lucio

6 Bibliografia

- [Barak et al., 2000] Barak, A.; Yair, A.; Awerbuch, B.; Borgstrom, A; Keren, A.. An opportunity cost approach for job assignment in a scalable computing Cluster. *IEEE Transactions on parallel and distributed systems*.v.11,n.7.p.760-767. 2000.
- [Barrand e Feigenbaum, 1981] Barrand, A. Feigenbaum, E.A. *The Handbook of Artificial Intelligence*, volume I. William Kaufmann Inc., Los Altos, California, 1981.
- [Betts et al., 2000] Betts, A.; Meadows.M.; Walley,P. Call centre capacity management, *International Journal of Service Industry Management*. p. 185–196. 2000.
- [Bittencourt, 2001] Bittencourt, G. *Inteligência Artificial Ferramentas e Teorias*. Editora da UFSC, ISBN 85-328-0138-2, Florianópolis, SC, 2a edição. p.362. 2001.
- [Bittencourt, 2005] Bittencourt, G. *Introdução aos Sistemas Multiagentes Histórico e Fundamentos*. Porto Alegre: UFRGS, 2005. Palestra proferida no Brazil Agents School 2005, patrocinado pela SBC, em Porto Alegre, em 21/09/2005.
- [BRASIL, 2004] BRASIL. Agência Nacional de Energia Elétrica - ANEEL. Resolução Normativa 57, de 12 de Abril de 2004. Estabelece as condições de atendimento por meio de Central de Teleatendimento das concessionárias ou permissionárias, critérios de classificação dos serviços e metas de atendimento. *Diário Oficial* de 13/04/2004, seção 1 , p.48, v.141, n.70.
- [Brazier et al., 1999] Brazier, F.M.T.; Jonker; C.M.; Jungen, F.J.; Treur J.. Distributed scheduling to support a call center: a cooperative multiagent approach, *Applied Artificial Intelligence*, 13 (1–2). p.65–90.1999.
- [Brown et al., 2005] Brown, L.; Gans, N.; Mandelbaum, A.; Sakov, A.; Shen, H.; Zeltyn, S.; Zhao, L.. Statistical analysis of a telephone call centre: a queueing-science perspective. *Journal of the American Statistical Association*. p.36–50. 2005.
- [Bussmann e Sieverding, 2001] Bussmann, S.; Sieverding, J. *Holonic Control of an Engine Assembly Plant – An Industrial Evaluation*. In: *Proceedings of IEEE Conf. on SMC*. Tucson, AZ. 2001.
- [Castelfranchi et al., 1991] Castelfranchi, C., Miceli, M., and Cesta, A.. Dependence relations among autonomous agents. In Werner, E. and Demazeau, Y., editors, *Decentralized A.I. 3º Proceedings of the Third European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, 5–7 Ago.,Kaiserslautern, Germany, p. 215–227. 1991.

- [Charniak e McDermott, 1985] Charniak, E.; McDermott, D. Introduction to Artificial Intelligence. Addison-Wesley Publishing Company, Reading, MA, 1985.
- [Chung, 2004] Chung, C. Simulation modeling handbook : a practical approach. ISBN 0-8493-1241-8. Washington, D.C. p.17-34.2004.
- [Coelho, 2005] Coelho, H. Overview of BDI Agents. Porto Alegre: UFRGS, 2005. Palestra proferida no Brazil Agents School 2005, patrocinado pela SBC, em Porto Alegre, em 21/09/2005.
- [Cohen e Levesque, 1990] Cohen, P. R.; Levesque, H. J. Rational interaction as the basis for communication. In Cohen, P. R., Morgan, J., and Pollack, M. E., editors, Intentions in Communication, cap. 12, p. 221–255. MIT Press, Cambridge, MA, 1990.
- [Daniels, 1999] Daniels, M. Integrating Simulation Technologies with Swarm. Agent Simulation: Applications, Models and Tools, Argonne National Laboratory, Universidade de Chicago. Out., 1999.
- [David, 1998] David, N.. Modeling and implementing AND and OR Dependencies in social reasoning. Ph.D Thesis. Faculty of Sciences. Universidade de Lisboa, Portugal. 1998.
- [Davidsson, 2000] Davidsson, P. Multi Agent Simulation: Beyond Social Simulation, Editors S. Moss and P. Davidson: MABS 2000, p. 97-107. 2000.
- [Deery et al., 2002] Deery, S.; Iverso, R.; Walsh, J. Work relationships in telephone call centres: Understanding emotional exhaustion and employee withdrawal. The Journal of Management Studies. v.39 ,n.(4), p.471–496. 2002.
- [Dixit e Pindyck, 1994] Dixit, A.; Pindyck, R. Investment Under Uncertainty, Princeton University Press, Princeton NJ (USA), 1994.
- [Dubiel e Tsimhoni, 2005] Dubiel, B.; Tsimhoni, O.; Integrating Agent based modeling into a discrete event simulation. Proceedings of the 2005 Winter Simulation Conference M.E. Kuhl, N.M. Steiger, F.B. Armstrong, and J.A. Joines, eds. p. 1029–1037. 2005.
- [Durfee, 1991] Durfee, E.H. The distributed artificial intelligence melting pot. IEEE Transactions on Systems, Man and Cybernetics. v.21, n.6., p.1301.1306 Special Issue on Distributed Artificial Intelligence. Nov., 1991.
- [Enoki et al., 2001] Enoki, H.; Kitamura, Y.; Tatsumi, S.; Kitamura, S.. Job allocation mechanism to support a call center: multiagent approach based on market economic model. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, Tucson. p.1119–1124. 2001.
- [Erlang, 1917] Erlang, A.K. Solutions of some problems in the theory of probabilities fo significance in automatic telephone exchanges. Electroteknikeren.v.13;p.5-13, 1917.
- [Gans et al., 2006] Gans, N.; Koole, G.; Mandaulbaum, A.. Telephone Call Centers: a

- Tutorial and Literature Review. Disponível em: <www.columbia.edu/~ww2040/tutorial.pdf>. Acesso em: 16 maio 2006.
- [Goldbarg e Luna, 1999] Goldbarg, M. C.; Luna, H. P. L., Otimização Combinatória e Programação Linear Modelos e Algoritmos, Editora Campus, ISBN 85-352-0541-1.2a edição. 1999.
- [Hillier e Lieberman, 2001] Hillier, F.S.; Lieberman, G.J.; Introduction to Operations research. Editora McGraw-Hill, ISBN 0-07-232169-5, 7a edição. p.838. 2001.
- [<http://www.callcenter.inf.br>] ESTATÍSTICAS: Mercado brasileiro caminha para maturidade, 2007. Disponível em <[http://www.callcenter.inf.br/default.asp?sp=materia_integra.asp &secao=8&codigo=7211](http://www.callcenter.inf.br/default.asp?sp=materia_integra.asp&secao=8&codigo=7211)>. Acesso em: 27 jan. 2007.
- [JADE, 2005] JADE. Disponível em: <<http://jade.tilab.com/>>. Acesso em: 05 Set. 2005.
- [Jennings et al., 1998] Jennings, N.R.; Sycara, K.; Wooldridge, M. A Roadmap of Agent Research and Development. Autonomous Agents and Multi-Agent Systems. Boston, p.275–306, 1998.
- [Johnson e Lancaster, 2006] Johnson, P. Lancaster, A. Swarm User Guide. Disponível em: <<http://www.swarm.org/swarmdocs-2.1.1/userbook/userbook.html>>. Acesso em: 01 fev. 2006.
- [Kohler et al., 2005] Kohler, T.A.; Gumerman, J.G.; Reynolds , R.G. Simulating Ancient Societies. Scientific American. p.77-84. Jul., 2005.
- [Klungle, 1999] Klungle, R. Simulation of a claims call center: a success and a failure. WSC'99. 1999 Winter Simulation Conference Proceedings. 'Simulation—A Bridge to the Future', IEEE, Piscataway, p. 1648–1653. 1999.
- [Koole e Mandelbaum, 2002] Koole, G.; Mandelbaum, A. Queueing models of call centers: An introduction. Annals of Operations Research. First Madrid Conference on Queueing Theory. p.41–59, 2002.
- [Koole, 2005] Koole, G. Call Centers Mathematics. 2005. Disponível em <<http://www.math.vu.nl/~koole/ccmath>>. Acesso em: 24 fev. 2005.
- [Koole, 2007] Koole, G. Erlang X Calculator. 2005. Disponível em <<http://www.math.vu.nl/~koole/ccmath/ErlangX/index.php>>. Acesso em: 4 fev. 2007.
- [L'Ecuyer e Avyramidis, 2005] L'Ecuyer, P.; Avyramidis, A. N.; Modelling and simulation of call center. Proceedings of the 2005 Winter Simulation Conference M.E. Kuhl, N.M. Steiger, F.B. Armstrong, and J.A. Joines, eds., p. 144–151. 2005.
- [Lesser, 2005] Lesser, V.R. Cooperation and Coordination in Multi-Agents Systems. Porto Alegre: UFRGS, 2005. Palestra proferida no Brazil Agents School 2005, patrocinado pela SBC, em Porto Alegre, em 21/09/2005.

- [Lopes et al., 2006] Lopes, C.L.; Ferreira, R.P.M.; Alves, H. B. V. R. Simulação Multiagentes aplicada no apoio à tomada de decisão e no problema de roteamento baseado em habilidades em centrais de teleatendimento. In.:SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 28,2006, Góias. Anais... Goiânia: Universidade Católica de Góias, p. 2225 -2236. 2006.
- [Macal e North, 2005] Macal, C. M.; North, M. J.; Tutorial on Agent Based Modeling and Simulation. Proceedings of the 2005 Winter Simulation Conference M.E. Kuhl, N.M. Steiger, F.B. Armstrong, and J.A. Joines, eds.2005.
- [Malin et al., 2000] Malin, J. T; Kowing, J.; Schreckenghost, D.; Bonasso, R. P.; Nieten, J.; Graham, J. S.; Fleming, L. D.; MacMahon, M.; Thronesbery, C. Multi-agent diagnosis and control of an air revitalization system for life support in space. In Proceedings of the IEEE Aerospace Conference, USA. Março, p. 309–326. 2000.
- [Mandaulbaum, 2005] Mandaulbaum, A. Call Centers (centres) Research Bibliography with Abstracts, 2004. Disponível em: <<http://iew3.technion.ac.il/serveng/References/ccbib.pdf>>. Acesso em: 08 Out. 2005.
- [Mandaulbaum et al., 2006] Mandaulbaum, A. Sakov, Anat. Zeltyn, S. Empirical Analysis of a call center. Março, 2001. Disponível em: <<http://iew3.technion.ac.il/serveng/References/ccdata.pdf>>. Acesso em: 08 mar. 2006.
- [Mandaulbaum, 2006] Mandaulbaum, A.. Informações Simulação. [mensagem pessoal]. Mensagem recebida por <clvl@excite.com> em 26 Out. 2006.
- [Mazzuchi e Wallace, 2004] Mazzuchi, T. A; Wallace, R. B. Analyzing Skill-Based Routing Call Center Using Discrete-Event Simulations and Design Experiment. Proceedings of the 2004 Winter Simulation Conference R. G. Ingalls I, M.D. Rossetti, J.S. Smith, and B.A. Peters, eds. 2004.
- [McCarthy e Hayes, 1969] McCarthy, J; Hayes, P. J. Some philosophical problems from the standpoint of artificial intelligence. In D. Michie and B. Meltzer, editors, Machine Intelligence 4, p. 463.502. 1969.
- [Mehrotra e Fama, 2003] Mehrotra, V.; Fama, J. Call Center Simulation Modeling: Methods, Challenges, and Opportunities. Proceedings of the 2003 Winter Simulation Conference S. Chick, P.J. Sanchez, D.Ferrin, and D.J. Morrice eds. p. 135–142. 2003.
- [Middendorf, 2005] Middendorf, M. Ant Colony Optimization and Particle Swarm Optimization. Porto Alegre: UFRGS, 2005. Palestra proferida no Brazil Agents School 2005, patrocinado pela SBC, em Porto Alegre, em 21/09/2005.
- [Pechoucek e Marik, 2006] Pechoucek, M.; Marik, V. Review of Industrial Deployment of Multi-Agent Systems. Disponível em:

- <<http://agents.felk.cvut.cz/teaching/33ui2/on-applications.pdf>>.
Acesso em: 05 Maio 2006.
- [Perros, 2003] Perros, H. Computer Simulation Techniques: The definitive introduction. Computer Science Department, University Raleigh, NC.p.1-2. 2003.
- [Pichitlamken et al., 2003] Pichitlamken, J.; Deslauriers, A.; L'Ecuyer, P.; Avramidis, A. N.; Modelling and Simulation of a telephone Call center. Proceedings of the 2003 Winter Simulation Conference S. Chick, P.J. Sanchez, D.Ferrin, and D.J. Morrice eds. p. 1805–1812. 2003.
- [RePast, 2005] RePast. Disponível em: <<http://repast.sourceforge.net/>>. Acesso em: 05 Set 2005.
- [Saltzman e Mehrotra, 2004] Saltzman, R. Mehrotra, V. A Manager-Friendly Platform for Simulation Modeling and Analysis of Call Center Queueing Systems.Proceedings of the 2004 Winter Simulation

- [Wallace e Whitt, 2006] Wallace, R.B.; Whitt, W. Resource Pooling and Staffing in Call Centers with Skill Based Routing.2006. Disponível em <<http://www.cs.cmu.edu/~harschol/WORMS04/people/whitt/whitt.pdf>>. Acesso em: 20 fev. 2006.
- [Whitt, 2002] Whitt, W.. Stochastic Models for the Design and Management of Customer Contact Centers: Some Research Directions. New York. Mar., 2002
- [Wooldridge, 1999] Wooldridge, M. Intelligent agents. In Weiß, G., editor, Multiagent Systems—A Modern Approach to Distributed Artificial Intelligence, p. 27–77. MIT Press, Cambridge, MA, 1999.
- [Valladares, 2007] Valladares, M.G.C. Informações Simulação. [mensagem pessoal]. Mensagem recebida por <claudio.lucio@coresynesis.com.br> em 3 jan. 2007.
- [Van Dijk, 1998] Van Dijk, N.M. On hybrid combination of queueing and simulation. Simulation: Past, Present and Future. 12th European Simulation Multiconference 1998. ESM'98. San Diego, CA, USA, p.731–735. 1998.
- [Zohar et al., 2002] Zohar, E. Mandelbaum, A. Shimkim, N. Adaptive Behavior of Impatient Customers in Tele-Queues: Theory and Empirical Support. Management Science. Abr., 2002.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)