Campus de São Carlos

UMA CONTRIBUIÇÃO PARA AVALIAR O
DESEMPENHO DE SISTEMAS DE TRANSPORTE
EMERGENCIAL DE SAÚDE

Renata Algisi Takeda

Orientador: Prof. Tit. João Alexandre Widmer
Co-Orientador: Prof. Dr. Reinaldo Morabito

UNIVERSIDADE DE SÃO PAULO

**ESCOLA DE ENGENHARIA
DE SÃO CARLOS**

UMA CONTRIBUIÇÃO PARA AVALIAR O DESEMPENHO DE SISTEMAS DE TRANSPORTE EMERGENCIAL DE SAÚDE

Renata Algisi Takeda

Tese apresentada à Escola de Engenharia
de São Carlos, da Universidade de São
Paulo, como parte dos requisitos para
obtenção do título de Doutora em
Transportes

ORIENTADOR: Prof. Tit. João Alexandre Widmer
CO-ORIENTADOR: Prof. Dr. Reinaldo Morabito

São Carlos
2000

T136c

Takeda, Renata Algisi

Uma contribuição para avaliar o desempenho de sistemas de transporte emergencial de saúde / Renata Algisi Takeda. -- São Carlos, 2000.

Tese (Doutorado) -- Escola de Engenharia de São Carlos-Universidade de São Paulo, 2000.

Área: Transportes.

Orientador: Prof. Tit. João Alexandre Widmer.

Co-orientador: Prof. Dr. Reinaldo Morabito.

1. Modelo hipercubo de filas. 2. Serviços de emergência. 3. Ambulância. 4. Medidas de desempenho. I. Título.

FOLHA DE APROVAÇÃO

Candidata: Bacharela **RENATA ALGISI TAKEDA**

Tese defendida e aprovada em 01-12-2000
pela Comissão Julgadora:

Prof. Titular **JOÃO ALEXANDRE WIDMER (Orientador)**
(Escola de Engenharia de São Carlos - Universidade de São Paulo)

Profa. Doutora **MIRIAN BUSS GONÇALVES**
(Universidade Federal de Santa Catarina - UFSC)

Prof. Doutor **ROBERTO DIÊGUES GALVÃO**
(Universidade Federal do Rio de Janeiro - UFRJ)

Prof. Doutor **SOLON VENÂNCIO DE CARVALHO**
(Instituto Nacional de Pesquisas Espaciais - INPE)

Prof. Titular **PAULO MORELATO FRANÇA**
(Universidade Estadual de Campinas - UNICAMP)

Prof. Doutor **EDSON MARTINS DE AGUIAR**
Coordenador do Programa de Pós-Graduação em
Engenharia de Transportes

JOSÉ CARLOS A. CINTRA
Presidente da Comissão de Pós-Graduação da EESC

Não se preocupe em entender.
Viver ultrapassa todo entendimento...

(Clarice Lispector)

Aos meus pais, Humberto e Adurse,
e ao Marcelo.

Valeu a pena ?
Tudo vale a pena, quando a alma não é pequena.

(Fernando Pessoa)

AGRADECIMENTOS

A Deus.

Aos meus orientadores, Prof. João Alexandre Widmer e Prof. Reinaldo Morabito, pela orientação, incentivo, dedicação e amizade demonstrados durante estes anos de convivência.

Ao CNPq, pela bolsa de estudos concedida.

A todos os serviços de atendimento médico de urgência que se dispuseram a colaborar com o desenvolvimento deste estudo, em especial, à coordenação do SAMU-192 de Campinas, na pessoa da Dra. Arine Campos Oliveira Assis, pelo apoio e incentivo demonstrados, proporcionando uma maior interação entre a universidade e a realidade de alguns problemas sociais brasileiros.

A todos os meus colegas de pós-graduação, pela amizade e momentos de companheirismo vividos durante estes quatro anos de trabalho.

Aos amigos, em especial, Cynthia P. Lotti, Gisele C. F. Pileggi, María del Pilar A. Rivas, Fernando C. Mendonça e Leonardo P. Maia, pelas importantes contribuições em todo o desenvolvimento desta pesquisa e cultivo de nossa amizade.

A todos os funcionários do Departamento de Transportes, pelo apoio e dedicação durante todo o tempo que permaneci nesta instituição.

Ao Rogério T. Kondo, analista de sistemas do CISC, cuja ajuda na área computacional foi fundamental para a conclusão desta pesquisa.

Finalmente, dois agradecimentos muito especiais...

... ao Marcelo, por toda a colaboração não só destes últimos meses, na elaboração do trabalho final, mas por sua compreensão, carinho e pelo amor demonstrado em todos os momentos e

... aos meus pais, por todo o incentivo e amor dedicados, e principalmente por sempre terem me mostrado que nunca devemos desistir de um grande sonho.

SUMÁRIO

LISTA DE FIGURAS.....	i
LISTA DE TABELAS.....	iii
LISTA DE ABREVIATURAS E SIGLAS.....	v
LISTA DE SÍMBOLOS.....	vi
RESUMO.....	vii
<i>ABSTRACT</i>	viii
1 INTRODUÇÃO	1
1.1 Considerações iniciais	1
1.2 Definição do problema	4
1.3 Objetivos.....	5
1.4 Método adotado e hipótese principal.....	6
1.5 Estrutura do trabalho	7
2 SISTEMAS DE ATENDIMENTO EMERGENCIAL	9
2.1 Classificação dos serviços de atendimento à população	10
2.2 Serviços de emergência	11
2.2.1 <i>Características dos serviços de emergência</i>	12
2.3 Principais problemas pesquisados	14
2.3.1 <i>Número de unidades de atendimento em uma região</i>	14
2.3.2 <i>Programação de equipes</i>	18
2.3.3 <i>Problemas de localização</i>	19
2.4 Sobre o modelo hipercubo de filas	29
3 MODELO HIPERCUBO DE FILAS	32
3.1 Inserindo prioridades	36
3.2 Hipóteses do modelo	37
3.3 Transições entre estados	40
3.4 Equações de equilíbrio	41
3.5 Insuficiência das equações de equilíbrio	45
3.5.1 <i>Adaptações com relação às limitações da fila</i>	47
3.5.2 <i>Relação com o modelo clássico M/M/N</i>	49
3.6 Desempenho do sistema	50
3.6.1 <i>Workload</i>	50
3.6.2 <i>Frequências de despachos</i>	51
3.6.3 <i>Tempos médios de viagem</i>	52

3.7	Calibração dos tempos de atendimento	54
3.8	A solução do modelo	55
3.8.1	<i>Método exato</i>	55
3.8.2	<i>Método aproximado</i>	57
4	A PESQUISA DE CAMPO	58
4.1	Breve histórico sobre o serviço	60
4.2	Caracterização dos sistemas	62
4.2.1	<i>Descrição dos oito cenários observados</i>	63
4.2.2	<i>Principais diferenças encontradas</i>	64
4.2.3	<i>Escolha de um cenário para a coleta de dados</i>	68
4.3	O SAMU-192 de Campinas	69
4.3.1	<i>Coleta de dados</i>	72
5	MODELAGEM DO SAMU-192 E ANÁLISE DOS RESULTADOS	83
5.1	Hipóteses da modelagem	83
5.1.1	<i>Átomos geográficos</i>	84
5.1.2	<i>Chegadas Poissonianas independentes</i>	85
5.1.3	<i>Tempos de viagem</i>	90
5.1.4	<i>Servidores</i>	91
5.1.5	<i>Localização dos servidores</i>	92
5.1.6	<i>Despacho dos servidores</i>	92
5.1.7	<i>Política de despacho dos servidores</i>	93
5.1.8	<i>Tempos de atendimento</i>	95
5.1.9	<i>Relação entre o tempo de atendimento e o tempo de viagem</i>	97
5.2	Possíveis estados do sistema	98
5.3	Equações de balanço	103
5.4	Resultados	105
5.4.1	<i>Workloads</i>	106
5.4.2	<i>Tempos médios de resposta</i>	107
6	AVALIAÇÃO DE CENÁRIOS ALTERNATIVOS	111
6.1	Descentralização das ambulâncias	113
6.1.1	<i>Cenário 1: descentralização de uma ambulância</i>	114
6.1.2	<i>Cenário 2: descentralização de duas ambulâncias</i>	118
6.1.3	<i>Cenário 3: descentralização de quatro ambulâncias</i>	120
6.1.4	<i>Cenário 4: descentralização de cinco ambulâncias</i>	124
6.1.5	<i>Cenário 5: descentralização de seis ambulâncias</i>	127
6.2	Aumentando o número de ambulâncias no sistema	131
6.2.1	<i>Cenário 6: sistema operando com um novo VSB</i>	131

6.2.2	<i>Cenário 7: sistema operando com dois novos VSB.....</i>	134
6.2.3	<i>Cenário 8: sistema operando com dois novos VSB e um novo VSA... </i>	136
6.3	Comparações entre as alternativas operacionais	137
6.3.1	<i>Workload</i>	137
6.3.2	<i>Probabilidades</i>	139
6.3.3	<i>Tempos médios de resposta.....</i>	140
6.3.4	<i>Tempo total de atendimento</i>	142
6.3.5	<i>Atendimentos backup.....</i>	142
7	CONCLUSÕES E PERSPECTIVAS.....	144
7.1	Conclusões.....	145
7.2	Perspectivas para futuras pesquisas	147
	ANEXO A: Caracterização de um sistema de filas.....	149
	ANEXO B: Questionário aplicado nos serviços de emergência	156
	ANEXO C: Definição do período de pico	158
	ANEXO D: Testes de aderência	160
	REFERÊNCIAS BIBLIOGRÁFICAS.....	163
	BIBLIOGRAFIA CONSULTADA.....	171
	APÊNDICE I: Portaria nº 824/GM.....	I-1
	APÊNDICE II: Ficha de regulação médica - SAMU de Campinas.....	II-1
	APÊNDICE III: Exemplo de planilha de coleta de dados.....	III-1
	APÊNDICE IV: Exemplos de equações de balanço - sistema com 10 servidores.....	IV-1

LISTA DE FIGURAS

FIGURA 3.1 - Representação da área de atuação de um sistema	33
FIGURA 3.2 - Espaço de estados para sistemas com três servidores.....	34
FIGURA 3.3 – Espaço de estados para sistemas com 3 servidores e cauda infinita .	35
FIGURA 4.1 – Visão geral dos veículos utilizados pelo SAMU-192 de Campinas..	70
FIGURA 4.2 – Densidade demográfica na área de atuação do SAMU em Campinas.	72
FIGURA 4.3 – Atendimentos emergenciais realizados pelo SAMU.....	73
FIGURA 4.4 – Distribuição do número de atendimentos ao longo do dia	74
FIGURA 4.5 – Atendimentos emergenciais realizados pelo SAMU (nov./1997).....	77
FIGURA 4.6 – Distribuição dos atendimentos ao longo de 24 horas (nov./1997)	77
FIGURA 4.7 – Número total de chamados nas regiões (nov./1997)	78
FIGURA 4.8 – Distribuição percentual do total de chamados (nov./1997)	78
FIGURA 4.9 – Número de chamados no período das 10 às 14 horas.....	81
FIGURA 4.10 – Número total de chamados em cada região no período de pico	81
FIGURA 5.1 - Distribuição espacial do sistema	84
FIGURA 5.2 – Átomos geográficos biparticionados do SAMU-192	85
FIGURA 5.3 – Decomposição de um Processo de Poisson de parâmetro $\lambda > 0$	88
FIGURA 5.4 - Distribuição dos intervalos entre chegadas sucessivas ao sistema. ...	89
FIGURA 5.5 – Diagrama de fluxos para o estado (0000000000).....	99
FIGURA 5.6 – Diagrama de fluxos para o estado (1111111111).....	100
FIGURA 5.7 - Diagrama de fluxos para o estado (1000000101)	101
FIGURA 5.8 – Aleatoriedade da política de despachos: estado (1000000101).....	102
FIGURA 5.9 – <i>Workloads</i> das ambulâncias para o cenário original	106
FIGURA 5.10 – Tempos médios de resposta para cada átomo (cenário original)	109
FIGURA 5.11 – Tempos médios de resposta das ambulância (cenário original)	109
FIGURA 6.1 - Primeiro cenário (1 VSB em O_B).....	114
FIGURA 6.2 – Diagrama de fluxos para o estado (10000000101): cenário 1.....	116
FIGURA 6.3 – Segundo cenário (2 VSB em O_B)	118
FIGURA 6.4 – Terceiro cenário (1 VSB em N_B , S_B , L_B e O_B).....	120
FIGURA 6.5 – Quarto cenário (1 VSB em N_B , S_B e L_B , 2 VSB em O_B).....	124
FIGURA 6.6 – Quinto cenário (1 VSB em S_B e L_B , 2 VSB em N_B e O_B).....	128
FIGURA 6.7 - Sexto cenário: aquisição de um novo VSB.....	132
FIGURA 6.8 - Sétimo cenário: aquisição de dois novos VSB.....	134
FIGURA 6.9 – Oitavo cenário: aquisição de dois novos VSB e um novo VSA	136
FIGURA 6.10 – Comparação entre as <i>workloads</i> médias do sistema	138

FIGURA 6.11 – Impactos do aumento do número de servidores sobre P_Q e P_S	139
FIGURA 6.12 – Impactos do aumento do número de ambulâncias sobre P_0	140
FIGURA 6.13 – Tempos médios de resposta de cada ambulância	141
FIGURA 6.14 – Comparação entre os tempos médios de resposta a cada átomo ...	141
FIGURA 6.15 – Tempos de atendimento para os veículos do sistema.....	142
FIGURA 6.16 – Despachos inter-átomos e chamados atendidos por <i>backup</i>	143
FIGURA A.1 – Diagrama de fluxos para o modelo M/M/1	153
FIGURA A.2 – Diagrama de fluxos para o modelo M/M/c.....	154
FIGURA A.3 – Diagrama de fluxos para o modelo M/M/c/K.....	155

LISTA DE TABELAS

TABELA 4.1 - Cidades participantes da pesquisa e suas populações	62
TABELA 4.2 - Resultados médios relativos aos <i>chamados de emergência</i>	67
TABELA 4.3 - Resultados médios relativos aos <i>atendimentos de emergência</i>	67
TABELA 4.4 – Medidas de desempenho obtidas para o SAMU-192	75
TABELA 4.5 – Dias observados durante o mês de novembro de 1997.....	76
TABELA 4.6 – Medidas de desempenho para o SAMU (nov./1997)	79
TABELA 4.7 – Taxas médias de chegadas para possíveis períodos de pico.....	80
TABELA 4.8 – Medidas de desempenho obtidas para o período de pico	82
TABELA 5.1 – Proporção de chamados em cada átomo do sistema.....	87
TABELA 5.2 – Intervalos médios entre chegadas sucessivas ao sistema	89
TABELA 5.3 – Taxas médias de chegadas para cada átomo	90
TABELA 5.4 – Tempos médios de viagem (minutos).....	91
TABELA 5.5 –Exemplo 1: Matriz de preferência de despachos (cenário original) ..	94
TABELA 5.6 – Exemplo2: Matriz de preferência de despachos (cenário original) ..	94
TABELA 5.7 – Taxas médias de atendimento para cada ambulância	96
TABELA 5.8 – Tempos médios de resposta das ambulâncias (cenário original)	107
TABELA 5.9 –Tempos médios de resposta para cada átomo (cenário original).....	108
TABELA 6.1 – Medidas de desempenho para o primeiro cenário	117
TABELA 6.2 – Medidas de desempenho para o segundo cenário	119
TABELA 6.3 – Indicadores de desempenho para o terceiro cenário	121
TABELA 6.4 – Tempos médios de resposta em cada átomo(cenários 1, 2 e 3).....	122
TABELA 6.5 – Freqüências de despachos para cada ambulância (cenário 3)	123
TABELA 6.6 – Freqüências de atendimentos para cada átomo (cenário 3).....	124
TABELA 6.7 – Indicadores de desempenho para o quarto cenário.....	125
TABELA 6.8 – Tempos médios de resposta para cada átomo (cenário 4)	126
TABELA 6.9 – Freqüências de despachos para cada ambulância (cenário 4)	126
TABELA 6.10 – Freqüências de atendimentos para cada átomo (cenário 4).....	127
TABELA 6.11 – Indicadores de desempenho para o quinto cenário.....	128
TABELA 6.12 – Tempos médios de resposta para cada átomo (cenário 5)	129
TABELA 6.13 – Freqüências de despachos para cada ambulância (cenário 5)	129
TABELA 6.14 – Freqüências de atendimentos para cada átomo (cenário 5).....	130
TABELA 6.15 – Indicadores de desempenho para o sexto cenário	132
TABELA 6.16 – Freqüências de despachos para cada ambulância (cenário 6)	133
TABELA 6.17 – Indicadores de desempenho para o sétimo cenário	135
TABELA 6.18 – Indicadores de desempenho para o sétimo cenário	137

TABELA D.1 – Teste $K-S$ para os intervalos entre chegadas sucessivas	160
TABELA D.2 – Teste $K-S$ para os tempos de atendimento ($\alpha = 5\%$)	161
TABELA D.3 – Quadro de ANOVA, ao nível de significância $\alpha = 5\%$	162

LISTA DE ABREVIATURAS E SIGLAS

SAMU	- Serviço de Atendimento Médico de Urgência
A	- Classe avançada
abr.	- Abril
ago.	- Agosto
ANOVA	- Análise de Variância
B	- Classe básica
BR 101	- Rodovia Federal 101
BR 111	- Rodovia Federal 111
C	- Região Central
CA	- Canadá
C _A	- Região Central geradora de chamados avançados
C _B	- Região Central geradora de chamados básicos
CEP	- Código de endereçamento postal
CRM	- Inscrição no Conselho Regional de Medicina
CS	- Centro de Saúde
D	- Processo determinístico
dez.	- Dezembro
D-P	- Desvio-padrão
E _k	- Distribuição de Erlang de ordem k
EUA	- Estados Unidos da América
F	- Valor da estatística de Fischer
F(x _i)	- Função distribuição de probabilidade estimada
FCFS	- <i>First come, first served</i>
fev.	- Fevereiro
G(x _i)	- Distribuição acumulada de probabilidades observada
gl	- Grau de liberdade
h	- Hora (s)
hh:mm	- Hora:minuto
IBGE	- Instituto Brasileiro de Geografia e Estatística
jan.	- Janeiro
jul.	- Julho
jun.	- Junho
km	- Quilômetro
K-S	- Método de Kolmogorov-Smirnov
L	- Região Leste
L _A	- Região Leste geradora de chamados avançados
L _B	- Região Leste geradora de chamados básicos
LCFS	- <i>Last-come, first-served</i>
LP	- <i>Linear Programming</i>
M/G/∞	- Sistema de filas com processo de chegadas Markoviano, processo de atendimento genérico e “infinitos” servidores
M/G/N/N	- Sistema de filas com processo de chegadas Markoviano, processo de atendimento genérico, N servidores e sala de espera com limitação de N usuários

M/M/c	- Sistema de filas com processo de chegadas Markoviano, processo de atendimento Markoviano e c servidores
M/M/N/N	- Sistema de filas com processo de chegadas Markoviano, processo de atendimento Markoviano, N servidores e sala de espera com limitação de N usuários
mar.	- Março
MB	- Megabites
min	- Minuto(s)
MQ	- Quadrado médio
N	- Região Norte
N	- Não
N _A	- Região Norte geradora de chamados avançados
N _B	- Região Norte geradora de chamados básicos
nov.	- Novembro
O	- Região Oeste
O _A	- Região Oeste geradora de chamados avançados
O _B	- Região Oeste geradora de chamados básicos
out.	- Outubro
PCAM	- <i>Patrol Car Allocation Model</i>
PRI	- Priority
PS	- Pronto-Socorro
PSQ	- Veículo psiquiátrico
S	- Região Sul
S	- Sim
S _A	- Região Sul geradora de chamados avançados
S _B	- Região Sul geradora de chamados básicos
SC	- Estado de Santa Catarina, Brasil
set.	- Setembro
SIRO	- <i>Service in random order</i>
SQ	- Soma de quadrados
St.	- <i>Saint</i>
UTI	- Unidade de tratamento intensivo
VRS	- Veículo de remoção simples
VSA	- Veículo de suporte avançado
VSB	- Veículo de suporte básico
X/Y/c	- Sistema de filas onde X representa o processo de chegadas, Y representa o processo de atendimento e c servidores
$X^\alpha/Y^{\beta,\gamma}/c/m/$	Sistema de filas com c servidores, sala de espera limitada em m usuários, população de N usuários, processo de chegadas em grupos de, no máximo, α usuários e processo de atendimento em grupos de no máximo β e no mínimo γ usuários
N/disciplina	

LISTA DE SÍMBOLOS

$\bar{T}$	- Tempo médio de viagem para o sistema
$\bar{T}_Q$	- Tempo médio de espera de chamados em fila
$\bar{T}_j$	- Tempo médio de viagem para o átomo j
$\bar{T}_{U_n}$	- Tempo médio de resposta do servidor n
#	- Número
%	- Porcentagem
[B]	- Soma das possíveis taxas de transição para o estado $B=[b_1 b_2 \dots b_N]$
$A(t)$	- Número de eventos ocorridos no intervalo de tempo $[0, t]$
B	- Representação vetorial de um estado B qualquer do sistema
E_{nj}	- Conjunto dos estados nos quais um chamado do átomo j é atendido pelo servidor n
f_{nj}	- Frequência de despachos do servidor n ao átomo j
$f_{nj}^{[1]}$	- Frequência de todos os despachos do servidor n ao átomo j, que não implica em tempo de espera para o usuário
$f_{nj}^{[2]}$	- Frequência de todos os despachos do servidor n ao átomo j sujeitos a esperas pelo atendimento
h	- Incremento no processo $A(t)$
$L = [l_{nj}]$	- Matriz de localização dos servidores
l_{nj}	- Elemento correspondente à linha n e coluna j da matriz L
N	- Número de servidores do sistema
N_A	- Número de átomos do sistema
$p[b_1 b_2 \dots b_N]$	- Probabilidade de encontrar o sistema no estado $B=[b_1 b_2 \dots b_N]$
$P_0(\lambda)$	- Processo de Poisson de parâmetro $\lambda > 0$
$p_{00\dots 0}$	- Probabilidade de sistema vazio (com todos os servidores livres)
p_B	- Probabilidade de equilíbrio do estado B
p_{ijk}	- Probabilidade do sistema encontrar-se no estado (ijk)
p_j	- Proporção de chamados no átomo j
p_k	- Probabilidade de estado (k) de um sistema clássico de filas
p_Q	- Probabilidade de formação de fila
p_S	- Probabilidade de saturação do sistema
R^N	- Espaço N-dimensional
S_k	- Estado no qual se encontram k usuários aguardando no sistema
t_{nj}	- Tempo médio de viagem para o servidor n deslocar-se até o átomo j
TR	- Tempo de resposta
x_i	- Ponto representativo da classe i de uma distribuição de frequências (amostra cujos dados estão organizados em classes)
Y	- Processo de contagem
Y_j	- Frequência da classe j
α	- Nível de significância de um teste de hipóteses
λ	- Taxa média de chegadas ao sistema
$\lambda[j]$	- Taxa média de chegadas no átomo j
λ_j	- Taxa média de chegadas no átomo j
μ	- Taxa média de atendimento do sistema

$\mu[n]$	- Taxa média de atendimento do servidor n
ρ	- Carga média de trabalho do sistema
ρ_N	- <i>Workload</i> do servidor N
τ_{ij}	- Tempo médio de viagem do átomo i ao átomo j

RESUMO

TAKEDA, Renata Algisi (2000). *Uma contribuição para avaliar o desempenho de sistemas de transporte emergencial de saúde*. São Carlos. 176 p. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

A rapidez na realização do atendimento às vítimas é uma das maiores necessidades de serviços de atendimento médico de urgência, e o tempo decorrido entre o instante da ocorrência da solicitação pelo serviço e o início do atendimento, denominado tempo de resposta, é um dos principais fatores que influenciam o desempenho do sistema. Este tempo depende de uma reunião de fatores como condições de tráfego, dia e período do dia, número de veículos disponíveis e suas localizações, capacitação profissional da equipe, etc. Apresenta-se neste trabalho uma análise do desempenho do serviço oferecido na cidade de Campinas-SP, tratando o problema por meio do modelo hipercubo de filas, que considera as variações aleatórias dos processos de chegadas e atendimento dos chamados. Sua aplicação produz uma ampla variedade de indicadores de desempenho para o sistema, que são comparados com os valores reais observados, para validar a hipótese de aplicação do modelo. Os resultados de sua aplicação para configurações operacionais alternativas, tais como descentralização e aumento do número de ambulâncias, mostraram uma elevação significativa do nível de serviço oferecido ao usuário. Conclui-se que o modelo constitui uma importante ferramenta de análise para este tipo de sistema, auxiliando na tomada de decisões estratégicas e operacionais do sistema.

Palavras-chave: modelo hipercubo de filas; serviços de emergência; ambulância; medidas de desempenho.

ABSTRACT

TAKEDA, Renata Algisi (2000). *A contribution to evaluate the performance of emergency health transportation systems*. São Carlos. 176 p. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

One of the major concerns of medical emergency systems is to provide the fastest possible medical attention for the victims. The time elapsed between the emergency call and the assistance, called the response time, is one of the main factors that influence the system's performance. This time lapse depends on traffic conditions, the day of the week and time of day, the number of available vehicles and their location, the rescue team's professional qualifications, etc. This work consists of an analysis of the performance of the emergency service available in Campinas, SP, and deals with the problem using the hypercube queuing model, which considers stochastic variations of the arrival and assistance processes. The application of this model produces a wide variety of system performance indicators, which are compared with the real observed values to validate the model's hypothetical application. Application of the model in alternative operational scenarios, such as decentralization and a greater number of ambulances, showed a significant increase in the quality of the service offered to the user. It was concluded that the model constitutes an important analytical tool for this type of system, serving as an aid for strategic and operational decision-making.

Keywords: hypercube queueing model; emergency services; ambulance; measure of effectiveness.

1 INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

A qualidade de vida da população, especialmente das áreas urbanas, está ligada a uma variedade de serviços públicos que podem ser classificados em três grandes grupos: serviços de rotina, semi-emergenciais e serviços de emergência, dentre os quais destaca-se o Serviço de Atendimento Médico de Urgência (SAMU). Quando projetado e operado com eficiência, o SAMU pode salvar vidas; por outro lado, quando ineficiente, é um potencial responsável pelo agravamento dos casos.

A rapidez na realização do atendimento é uma das maiores exigências destes sistemas, e o tempo decorrido entre o instante da ocorrência da solicitação pelo serviço e o início do atendimento junto às vítimas, denominado tempo de resposta, é um dos principais fatores que influenciam o nível de serviço destes sistemas. Este tempo é resultante de um conjunto de fatores, como por exemplo, condições do tráfego local, dia da semana, período

do dia, número de veículos disponíveis e suas localizações, experiência e capacitação profissional da equipe, características operacionais da frota, etc.

Serviços emergenciais como os SAMU's apresentam alto grau de incerteza, e normalmente sua eficiência é medida através do tempo médio de resposta a um chamado, ou seja, o tempo que uma vítima espera em média para começar a receber algum tipo de atendimento. Quanto maior o grau de incerteza envolvido e maior a necessidade de se obter respostas rápidas, menor deve ser a taxa de utilização dos operadores e equipamentos do sistema. Caso contrário, o nível de serviço oferecido pode ser sensivelmente deteriorado. Neste contexto, quando bem dimensionados, geralmente ocorrem longos períodos em que os operadores e equipamentos permanecem desocupados.

Muitas pesquisas vêm sendo desenvolvidas no sentido de se obter métodos para analisar e dimensionar tais sistemas, de forma a elevar o nível de serviço oferecido e também racionalizar os recursos exigidos. No entanto, uma limitação dos estudos é que muitos deles não consideram a natureza probabilística dos processos de chegada e atendimento dos chamados, e não levam em conta o fato de que as ambulâncias nem sempre estão disponíveis para iniciar um atendimento. Dentre as importantes contribuições, destacam-se trabalhos onde a atenção é concentrada em questões tais como localização de bases, roteirização de veículos, zoneamento da área de atuação do sistema e problemas de congestionamento.

A modelagem integrada destas questões é muito complexa, pois geralmente os sistemas reais são compostos por um grande número de veículos, as solicitações por serviço ocorrem temporária e espacialmente, existe cooperação entre veículos de áreas distintas, podem ocorrer múltiplos despachos para atender a um mesmo chamado, o tempo médio de viagem varia de acordo com a região, o dia e o período do dia, e existe a possibilidade de formação de filas de espera.

Alguns pesquisadores dedicaram-se exclusivamente a problemas de localização de bases (TOREGAS et al., 1971, ANDERSON & FONTENOT, 1992 e LOUVEAUX, 1993), enquanto outros consideraram o problema de projeto das áreas de cobertura específicas (KEENEY, 1972 e LARSON, 1974a). DASKIN & STERN (1981) usaram modelos de cobertura de conjuntos com o objetivo de determinar o número necessário de veículos para cobrir cada região individualmente, e também para um conjunto de regiões vizinhas.

Outra maneira de se abordar o problema é por meio de modelos de simulação. SAVAS (1969) usou um modelo de simulação na cidade de Nova Iorque para mostrar que o tempo médio de resposta a um chamado pode ser reduzido redistribuindo as ambulâncias em suas bases. SWOVELAND et al. (1973) aplicaram um modelo de simulação para determinar o tempo médio de resposta das ambulâncias em toda a extensão de Vancouver, para determinar novas localizações e novas configurações das áreas de cobertura do sistema. FITZSIMMONS (1973) desenvolveu um modelo baseado em teoria de filas para analisar a localização de ambulâncias na cidade de Los Angeles, que considera como principal fator o tempo médio de viagem de cada veículo a cada chamado.

No presente trabalho utilizou-se o *modelo hipercubo de filas*, desenvolvido por LARSON (1974a) e estudado por diversos autores (SWERSEY, 1994), para analisar o desempenho atual do serviço oferecido pelo SAMU da cidade de Campinas-SP. O modelo hipercubo, baseado em teoria de filas espacialmente distribuídas, não é um modelo de otimização que determina uma configuração ótima para o sistema, mas fornece uma completa avaliação de desempenho de cada configuração sugerida (HALPERN, 1977).

No Brasil, alguns exemplos importantes de aplicação do modelo são: a localização de ambulâncias em um trecho da BR 111 – SC (GONÇALVES et

al., 1994 e GONÇALVES et al., 1995), e o balanceamento das *workloads* de ambulâncias no sistema Anjos do Asfalto da Rodovia Presidente Dutra (MENDONÇA & MORABITO, 2000). Um exame recente do uso do modelo hipercubo na solução de problemas de localização probabilísticos foi apresentado em CHIYOSHI et al. (2000).

1.2 DEFINIÇÃO DO PROBLEMA

O sistema de transporte emergencial de saúde que atua em grande parte das cidades brasileiras é deficiente em virtude da escassez de recursos disponíveis e da ausência de interesses e investimentos no setor. Somente em meados de 1999, ao se regulamentar as diretrizes para atendimentos realizados por planos de saúde no país, o Ministério da Saúde percebeu a necessidade de normatizar a atividade médica na área de urgência, em fase pré-hospitalar (Portaria nº 824/GM, de 24 de junho de 1999, e que pode ser encontrada, na íntegra, na página da internet do Ministério da Saúde – www.saude.gov.br/portarias/1999.htm).

O atendimento emergencial de saúde é um desafio para todas as nações, pois, independente do grau de urgência envolvido, somente por meio de uma rigorosa organização é possível oferecer um serviço de boa qualidade.

Em geral, estes serviços são de responsabilidade do poder público e, em diversas situações, não se observa a existência de um planejamento adequado à realidade envolvida. O tipo de serviço a ser oferecido deve ser escolhido de acordo com as características da demanda local, considerando as configurações geográficas do município, tipos de atendimentos mais solicitados, regiões de altos índices de demanda, entre outros fatores.

Observa-se, na maioria dos casos, ausência de avaliações periódicas das condições operacionais do sistema, e a legislação hoje existente não define

indicadores de desempenho que revelem, com segurança, o nível de serviço que está sendo oferecido.

Na revisão da literatura realizada foram encontrados vários métodos que auxiliam no dimensionamento e avaliação deste tipo de sistema, cada um enfocando um determinado tipo de problema. No entanto, poucos são os estudos realizados sobre cenários brasileiros, especialmente com a finalidade de oferecer às entidades representativas do setor um instrumento que forneça subsídios à análise e avaliação de políticas operacionais alternativas, através da modelagem do sistema.

Com foco neste problema, o presente trabalho procura contribuir com o gerenciamento dos sistemas urbanos de transporte médico-emergencial brasileiros apresentando uma ferramenta de análise robusta, na tentativa de avaliar não só o nível de serviço oferecido, mas também os impactos de decisões estratégicas e operacionais possíveis de ser aplicadas nos sistemas.

1.3 OBJETIVOS

Os objetivos principais deste trabalho são três:

1. Apresentar um modelo para bem representar a operação de sistemas de transporte por ambulância em condições operacionais típicas de cidades brasileiras, de forma a ser utilizado para indicar alternativas operacionais viáveis para o sistema;
2. Realizar um estudo de caso para ilustrar a aplicabilidade e eficiência do modelo proposto, além de fazer a validação da técnica através de uma análise comparativa entre os dados reais do estudo de caso e os resultados obtidos através da aplicação do modelo; e

3. Desenvolver uma avaliação de desempenho considerando diversas configurações operacionais alternativas para o sistema, estabelecendo uma relação entre as características operacionais do serviço e os níveis de atendimento proporcionados aos usuários.

1.4 MÉTODO ADOTADO E HIPÓTESE PRINCIPAL

O método adotado neste trabalho baseia-se no pensamento de Hempel, denominado “método da hipótese”, onde se cria uma hipótese como tentativa de resposta ao problema em estudo e então a submete à verificação empírica (HEMPEL, 1974).

Neste sentido, a hipótese fundamental a ser verificada nesta pesquisa é a de que o modelo hipercubo de filas constitui uma ferramenta eficaz para a avaliação de serviços urbanos de transporte de emergência no Brasil, podendo ser considerado como instrumento de apoio às decisões estratégicas e operacionais para este tipo de sistema.

A fase de criação corresponde à validação da técnica de modelagem com base em parâmetros como taxas médias de chegadas de chamados emergenciais para cada região e tempos médios de atendimentos para cada servidor obtidos em um sistema nacional, e à aplicação do modelo calibrado para avaliar configurações operacionais alternativas para o sistema.

Para a realização desta fase, é necessário, inicialmente, conhecer alguns sistemas que atuam em cidades brasileiras, diferenciando-se cenários de cidades pequenas, médias e grandes e identificar os principais fatores que determinam o nível de serviço oferecido.

A aplicação se dá a partir da verificação das hipóteses da modelagem e em seguida, a verificação empírica através da comparação entre os valores dos

indicadores de desempenho gerados pelo modelo e os observados no sistema real.

1.5 ESTRUTURA DO TRABALHO

Para se concluir a pesquisa, foi necessário cumprir uma seqüência de etapas, mostradas em detalhes ao longo do trabalho, de acordo com a seguinte estrutura:

Capítulo 2 – Sistemas de Atendimento Emergencial

Neste capítulo, apresentam-se as características gerais dos serviços de atendimento de emergência. Procura-se destacar a importância de um bom gerenciamento, visto que falhas no atendimento podem trazer conseqüências catastróficas. Ao mostrar os tipos de serviços emergenciais e os seus principais problemas, também é incorporada uma revisão da literatura existente.

Capítulo 3 – Modelo Hipercubo de Filas

Este capítulo trata da formulação do modelo hipercubo proposto para representar, analiticamente, sistemas de atendimento de emergência de cidades brasileiras. São apresentadas as hipóteses que validam a aplicação do modelo, construção do espaço de estados e as soluções de equilíbrio para o sistema. Finalmente, são apresentados os principais indicadores de desempenho produzidos pelo modelo utilizando os resultados das probabilidades de equilíbrio geradas.

Capítulo 4 – A Pesquisa de Campo

Nesta etapa, apresenta-se um relatório da pesquisa de campo realizada e calcula-se as estatísticas descritivas sobre a base de dados obtida. O cenário escolhido para estudo foi a cidade de Campinas-SP, onde foi possível encontrar apoio dos profissionais responsáveis pelo serviço, e interesse em

participar da pesquisa, integrando o meio acadêmico ao meio onde está inserido o problema, além da disponibilidade de informações para se criar uma base de dados confiável.

Capítulo 5 – Modelagem do SAMU-192 e Análise dos Resultados

Dando continuidade ao Capítulo 4, é realizado o estudo de caso para validar o modelo hipercubo como ferramenta para representar serviços de transporte por ambulância em cidades brasileiras, considerando informações fornecidas por uma cidade de grande porte.

Validadas as hipóteses e construído o sistema de equações de balanço, o modelo é resolvido computacionalmente produzindo indicadores de desempenho para a configuração original do sistema. Os valores obtidos são comparados com os observados e assim determinados os possíveis desvios produzidos com a aplicação da modelagem.

Capítulo 6 – Avaliação de Cenários Alternativos

A parte inicial deste capítulo consiste da construção de cenários alternativos baseados na experiência adquirida com profissionais do setor e problemas já estudados (apresentados na bibliografia), com o objetivo de se alcançar níveis de serviço superiores para o sistema. Em seguida, é apresentada uma análise de sensibilidade entre os resultados obtidos para os diversos cenários, no sentido de se avaliar as vantagens e desvantagens destes cenários.

Capítulo 7 – Conclusões e Perspectivas

Neste capítulo são apresentadas as principais conclusões obtidas ao longo do desenvolvimento do trabalho. Para finalizar, são descritas algumas perspectivas para futuras pesquisas nesta área.

2 SISTEMAS DE ATENDIMENTO EMERGENCIAL

Este capítulo trata das características dos serviços de atendimento à população, em especial os de responsabilidade do setor público, como é o caso dos serviços de emergência. São apresentados os vários problemas enfrentados por planejadores e operadores destes sistemas e ferramentas existentes na literatura para solucioná-los, com o objetivo de conhecer uma técnica eficaz para a representação e avaliação do serviço de transporte emergencial de saúde que opera na maioria das cidades brasileiras.

A exigência da implantação de instalações que comportem serviços que atendam às necessidades básicas do cidadão tem sido crescente nos últimos anos, devido, sobretudo, ao desenvolvimento complexo das estruturas sociais e econômicas, e do crescimento desordenado das cidades a partir do início da década de 70. Serviços como habitação, energia elétrica, rede de água, saneamento e sistemas de comunicações, bem como limpeza, segurança e saúde, são atividades presentes no dia a dia da população.

Em geral, estes serviços são administrados pelo poder público local e, em grande parte dos casos, não se observa a existência de um planejamento

adequado à realidade envolvida. As decisões partem de alguns grupos ou de alguns setores, e as soluções devem atender a objetivos de uma ampla comunidade; estas decisões devem ser estratégicas, dado que envolvem grande quantidade de recursos, e geralmente efeitos de médio e longo prazos. (ALMEIDA & GONÇALVES, 1996)

Planejar um sistema consiste em um processo de previsão de necessidades e racionalização dos recursos materiais e humanos disponíveis, a fim de se alcançar determinadas metas em etapas e prazos pré-definidos. Ou seja, planejar é uma atividade com dimensões técnica, administrativa e política, principalmente em se tratando de bens de serviços públicos. Neste caso, os objetivos do planejamento estão intimamente relacionados com a satisfação das necessidades da população, sejam elas individuais ou coletivas (BELTRAMI, 1977; SWERSEY & INGALL, 1986).

2.1 CLASSIFICAÇÃO DOS SERVIÇOS DE ATENDIMENTO À POPULAÇÃO

Segundo GONÇALVES (1994), os serviços urbanos de atendimento à população podem ser classificados em três grandes grupos: os serviços de rotina, os semi-emergenciais e os de emergência.

Os serviços de rotina ou periódicos são regidos por padrões estatísticos bem definidos, sem grandes variações, uma vez que são serviços previsíveis, podendo assim oferecer um bom nível de serviço, com uma alta taxa de ocupação de seus operadores e equipamentos. Exemplos destes serviços são: coleta domiciliar de lixo, entrega de jornais e correspondências, etc.

Os serviços semi-emergenciais apresentam um alto grau de incerteza, não podendo se prever o instante da ocorrência de uma solicitação pelo serviço. Nestes casos, em particular, longas esperas não implicam em graves

consequências para os usuários, não sendo comprometido também o desempenho do sistema. Dentre os principais serviços semi-emergenciais encontram-se os reparos em redes de energia elétrica, água, telefone, etc.

Nos serviços de emergência, o grau de incerteza envolvido é muito alto, e normalmente a sua eficiência é medida através do tempo de resposta a um chamado de urgência, ou seja, o tempo de espera de uma vítima para começar a receber algum tipo de atendimento. Nestes casos, atrasos no início do atendimento podem ter consequências catastróficas. Quanto maior o grau de incerteza envolvido e maior a necessidade de se obter respostas rápidas, menor deve ser a taxa de utilização dos operadores e equipamentos do sistema. Caso contrário, o nível do serviço oferecido pelo sistema é deteriorado. Desta forma, sistemas emergenciais bem dimensionados apresentam operadores e equipamentos desocupados por um longo período de tempo.

Uma análise determinística de tais fenômenos pode levar a conclusões desacertadas, uma vez que nem sempre são incorporadas as variabilidades dos processos. Portanto, para uma boa avaliação de desempenho de sistemas emergenciais, é indispensável o uso de ferramentas probabilísticas, obtendo-se assim, indicadores que melhor refletem a realidade observada.

2.2 SERVIÇOS DE EMERGÊNCIA

A rapidez na realização do atendimento é uma das maiores necessidades destes serviços, e o tempo decorrido entre o instante da ocorrência da solicitação pelo serviço e o início do atendimento junto às vítimas é um dos principais fatores que influenciam o desempenho do sistema.

Atividades como polícia, bombeiros e transporte por ambulância vêm sendo foco de vários grupos de pesquisadores, desde o início dos anos 70 (CHAIKEN & LARSON, 1972). Países como Estados Unidos, Canadá e França foram pioneiros nos estudos, produzindo melhorias significativas no

nível de serviço oferecido à população da cidade de Nova Iorque, em serviços como patrulhamento policial (LARSON, 1972) e estratégias de decisão em situações de emergência do corpo de bombeiros (INGALL et al., 1975).

Pesquisas sobre atendimento médico de urgência iniciaram-se também na mesma época. SAVAS (1969), por exemplo, utilizou simulação computacional para analisar as conseqüências da variação do número de ambulâncias em operação sobre o nível de serviço oferecido na cidade de Nova Iorque.

2.2.1 Características dos serviços de emergência

2.2.1.1 Ambulâncias

O transporte emergencial de saúde é de importância vital, porém, ainda pouco valorizado pela sociedade. Pequenas melhorias no tempo de resposta a um chamado podem provocar efeitos significativos nos índices de salvamento de vítimas.

No final da década de 70 surge um novo conceito de atendimento médico de urgência. A ambulância deixa de ser um simples veículo de remoção de vítimas, e passa a ser complementada por um suporte avançado de vida composto por equipamentos, medicamentos e profissionais especializados. Esta nova estrutura permite que o tratamento se inicie no local da ocorrência, estendendo-se durante todo o transporte.

Neste contexto, MAYER (1979) examinou 525 casos de paradas cardíacas ocorridas no período de um ano, na cidade de Seattle. Foram considerados apenas os casos onde as vítimas apresentaram chances de sobrevivência, correspondendo a 60% das ocorrências, tendo fibrilação ventricular como causa. As análises estabeleceram uma relação estatística

significativa entre as taxas de sobrevivência das vítimas e o tempo gasto por uma equipe para chegar ao local da ocorrência e iniciar o tratamento.

Desde então, outras pesquisas foram desenvolvidas sobre o assunto. Os exames em KOLESAR & SWERSEY (1986) e SWERSEY (1994) apresentam análises detalhadas sobre métodos que têm sido aplicados para solucionar problemas nesta área.

2.2.1.2 Bombeiros

Atualmente, o corpo de bombeiros de uma cidade realiza tipos distintos de atendimento: o socorro em situações de incêndio e destruição, que é o princípio básico de suas atividades, e situações de salvamento de vidas (cooperando com os serviços prestados pelas secretarias municipais de saúde). Estas duas funções são desempenhadas por equipes especializadas e veículos diferentes. No primeiro caso, são caminhões cujos equipamentos são de combate ao fogo, vazamentos de gases tóxicos e salvamento de vítimas; no segundo caso, o atendimento é realizado pela equipe de resgate, em um veículo semelhante a uma ambulância, cujo objetivo é dar suporte básico de vida, como por exemplo, em casos de acidentes e emergências que exijam a remoção das vítimas a um centro médico.

Os principais problemas encontrados neste sistema são: determinar o número de unidades necessárias e suas localizações, dimensionamento da equipe de profissionais, determinar quantas equipes alocar para realizar um atendimento, dentre outros. SWERSEY (1982) estudou, através de modelos Markovianos de filas, situações para as quais é mais coerente enviar uma equipe mínima para atender o chamado, considerando a possibilidade de, em seguida, ocorrer um chamado ainda mais grave.

2.2.1.3 Polícia

A operação de patrulhamento é realizada em viaturas que são alocadas para determinados setores de uma cidade, podendo operar com um ou mais policiais, dependendo do índice de criminalidade associado ao seu setor de cobertura. Os chamados ocorrem em uma central telefônica e são encaminhados, via rádio, para o veículo responsável pelo setor. Caso o veículo não esteja disponível e o chamado seja de alta prioridade, ele é redirecionado para ser atendido por uma equipe vizinha que esteja disponível no momento; caso contrário, passa a esperar em uma fila espacialmente distribuída.

Os setores de uma cidade são projetados de forma a balancear a carga de trabalho das viaturas, e contribuem para produzir resultados expressivos em se tratando do tempo de resposta a um chamado (LARSON, 1974c).

Estudos avançados vêm sendo desenvolvidos ao longo do tempo, enfocando problemas operacionais como número de unidades necessárias para atender uma região, localização das unidades de atendimento, políticas de despacho dos veículos, programação das equipes de atendimento ao longo do dia, avaliação do desempenho do sistema, dentre outros. Várias são as ferramentas utilizadas para tratar tais problemas, e a precisão dos resultados é dada em função do nível de detalhamento desejado para a análise.

2.3 PRINCIPAIS PROBLEMAS PESQUISADOS

2.3.1 Número de unidades de atendimento em uma região

O principal objetivo deste tipo de pesquisa é encontrar o número ideal de unidades de atendimento destinadas a uma determinada região, sem se preocupar com a localização de tais unidades.

2.3.1.1 Método da raiz quadrada

Uma maneira simples de determinar o número de unidades de atendimento em uma região é através do método da raiz quadrada. Ele consiste em estimar o tempo médio de viagem como função do número de unidades em operação (LARSON & ODoni, 1981).

LARSON (1972) e, posteriormente, KOLESAR & BLUM (1973) mostraram que a distância média percorrida em uma viagem dentro de uma região é inversamente proporcional à raiz quadrada do número de unidades da região. A constante de proporcionalidade depende da configuração do tráfego local e localização da unidade, e pode ser determinada por ferramentas geográficas ou simulação.

Uma das grandes aplicações deste método deu-se nas décadas de 70 e 80 na cidade de Nova Iorque, onde foram autorizadas as operações de 6 batalhões do corpo de bombeiros e relocados 7 outros (INGALL et al., 1975). Posteriormente, TIEN & CHIU (1985) utilizaram o método para avaliar a diminuição nos custos quando se mantém fechados alguns batalhões das 2 às 10 horas, devido ao baixo índice de ocorrências, reduzindo também o custo com recursos humanos.

2.3.1.2 Modelos de filas

Por se tratar de sistemas emergenciais, é fundamental considerar a natureza estocástica dos processos; neste sentido, vários modelos analíticos de filas podem ser aplicados, produzindo bons resultados do ponto de vista prático.

BELL & ALLEN (1969) representaram um sistema emergencial de saúde através do modelo $M/G/\infty$, e determinaram o número de ambulâncias necessárias para atender a demanda, considerando que a probabilidade de se encontrar todas as ambulâncias ocupadas é menor ou igual à soma das

probabilidades de se encontrar cada veículo particular ocupado. Para a aplicação do modelo $M/G/\infty$, não é necessário especificar a distribuição dos tempos de atendimento, bastando apenas estimar a taxa média de atendimento. Por não ser um modelo de filas espacialmente distribuídas, os autores consideraram um sistema com todas as ambulâncias posicionadas em uma base central, juntamente com a central telefônica de chamados.

O mesmo modelo foi utilizado por CHAIKEN¹ apud SWEREY (1994) para encontrar o número de equipes de bombeiros necessárias para atender a região de cobertura de um batalhão. Porém, o problema diferencia-se do caso anterior devido ao fato de que é comum a ocorrência de casos em que é necessário despachar mais de uma equipe para realizar o atendimento.

TAYLOR & TEMPLETON (1980) descreveram uma aplicação na região central da cidade de Toronto, cujos usuários foram separados em duas classes distintas: aqueles que necessitam de atendimento imediato (casos cardíacos, por exemplo), denominados *emergências*, e os que podem aguardar para receber o serviço (transferências entre hospitais), denominados *urgências*. Para efeitos de análise, foi considerado um único processo de chegadas Poissoniano, com proporções distintas de chamados relativos a cada classe, e tempos de atendimento exponencialmente distribuídos. O modelo produziu três importantes indicadores de desempenho para o sistema: a probabilidade de uma emergência encontrar um determinado número de ambulâncias disponíveis, o tempo médio de espera de uma urgência e a probabilidade do tempo de espera de uma urgência ser superior a um valor considerado

¹ CHAIKEN, J. (1971). *The number of emergency units busy at alarms which require multiple servers*. Santa Monica, CA, The Rand Corporation (R-531-NYC/HUD) apud SWERSEY, A. J. (1994). The deployment of police, fire, and emergency medical units. *Handbooks in OR & MS*, S. M. Pollock, et al., v. 6, p. 151 – 200.

aceitável. Com estas medidas, foi possível determinar um número ótimo de veículos para realizar as operações, elevando o nível de serviço oferecido.

Nos serviços de patrulhamento, CHAIKEN & DORMONT (1978a) apresentaram um modelo de alocação de veículos de patrulha (*Patrol Car Allocation Model – PCAM*) que é uma extensão do modelo desenvolvido por LARSON (1972). Sua base é o modelo M/M/c com classes de prioridade. Para cada delegacia são especificados o setor de responsabilidade do patrulhamento, a extensão das ruas, velocidade de resposta das viaturas e índices criminais, além das taxas médias de chamados por hora, ao longo do dia, e durante a semana (cada hora do dia é tratada individualmente, segundo o modelo M/M/c, conforme KOLESAR et al., 1975).

CHAIKEN & DORMONT (1978b) descreveram um algoritmo para encontrar a solução computacional do modelo PCAM, que calcula medidas de desempenho como número ótimo de viaturas necessárias para a operação, frequência da patrulha, tempo médio de viagem até o local da ocorrência e tempo médio de espera para cada classe de prioridade. Resultados com base em informações do serviço em operação na cidade de St. Louis, EUA, apontaram que os chamados emergenciais da comunidade consomem uma pequena fração do tempo de operação das viaturas. A maior parte do tempo é gasta em atividades como transporte de prisioneiros, manutenção, alimentação da equipe, etc.

Novas versões mais aprimoradas para o PCAM foram surgindo, conforme a necessidade de aplicação nos mais variados cenários. GREEN (1984) e GREEN & KOLESAR (1984) incorporaram ao modelo a análise de situações em que é necessário enviar mais de uma viatura para atender uma ocorrência. O novo modelo pressupõe chamados chegando segundo um processo de Poisson, e várias classes de prioridades, cada uma apresentando uma particular distribuição do número de viaturas despachadas para atender

uma ocorrência. Os tempos de atendimento são aleatórios, independentes e identicamente distribuídos, e são calculados a partir do momento em que todas as viaturas destinadas ao atendimento estão disponíveis.

2.3.2 Programação de equipes

Um dos grandes problemas enfrentados em países em desenvolvimento é a organização dos serviços de atendimento à comunidade. Por um lado, há um grande número de pessoas que necessitam receber atendimento e por outro, a eficiência dos serviços, em geral, apresenta problemas. A escassez de recursos para a aquisição de equipamentos e materiais e contratação de pessoal especializado é uma constante.

Pesquisas neste sentido foram desenvolvidas nas últimas décadas para sistemas de emergência de vários países desenvolvidos, dentre eles, os Estados Unidos, as quais respeitando-se as características de cada sistema em particular, podem vir a ser aplicadas em países em desenvolvimento.

KAPLAN (1979) aplicou um modelo simples para comparar o emprego de um ou dois policiais em viaturas de patrulhamento policial em San Diego. Foram avaliados os efeitos sobre a cobertura, o tempo de resposta da viatura, a frequência da patrulha, detecção de crimes, etc. Um dos importantes resultados foi com relação à diminuição significativa (em torno de 50%) do tempo de resposta a um chamado emergencial quando se trabalha com duas viaturas patrulhando uma determinada área, cada uma com apenas um policial, ao invés de apenas uma viatura e dois homens.

Problema semelhante foi estudado por GREEN & KOLESAR (1984) para analisar a eficiência do sistema de patrulhamento na cidade de Nova Iorque, quando na viatura estão dois ou apenas um policial. A questão levantada pelos pesquisadores foi com relação ao número necessário de viaturas com apenas um policial e com dois policiais para atender os

chamados de alta prioridade, admitindo mesmo tempo de espera para todos estes chamados.

TAYLOR & HUXLEY (1989) desenvolveram uma aproximação que foi aplicada para programar os turnos de trabalho de policiais em São Francisco. Foram avaliados os efeitos da mudança de 5 dias e 8 horas/dia para um quadro de 4 dias, 10 horas/dia. Inicia-se com a previsão do número de policiais necessários por hora, e determina-se o horário de início de turno e número de policiais em cada turno, por meio de uma função multi-objetivo não linear que determina o número mínimo de policiais necessários em todos os períodos da semana.

TRUDEAU et al. (1989) e AUBIN (1992) apresentaram métodos para programar equipes médicas de emergência utilizados com sucesso em Montreal, CA. O problema é formulado como um problema de programação inteira cuja função objetivo é minimizar os custos com pessoal devendo satisfazer as necessidades horárias do sistema.

Muitas pesquisas em programação de equipe também foram realizadas em outras áreas, como por exemplo, tripulação de aviões, equipe de enfermeiros, etc.

2.3.3 Problemas de localização

O objetivo dos problemas de localização é determinar as localizações das unidades de atendimento. Os modelos de localização são classificados em duas grandes categorias: os modelos determinísticos e os modelos probabilísticos.

O primeiro caso considera que, sempre que solicitado, o sistema possui uma unidade disponível para atender; estes modelos são freqüentemente aplicados para localizar batalhões do corpo de bombeiros. Os modelos

probabilísticos são mais indicados para localizar bases para ambulâncias, projetar setores para patrulhamento policial e decidir quantos e quais unidades enviar para realizar um atendimento.

2.3.3.1 Modelos determinísticos

As pesquisas envolvendo modelos determinísticos são, em sua grande maioria, direcionadas para o serviço do corpo de bombeiros, pois trata-se de um sistema com uma alta taxa de disponibilidade das viaturas – em torno de 95% (SWERSEY, 1994).

Para determinar a localização das unidades, cada estudo segue um caminho particular, como por exemplo, através dos problemas de tempos médios de viagem que reflete os tempos médios de resposta (solucionados, muitas vezes pelo método de *p-medianas* ou pelo método de *p-centros*), ou problemas de *cobertura de conjuntos*, dentre outros.

a) Problemas de tempos médios de viagem

Segundo NOVAES (1989), o problema matemático clássico de localização é devido a Steiner que, em 1837, analisou a questão da localização do ponto central, ou seja, para um sistema simples, onde são dados três pontos, determina-se um ponto, denominado “ponto central”, tal que a soma das distâncias até os três pontos dados seja mínima. Posteriormente, generalizou-se este estudo para problemas com um número finito de pontos. Esta teoria é conhecida como o problema de localização de *p-centros* ou “solução *minimax*” (CHRISTOFIDES, 1975).

Este problema deu origem a uma importante linha de pesquisa freqüentemente usada em análise de localização e distribuição de bens e serviços: a teoria do lugar central. Alguns estudos importantes foram desenvolvidos neste contexto. Dentre eles, destacam-se os trabalhos de

REVELLE & SWAIN (1970), WHITHE & CASE (1974), HODGART (1980) e BEAUMONT (1987).

A extensão do problema de *p-centros* deu origem a um novo problema, cujo objetivo é localizar um número fixo, p , de unidades de atendimento a fim de minimizar a distância média destas unidades até os pontos geradores de demanda por serviço. Este problema é conhecido como o problema de *p-medianas*, e é um dos métodos mais utilizados para resolver problemas de localização, em especial, para os casos de localização de serviços emergenciais. Para estes sistemas, em geral, a distância média é interpretada como sendo o tempo médio de resposta a um chamado.

HOGG (1968) desenvolveu um modelo para localizar batalhões do corpo de bombeiros em várias cidades da Inglaterra. No modelo:

- as cidades foram divididas em m pequenas sub-áreas;
- a taxa média de ocorrências em cada área foi determinada a partir do histórico dos sistemas;
- n localizações são especificadas inicialmente, e
- os tempos de viagem inter-áreas foram determinados através dos tempos de resposta a uma ocorrência ou utilizando a velocidade do veículo ao longo da malha rodoviária.

O principal objetivo é encontrar, para um certo número de batalhões, as localizações que minimizam o tempo total de viagem de todas as unidades, para todas as ocorrências.

REVELLE & SWAIN (1970) formularam um modelo baseado no problema de *p-medianas*, similar ao desenvolvido por HOGG (1968), cujo objetivo foi minimizar a distância média ou o tempo médio de viagem de um usuário para um conjunto de unidades de serviço. O modelo estimou o tempo

médio de viagem de uma equipe de bombeiros até o local da ocorrência, podendo também ser aplicado para casos como localização de hospitais, escolas, creches, bibliotecas, postos de correios, etc.

TOREGAS et al. (1971) reconheceram que o problema de *p-medianas* tem uma séria deficiência, pois minimizar o tempo de viagem do sistema pode implicar em um desbalanceamento dos tempos de viagem para cada nó de demanda, em particular.

Posteriormente, outros métodos de solução para o problema, com algoritmos mais eficientes, foram propostos. JARVINEN et al. (1972), por exemplo, aplicaram o algoritmo *branch and bound* em sistemas de tamanho moderado. GALVÃO (1980) utilizou o algoritmo *branch and bound* para derivar limitantes baseado no dual-LP das formulações de programação inteira. Ainda nesta direção, GALVÃO (1993) e GALVÃO & RAGGI (1989) apresentaram estudos sobre o uso de *relaxação Lagrangeana* na solução de problemas de programação inteira para otimizar localizações de instalações. Outros trabalhos podem ser encontrados em SWERSEY (1994).

No Brasil, SOUZA et al. (1998), por exemplo, desenvolveram um modelo para distribuir ambulâncias para o socorro em rodovias. O modelo é baseado no problema de *p-medianas*, com restrições relativas ao número de feridos e óbitos. A aplicação prática foi realizada em um trecho da BR 101, no Estado de Santa Catarina, constatando-se que é possível obter uma redução significativa do tempo de resposta e, sobretudo, uma redução do índice de vítimas fatais. Isto garante um nível de serviço mais elevado aos usuários.

LIMA & GONÇALVES (1999) aplicaram o problema de *p-medianas* para determinar a localização espacial de centros intermediários de serviços especializados de saúde no Estado de Santa Catarina, Brasil. A aplicação foi desenvolvida com o objetivo de planejar a distribuição geográfica de centros cardiológicos intermediários no estado, elevando assim a qualidade do

atendimento à população, que necessita de serviços especializados de cardiologia, mas está distante dos grandes centros presentes na capital.

b) Problemas de cobertura de conjuntos

Os problemas de cobertura de conjuntos tiveram como motivação inicial os sistemas de atendimento de emergência. Na localização de serviços de emergência busca-se, em geral, prover cobertura a áreas de demanda. A noção de cobertura implica na definição de uma distância (tempo de resposta) de serviço, que é a distância (tempo de resposta) crítica além da qual a área de demanda é considerada não coberta. Uma área de demanda é considerada coberta se está a menos da distância (tempo de resposta) crítica de pelo menos uma unidade de serviço existente, independente da unidade estar disponível quando o serviço é solicitado (CHIYOSHI et al., 2000).

TOREGAS et al (1971) estudaram o problema de localização de batalhões do corpo de bombeiros considerando que existem n nós geradores de demanda e que as instalações podem se dar em qualquer nó, através de problemas de cobertura de conjuntos solucionado por meio de programação inteira.

O modelo apresentado por TOREGAS et al. (1971) foi posteriormente utilizado por WALKER (1974) para posicionar equipamentos especializados do corpo de bombeiros em Nova Iorque, considerando restrições como índice de solicitações de serviços que necessitam destes tipos de equipamentos em cada área da cidade. SCHREUDER (1981) desenvolveu estudo semelhante para Rotterdam.

CHURCH & REVELLE (1974) reconheceram que os problemas de cobertura de conjuntos pode conduzir a um número excessivo de unidades de serviço ao se considerar que todos os nós de demanda devem ser cobertos.

Então, estenderam o problema considerando localizações com máxima cobertura esperada.

EATON et al. (1985) aplicaram o modelo de localização de máxima cobertura esperada para posicionar ambulâncias em Austin, Texas. O seu objetivo foi determinar o número de ambulâncias necessário para operar na cidade de tal forma que 95% dos chamados fossem atendidos dentro de 5 minutos.

Várias extensões foram desenvolvidas para esta modelagem. SWERSEY (1994) relatou grande parte destes trabalhos. Uma das grandes contribuições comentada por Swersey está apresentada em BATTA & MANNUR (1990) que ajustam o modelo para situações onde mais de uma unidade de serviço é necessária para atender a solicitação. Por exemplo, um sério incêndio onde é necessário enviar dois ou três caminhões do corpo de bombeiros, além do veículo resgate para socorrer as vítimas. Os autores especificaram a cobertura padrão para cada veículo despachado para o local da emergência, embutindo a restrição da máxima distância de percurso permitida para cada veículo.

Outros tipos de aplicações do modelo também foram realizadas com sucesso ao longo do tempo: localização de barreiras utilizadas em situações de vazamento de óleo (BELARDO et al. , 1984) e localização das bases da companhia de energia elétrica em Denver (PLANE & HENDRICK, 1977), dentre outras.

2.3.3.2 Modelos probabilísticos

Problemas de localização probabilísticos preocupam-se, de maneira geral, com a natureza estocástica dos sistemas reais. Nestes sistemas, um dado número de parâmetros tais como tempos de viagem, custo, localização dos

usuários e das unidades, índice de demanda e disponibilidade de servidores são variáveis aleatórias (CHIYOSHI et al., 2000).

O objetivo deste tipo de modelagem é localizar unidades de serviço, de forma a otimizar a medida de desempenho escolhida, para um conjunto de parâmetros considerados. Todos os modelos consideram as distribuições de probabilidade das variáveis aleatórias em estudo, que podem ser incorporadas em formulações padrão de programação matemática ou sob um enfoque da teoria das filas. Segundo OWEN & DASKIN (1998), alguns autores incorporam tais distribuições nas formulações de programação matemática, outros as utilizam em um enfoque de teoria das filas.

a) Tempos médios de viagem

SAVAS (1969) foi o precursor dos estudos de localização considerando as características probabilísticas do processo, para o caso das ambulâncias de um único distrito do Brooklyn, na cidade de Nova Iorque. Ele utilizou simulação como ferramenta, e avaliou os benefícios de trazer algumas ambulâncias com bases em regiões satélite para áreas com maior incidência de demanda. Os resultados simulados apontaram para uma redução em torno de 10% no tempo médio de resposta a um chamado emergencial.

VOLZ (1971) estudou o problema de localização de ambulâncias nas rodovias de acesso às entradas de Michigan, nos Estados Unidos, com o objetivo de minimizar o tempo de resposta a um chamado, dado um valor máximo aceitável. Para isto, estimou a função distribuição de probabilidade do número de ambulâncias ocupadas como sendo uma binomial, e utilizou o modelo de filas $M/G/\infty$ para descrever o desempenho do sistema. Para atender as hipóteses do modelo, foi considerado que todos os chamados convergiam para uma única central de chamados, e que também exerceu a função de base central para localizar as ambulâncias. A partir de então, os veículos passaram a

ser reposicionados em uma rede de transporte, onde cada nó representava os vários municípios de Michigan, e os arcos da rede, as respectivas rodovias de interligação. Os tempos de viagem entre os pontos foi considerado a partir do tipo de rodovia e da velocidade média de percurso desempenhada em cada rodovia.

Método semelhante foi utilizado por FITZSIMMONS (1973), que estimou a distribuição de probabilidade do número de ambulâncias ocupadas através da aplicação do modelo $M/G/\infty$. Em seguida, ele utilizou simulação para estimar a probabilidade de uma particular ambulância estar ocupada, dado um número de veículos ocupados. Este método foi aplicado com sucesso nos sistemas de Los Angeles, EUA e Melbourne, Austrália.

Em Vancouver, Canadá, SWOVELAND et al. (1973) estudaram o problema de localizar ambulâncias de tal forma a minimizar o tempo médio de resposta do sistema. Para isto, a cidade foi dividida em regiões com *workloads* aproximadamente balanceadas, fixando uma ambulância em cada região. O trabalho consistiu em determinar, para cada região, a probabilidade de, no instante da ocorrência do chamado, uma certa ambulância estar disponível para executar o serviço.

MIRCHANDANI & ODoni (1979) estenderam o problema de *p-medianas* para uma rede estocástica, na qual os tempos de viagem nos arcos e/ou a demanda em cada nó são variáveis aleatórias discretas. Os autores mostraram que existe uma solução ótima, e que o problema pode ser formulado como um problema de programação inteira. WEAVER & CHURCH (1983) desenvolveram uma aproximação computacional para resolvê-lo. BERMAN et al. (1987) descreveram um algoritmo com o objetivo de localizar uma unidade de serviço em uma rede a fim de minimizar o tempo médio de resposta a um chamado qualquer.

Muitos outros trabalhos foram desenvolvidos. Dentre eles destacam-se: JARVIS (1975), BERMAN & LARSON (1982), WEAVER & CHURCH (1985).

b) Problemas de cobertura de conjuntos

Os problemas de cobertura probabilísticos levam em conta a aleatoriedade na disponibilidade dos servidores, que é o fator mais importante em se tratando de atendimento de emergência.

DASKIN (1983) apresentou um modelo de localização com máxima cobertura. O autor desenvolveu uma heurística de substituição de vértices para resolver o problema. O algoritmo inicia-se sob a hipótese de que todos os servidores estão ocupados, praticamente durante todo o tempo. Nestas condições, todos os servidores devem ser posicionados no nó que cobre a maior parte da demanda. O modelo calcula as taxas de ocupação dos servidores cujas localizações devem mudar, e assim encontra a melhor configuração de localização dos servidores para o sistema. O algoritmo foi testado para uma rede de 55 nós, definida por SWAIN² apud CHIYOSHI et al. (2000), e a solução encontrada é ótima para o problema de localização com máxima cobertura. Este fato pode ser constatado utilizando-se a *heurística Lagrangeana* proposta por GALVÃO & REVELLE (1996) para este tipo de problema. O modelo foi posteriormente utilizado para localizar ambulâncias em Bangkok, Tailândia (FUJIWARA et al., 1987).

² SWAIN, R. (1971). *A decomposition algorithm for a class of facility location problems*. Ithaca,. PhD. Thesis, Cornell University apud CHIYOSHI, F.; GALVÃO, R. D.; MORABITO, R. (2000). O uso do modelo hipercubo na solução de problemas de localização probabilísticos. *Gestão & Produção*, v. 7, n. 2, p.146-74.

GOLDBERG et al. (1990) revisaram o modelo proposto por DASKIN (1983) que admite tempos de viagem determinísticos. Os autores consideraram os tempos de viagem como sendo variáveis aleatórias. Também foram estimadas, para todos os pares de nós da rede, a probabilidade de um veículo chegar ao local da ocorrência dentro de um intervalo de tempo estipulado, quando disponível. A diferença fundamental entre os dois modelos é que, em DASKIN (1983), a função objetivo diz respeito ao número de chamados que são cobertos (atendidos), ou seja, apresentam distância (tempos de resposta) dentro de um intervalo pré-estabelecido, e em GOLDBERG et al. (1990), considera-se como sendo sucessos, os chamados que apresentam distâncias (tempos de resposta) dentro deste intervalo, e fracasso quando são superiores a este valor.

Em outra direção, surgem, posteriormente, os problemas de localização de máxima disponibilidade, que são derivados dos problemas de localização de máxima cobertura. Estes problemas buscam localizar um número de servidores tal que o máximo número de solicitações de serviço tenha um servidor disponível para atendê-lo, a menos da distância (tempo) crítica, com determinado nível de confiabilidade.

REVELLE & HOGAN (1989) criaram um modelo que encontra o número mínimo de unidades de atendimento sujeito à restrição de que cada nó é coberto com probabilidade α . Os autores também assumem que as probabilidades individuais das unidades estarem ocupadas são independente, contudo podendo assumir valores distintos.

PIRKUKL & SCHILLING (1989) estenderam o problema introduzindo o atendimento *backup* quando a unidade de atendimento primário está ocupada, além das restrições de *workload* das unidades.

Demais detalhes sobre estes e outros trabalhos nesta área podem ser encontrados em OWEN & DASKIN (1998) e CHIYOSHI et al. (2000).

2.4 SOBRE O MODELO HIPERCUBO DE FILAS

Todos os problemas até então apresentados abordam técnicas para estudar diversos sistemas de atendimento público, em particular, os atendimentos emergenciais. Em todos os casos está inserida uma análise de desempenho do sistema em questão, pois todas as soluções encontradas apontam para critérios operacionais que melhoram o nível de serviço oferecido aos usuários.

Problemas como número de unidades de serviço em uma região, localização destas unidades ou então zoneamento de áreas tornam-se muito complexos quando se incorpora as características individuais destas unidades, como por exemplo, o seu estado (livre ou ocupado). Por este motivo, muitas das pesquisas nesta direção utilizam modelos de simulação, ou então admitem hipóteses simplificadoras para o caso analítico.

Alternativamente, pode-se utilizar o *modelo hipercubo de filas* desenvolvido por LARSON (1974a), que conserva a identidade das unidades de serviço. Este modelo constitui uma ferramenta analítica e descritiva que permite calcular uma ampla variedade de medidas de desempenho que podem auxiliar as decisões operacionais e de planejamento, dada uma configuração para o sistema (BRANDEAU & LARSON, 1986). Ele vem sendo utilizada até o presente momento para a análise de uma ampla variedade de serviços.

O modelo não constitui uma técnica de otimização, selecionando uma configuração ótima, mas fornece uma sofisticada avaliação de desempenho, permitindo assim a escolha da melhor alternativa operacional para o sistema (HALPERN, 1977).

Ele pode ser usado para analisar sistemas coordenados, onde o usuário que deseja receber algum tipo de serviço telefona para a central de chamados (que em geral coincide com a central de despachos) do sistema. O

administrador então despacha uma unidade de serviço mais próximo do local do chamado para realizar o atendimento. Caso nenhuma unidade esteja disponível, a solicitação entra em uma fila até que uma delas se torne disponível.

As complexidades geográficas e temporais da região também são abordadas pela análise, com base na teoria de filas espacialmente distribuídas e aproximações a partir do universo Markoviano.

Sistemas como bombeiros, ambulâncias, defesa civil, reparos em redes de energia elétrica, comunicações e água, entregas a domicílio (comida, remédios, por exemplo), guinchos e reparos mecânicos, visitas sociais, táxis, etc, podem ser representados por este modelo. A aplicação original do modelo foi para o problema de patrulhamento policial.

Um grande número de extensões do modelo foi reportado. CHELST & JARVIS (1979) modificaram o hipercubo para calcular a distribuição de probabilidade dos tempos de viagem. CHELST & BARLACH (1981) modelaram o envio de duas unidades de atendimento para o local do chamado.

Outras derivações importantes do modelo incluem as variações do tempo de atendimento, a calibração do tempo médio de atendimento e o algoritmo de barreiras. As variações do tempo médio de atendimento foram desenvolvidas por JARVIS (1975); a calibração do tempo médio de atendimento foi considerada por BRANDEAU & LARSON (1986) em uma aplicação na cidade de Boston, e o algoritmo de barreiras pode ser encontrado em LARSON & LI (1981).

LARSON & RITCH (1987) apresentaram uma reanálise do tempo de viagem dos veículos de patrulhamento da polícia de Nova Iorque, utilizando para isso o modelo hipercubo, e avaliam as consequências das novas estratégias de operação do sistema que passou a incorporar não somente a

região urbana da cidade, como também as áreas periféricas (subúrbios), estabelecendo níveis de prioridade nos chamados.

BATTA et al. (1989) usaram o hipercubo em conjunto com o modelo de máxima cobertura esperada de Daskin. Trata-se de um procedimento iterativo que utiliza o hipercubo para encontrar as probabilidades de se encontrar cada unidade de serviço ocupada. O processo inicia-se com um conjunto de localizações das unidades, e então insere o hipercubo para determinar a chance de estarem ocupadas, e baseado nestas probabilidades, calcula a cobertura esperada.

A questão de localização com máxima cobertura esperada, de grande valor o transporte médico de urgência, foi retomada por SAYDAM et al. (1994). Os autores utilizaram os modelos propostos por DASKIN (1983) e BATTA et al. (1989), dentre outros, para avaliar diversos cenários simulados de serviços médicos de urgência. Os resultados apontaram que, dentre os modelos utilizados, não há aquele que se destaca como mais preciso para os problemas de cobertura, e apóiam as recomendações de BATTA et al. (1989) ao indicar a utilização do modelo hipercubo para análises de pós-otimização.

No Brasil, alguns exemplos de aplicação do hipercubo são: o atendimento a interrupções na distribuição de energia elétrica em Florianópolis, Santa Catarina (ALBINO, 1994), a localização de ambulâncias em um trecho da BR 111 – SC (GONÇALVES et al. 1994 e GONÇALVES et al., 1995) e o balanceamento das *workloads* de ambulâncias no sistema Anjos do Asfalto da Rodovia Presidente Dutra (MENDONÇA, 1999 e MENDONÇA & MORABITO, 2000).

Observadas as potencialidades do modelo hipercubo e o seu rigor e precisão em descrever serviços emergenciais, o próximo capítulo apresenta a descrição analítica do modelo e discute os principais indicadores de desempenho que serão aplicados no estudo de caso apresentado no Capítulo 5.

3 MODELO HIPERCUBO DE FILAS

Este capítulo trata da formulação do modelo hipercubo de filas proposto para descrever sistemas em que servidores deslocam-se para oferecer algum tipo de serviço aos seus usuários (*server-to-customer service*), como é o caso do serviço de ambulâncias em um município ou região. De forma geral, serão apresentadas as hipótese que validam a aplicação do modelo, a descrição do espaço de estados e a solução do sistema de equações que produzem as probabilidades de equilíbrio como resposta. Com isto, são calculados diversos indicadores de desempenho que contribuem para o gerenciamento ou o planejamento do sistema.

O modelo desenvolvido por LARSON (1974a) e estudado por diversos autores (SWERSEY, 1994) constitui uma ferramenta específica para o planejamento e avaliação de sistemas de atendimento que envolvam demanda aleatória espacialmente distribuída, e cujo atendimento é realizado por servidores descentralizados ou não. Trata-se de uma importante ferramenta de apoio às decisões no sistema a partir dos indicadores de desempenho gerados

para determinar o nível de serviço oferecido. A essência do modelo é a teoria das filas, e uma breve revisão dos conceitos pode ser encontrada no Anexo A.

O modelo hipercubo baseia-se na partição da região de atuação do sistema em um conjunto finito de áreas geradoras de demanda, denominadas *átomos geográficos*. Cada átomo é considerado como uma fonte pontual independente de solicitação de serviço ao longo do tempo. Portanto, são consideradas as distribuições espacial e temporal dos chamados.

O atendimento é realizado por servidores (veículos) distribuídos na região e que, quando disponíveis, podem estar fixos em alguns pontos (bases), ou em movimento (neste caso, sua localização deve ser conhecida ao menos probabilisticamente). A área de cobertura primária de um servidor é definida pelo conjunto dos átomos que este servidor atende prioritariamente, ou seja, é o primeiro a ser chamado. O modelo também trata situações em que se tem átomos com mais de um servidor preferencial. Nos casos onde este servidor está ocupado, outros servidores são chamados para atender a solicitação, de acordo com ordem de preferência de atendimento da área onde ocorreu o evento.

A Figura 3.1 exemplifica uma região particionada em 13 átomos, e a distribuição espacial de 4 servidores fixos.

FIGURA 3.1 - Representação da área de atuação de um sistema

O espaço de estados descreve o número (N) e a disponibilidade dos servidores. Cada servidor pode se encontrar *livre* (0) ou *ocupado* (1) em um certo instante. Um estado particular do sistema é descrito pela lista ordenada de servidores que estão livres e ocupados. A representação é vetorial, com dimensão igual ao número de servidores do sistema, ordenados do primeiro ao último, da direita para a esquerda nas posições do vetor.

Por exemplo, para um sistema com três servidores, o estado (011) corresponde à situação onde o servidor número 1 está ocupado, o servidor número 2 está ocupado e o servidor número 3 está livre. Desta maneira, um sistema com três servidores tem seu espaço de estados em $\mathbf{R}^3$, e é dado pelo conjunto de vértices do cubo de arestas iguais a 1. À medida que o número de servidores cresce, a representação se dá no espaço N -dimensional ($\mathbf{R}^N$), ou seja, pelo conjunto de vértices de um *hipercubo*, de onde deriva o nome do modelo. A Figura 3.2 mostra o espaço de estados para sistemas com $N = 3$.

FIGURA 3.2 - Espaço de estados para sistemas com três servidores

O modelo trata tanto de sistemas em que não é permitida a formação de fila, quanto de sistemas em que os chamados quando chegam e não há servidor disponível aguardam para receber o atendimento. Neste caso, o espaço de estados deve ser acrescido de uma cauda junto ao estado totalmente ocupado (11...1), representando os usuários que esperam pelo serviço. Estes usuários são atendidos, à medida em que os servidores tornam-se livres, segundo a disciplina *FCFS*. A Figura 3.3 descreve o espaço de estados para um hipercubo com $N = 3$, considerando a formação de fila de espera (sem limitação de número de usuários aguardando por atendimento).

FIGURA 3.3 – Espaço de estados para sistemas com 3 servidores e cauda infinita

Quando o serviço analisado envolve critérios de prioridade no atendimento, o modelo pode ser adaptado para representar estes critérios. Este trabalho, por exemplo, adaptará o modelo original para representar um sistema onde existem classes distintas de usuários: os que necessitam do atendimento de uma equipe básica de suporte de vida e os que exigem um atendimento mais especializado.

3.1 INSERINDO PRIORIDADES

Muitos dos sistemas do tipo *server-to-customer* admitem implícita ou explicitamente critérios de prioridade para os chamados que recebem. Porém, nem sempre é necessário representar tais critérios diretamente no modelo, como por exemplo, em casos de planejamento da área de cobertura de um servidor ou então para problemas de localização das unidades de atendimento. Entretanto, existem casos onde estes critérios devem ser introduzidos, como por exemplo, para um sistema de transporte médico de urgência onde se tem uma ambulância especializada em atendimentos avançados, como é o caso das UTI's (unidades de tratamento intensivo) móveis, e deseja-se avaliar o nível de serviço que está sendo oferecido.

Este sistema opera com o seguinte critério: o chamado emergencial é avaliado e classificado como sendo do tipo avançado, ou seja, que necessita da intervenção de uma equipe de UTI, ou por exemplo, do tipo básico, onde uma equipe simples de socorro pode realizar o atendimento. Ao ocorrer um chamado do tipo avançado, a equipe de UTI é enviada. Caso não esteja disponível, uma equipe básica é enviada a fim de que as vítimas não sofram longas esperas para receber os primeiros socorros.

Nestes casos, o modelo pode ser adaptado para refletir os critérios de prioridade, bem como as regras de despacho de servidores.

Nesta situação, o analista deve particionar cada átomo do sistema em duas áreas distintas e independentes: uma gerando apenas chamados do tipo avançado, e outra gerando chamados do tipo básico. Para a área geradora de chamados avançados, os servidores primários são aqueles que realizam atendimento especializado, sendo que os demais comportam-se como *backup*. Para a área geradora dos chamados básicos a regra é a mesma: os servidores primários são aqueles que oferecem suporte básico de vida, ficando os

servidores avançados na posição de *backup*. Esta técnica foi denominada por LARSON & ODoni (1981) como sendo o “processo de camadas”.

Os indicadores de desempenho produzidos são relacionados a cada classe de usuário e cada tipo de servidor, separadamente.

O modelo também pode ser aplicado para sistemas com mais de duas classes de usuários; entretanto, o volume de informações produzidas pelo modelo vai se tornando impraticável, prejudicando o nível global da análise. Neste sentido, um *trade-off* entre o nível de detalhamento da representação do sistema e a habilidade de compreensão, interpretação e ação do analista é estabelecido (LARSON & ODoni, 1981).

Esquemas complexos como manutenção de chamados de baixa prioridade para serem atendidos pelo servidor primário, interrupção de um serviço de baixa prioridade para atender a um chamado de alta prioridade, ou o posicionamento seletivo dos chamados em função do seu grau de prioridade não são abrangidos pelo modelo. Nestes casos, técnicas mais flexíveis devem ser aplicadas (como por exemplo, a simulação computacional).

3.2 HIPÓTESES DO MODELO

A seguir, serão apresentadas as nove hipóteses críticas que devem ser verificadas para a aplicação do modelo, segundo a descrição de LARSON & ODoni (1981).

- 1) Átomos geográficos: a área na qual o sistema opera é particionada em N_A pequenas regiões que se complementam geograficamente, cada qual representando uma fonte independente de solicitações pelo serviço. Toda pequena região é atendida por ao menos um servidor do sistema.
- 2) Chegadas Poissonianas independentes: as solicitações por serviço (chamados) chegam independentemente em cada átomo j segundo

processos de Poisson, com taxas médias λ_j ($j=1,2,\dots,N_A$), constantes no tempo, possíveis de serem medidas ou estimadas. Embora esta hipótese pareça muito restritiva, ela é freqüentemente satisfeita em diversos sistemas reais.

- 3) Tempos de viagem: os tempos médios de viagem τ_{ij} ($i, j = 1, 2, \dots, N_A$) entre o átomo i e o átomo j deverão ser conhecidos ou estimados através dos conceitos de probabilidade geométrica.
- 4) Servidores: o sistema é composto por N servidores (distintos ou não) espacialmente distribuídos, que podem deslocar-se e atender qualquer um dos átomos. Esta é uma hipótese que pode ser facilmente relaxada para representar políticas de despacho particulares (por exemplo, MENDONÇA, 1999 modelou o sistema *Anjos do Asfalto*, na Rodovia Presidente Dutra, despachando apenas os dois servidores mais próximos do local da ocorrência).
- 5) Localização dos servidores: a localização de cada servidor, quando disponível, deve ser conhecida. No caso de servidores fixos, a localização é a sua base; no caso de servidores que se movem dentro de uma determinada área (por exemplo, viaturas de patrulhamento policial), esta localização deve ser conhecida ao menos probabilisticamente.
- 6) Despacho (ou alocação) de um servidor: apenas um servidor é despachado para atender um chamado. O modelo não representa de maneira adequada casos em que mais de um servidor é despachado para atender o mesmo chamado; porém, em certas situações reais, o conjunto de servidores despachados para uma grande catástrofe, por exemplo, pode ser visto como sendo um único servidor. Se não houver veículo disponível para realizar o atendimento, o chamado poderá entrar em uma fila (nos sistemas que permitem a formação de filas), ou então ocorrer a perda do

chamado (nos casos em que não se admite fila), podendo o evento ser transferido para outro sistema de atendimento.

- 7) Política de despacho dos servidores: há uma lista fixa de preferências de despacho para cada átomo. Se o primeiro servidor desta lista estiver disponível, ele é despachado para atender o chamado; caso contrário, o próximo servidor disponível da lista é despachado (*backup*). A lista de preferência de despachos é fixada e permanece inalterada durante a operação do sistema. Adaptações para os casos em que um átomo tem mais de um servidor preferencial também podem ser incorporadas ao modelo. Isto pode ser feito através da introdução da distribuição de frequências de despachos de cada servidor para cada átomo nas equações de balanço do sistema, ou considerando um número suficiente de listas de preferências de despacho geradas aleatoriamente, de forma a representar as possíveis chances dos servidores primários de cada átomo serem despachados para atender um chamado em cada cenário investigado. Este processo também deve ser repetido para os veículos *backup* (BURWELL et al., 1993).
- 8) Tempos de atendimento: o tempo total de atendimento de um chamado é composto por um tempo de preparo do servidor (*setup time*), tempo de viagem do servidor até o local da ocorrência, tempo de execução do serviço junto ao usuário (*tempo em cena*) e o tempo de retorno à base. Os servidores podem ter taxas médias de atendimento μ_n distintas, para cada servidor $n = 1, 2, \dots, N$. No caso de sistemas que permitem a formação de fila, o modelo é ainda mais eficiente à medida que os tempos médios de atendimento aproximam-se dos respectivos desvios-padrão, ou seja, tempos de atendimento exponencialmente distribuídos. Segundo os autores, desvios razoáveis desta hipótese não alteram sensivelmente a precisão dos resultados. Caso o sistema não permita a formação de filas, esta hipótese é ainda menos necessária, visto que os sistemas M/M/N/N e

M/G/N/N apresentam a mesma distribuição de equilíbrio (GROSS & HARRIS, 1974).

- 9) Dependência do tempo de atendimento em relação ao tempo de viagem: variações no tempo total de atendimento devidas às variações no tempo de viagem são assumidas como sendo de segunda ordem, quando comparadas às variações dos tempos junto ao paciente (em cena) e/ou tempo de preparação da equipe (*setup*). Esta hipótese, que limita a aplicabilidade do modelo, é freqüentemente verificada em serviços urbanos, e menos observada em serviços em rodovia (LARSON & ODoni, 1981).

Na prática, poucos são os sistemas reais que aderem completamente as nove hipóteses apresentadas. A decisão de se aplicar ou não o modelo deve levar em conta o quanto o sistema real não se ajusta à rigidez das hipóteses, contra as limitações ou a complexidade do uso de modelos analíticos alternativos (LARSON & ODoni, 1981).

3.3 TRANSIÇÕES ENTRE ESTADOS

No modelo hipercubo, as transições entre estados ocorrem de modo idêntico aos modelos clássicos de filas. Admite-se que apenas um servidor é despachado para atender um chamado, e que a probabilidade de chegarem dois chamados exatamente no mesmo instante, e a probabilidade de dois atendimentos terminarem exatamente no mesmo instante são nulas.

Em síntese, qualquer transição de um passo é permitida, e todas as transições com mais de um passo não são (processos de nascimento e morte). Esta característica é decorrente da natureza das distribuições dos intervalos entre solicitações e dos tempos de atendimentos (ambos admitidos como sendo processos Markovianos).

3.3.1 Taxas de transição

A taxa de transição de um estado do sistema para outro estado adjacente pode ser de natureza ascendente ou descendente.

Genericamente, as taxas de transições ascendentes são dadas pela soma da taxa de chegada de chamados da área de cobertura primária do servidor que passa para o estado “ocupado”, com a taxa de chegada de chamados dos átomos que tem o servidor em questão como o primeiro *backup*, e cujo servidor de prioridade máxima esteja ocupado. Raciocínio análogo é feito para as taxas descendentes.

3.4 EQUAÇÕES DE EQUILÍBRIO

A solução do modelo é dada partindo-se da construção do conjunto de equações de equilíbrio para o sistema, e os resultados baseiam-se nos valores das probabilidades de estado dos modelo clássicos de filas.

As equações de equilíbrio são definidas supondo-se que o sistema atinja equilíbrio, ou seja *steady state flow*. Para cada estado do sistema, o fluxo com que se entra neste estado deve ser igual ao fluxo com que se sai deste estado.

Considere, por exemplo, um simples sistema com $N_A = 3$ átomos e $N = 3$ servidores, sendo os servidores 1, 2 e 3 primários para os átomos 1, 2 e 3, respectivamente.

a) *Sistema vazio – estado (000): S_0*

Iniciando-se com o estado (000), onde todos os servidores estão disponíveis, tem-se que o sistema passa do estado (000) para o estado (001) quando ocorre um chamado com origem no átomo 1. A taxa de ocorrência deste evento é λ_1 . O mesmo acontece para o estado (010), com taxa λ_2 e para o

estado (100), com taxa λ_3 . Conseqüentemente, a taxa total de transição do estado (000) para outros estados é $\lambda = \lambda_1 + \lambda_2 + \lambda_3$.

No sentido inverso, o estado (000) pode ser alcançado a partir do estado (001) quando o servidor 1 conclui o atendimento, com taxa μ_1 , da mesma maneira, a partir do estado (010), com taxa μ_2 e a partir de (100), com taxa μ_3 .

Assim, a equação de equilíbrio para o estado (000) é:

$$\lambda \cdot p_{000} = \mu_1 \cdot p_{001} + \mu_2 \cdot p_{010} + \mu_3 \cdot p_{100} \quad (1)$$

onde p_B é a probabilidade de equilíbrio do estado B (por exemplo, p_{000} é a probabilidade de equilíbrio do estado B = (000)).

b) 1 servidor ocupado – estados (001), (010) e (100): S_1

Quando o sistema encontra-se em (001), qualquer chegada o levará para outro estado (com dois servidores ocupados) e também é possível a transição para o estado (000) quando o servidor 1 conclui o atendimento. Logo, a taxa de transição de (001) para outros estados será igual a $\lambda + \mu_1$. No sentido inverso, as transições de outros estados para (001) são:

- i. sistema no estado (000) e chegada de um chamado com origem no átomo 1;
- ii. sistema no estado (011) e conclusão do atendimento realizado pelo servidor 2, e
- iii. sistema no estado (101) e conclusão do atendimento realizado pelo servidor 3.

Portanto, a equação de equilíbrio do estado (001) é:

$$(\lambda + \mu_1) \cdot p_{001} = \lambda_1 \cdot p_{000} + \mu_2 \cdot p_{011} + \mu_3 \cdot p_{101} \quad (2)$$

Com raciocínio análogo, constrói-se as equações de equilíbrio para os estados (010) e (100):

$$(\lambda + \mu_2) \cdot p_{010} = \lambda_2 \cdot p_{000} + \mu_1 \cdot p_{011} + \mu_3 \cdot p_{110} \quad (3)$$

e

$$(\lambda + \mu_3) \cdot p_{100} = \lambda_3 \cdot p_{000} + \mu_1 \cdot p_{101} + \mu_2 \cdot p_{110} \quad (4)$$

c) 2 servidores ocupados – estados (011), (101) e (110): S_2

Quando o sistema encontra-se em (011), qualquer chegada o levará para o estado (111), e também é possível a transição para o estado (001) através da conclusão do atendimento realizado pelo servidor 2, bem como a transição para (010) através da conclusão do atendimento realizado pelo servidor 1. Assim, a taxa total de transição do estado (011) para outros estados é igual a $\lambda + \mu_1 + \mu_2$. No sentido inverso, tem-se, primeiramente, a transição a partir de (001). Este caso acontece quando ocorre a chegada de um chamado no átomo 2, e também quando ocorre um chamado no átomo 1, estando o servidor 1 ocupado, despachando-se então o primeiro *backup* (servidor 2). Outras possíveis transições são:

- i. sistema no estado (010) e chegada de um chamado com origem no átomo 1, e
- ii. sistema no estado (111) e conclusão do atendimento realizado pelo servidor 3.

Então, a equação de equilíbrio para (011) é:

$$(\lambda + \mu_2 + \mu_1) \cdot p_{011} = (\lambda_1 + \lambda_2) \cdot p_{001} + \lambda_1 \cdot p_{010} + \mu_3 \cdot p_{111} \quad (5)$$

De forma análoga, para os estados (101) e (110):

$$(\lambda + \mu_3 + \mu_1) \cdot p_{101} = \lambda_3 \cdot p_{001} + (\lambda_1 + \lambda_3) \cdot p_{100} + \mu_2 \cdot p_{111} \quad (6)$$

e

$$(\lambda + \mu_2 + \mu_3) \cdot p_{110} = (\lambda_3 + \lambda_2) \cdot p_{010} + \lambda_2 \cdot p_{100} + \mu_1 \cdot p_{111} \quad (7)$$

d) 3 servidores ocupados (sistema cheio) – estado (111): S_3

Supondo que não há usuários em fila, qualquer chegada ou conclusão de atendimento de algum servidor provocará a transição do sistema para fora deste estado. Por outro lado, o estado (111) pode ser alcançado a partir dos estados com dois servidores ocupados, através da chegada de um chamado em qualquer um dos átomos. Nesta situação, o único servidor livre é despachado para realizar o atendimento, sendo ele preferencial, 1º *backup* ou 2º *backup*. Por fim, o estado (111) também pode ser alcançado a partir do estado S_4 , ou seja, quando 4 usuários estão presentes no sistema, isto é, três recebendo atendimento e 1 aguardando em fila, através da conclusão do atendimento de qualquer um dos servidores. Sendo a probabilidade $p_4 = P(S_4)$, a equação de equilíbrio de (111) é:

$$(\lambda + \mu) \cdot p_{111} = \lambda \cdot p_{011} + \lambda \cdot p_{101} + \lambda \cdot p_{110} + \mu \cdot p_4$$

onde $\mu = \mu_1 + \mu_2 + \mu_3$.

Este procedimento poderia continuar para os estados $S_5, S_6, S_7, \dots$, resultando em um sistema infinito de equações. Entretanto, deve-se respeitar as condições de equilíbrio do sistema, que exigem que as taxas de transição entre os estados (111) e S_4 sejam iguais, isto é, $\lambda \cdot p_{111} = \mu \cdot p_4$. Se esta relação não fosse obedecida e, por exemplo, $\lambda \cdot p_{111} > \mu \cdot p_4$, o sistema estaria em estado transiente, e a cauda que representa a fila estaria em fase de crescimento.

Desta maneira, a equação de equilíbrio do estado (111) resume-se em:

$$\mu \cdot p_{111} = \lambda \cdot p_{011} + \lambda \cdot p_{101} + \lambda \cdot p_{110} \quad (8)$$

Observando as equações de (1) a (8), onde estão representadas todas as possíveis transições de estado para um sistema com três servidores, observa-se um sistema finito de equações, em relação às variáveis p_{000} , p_{001} , p_{010} , p_{100} , p_{011} , p_{101} , p_{110} e p_{111} .

As probabilidades de equilíbrio dos demais estados (S_4 , S_5 , S_6 , ...), isto é, p_4 , p_5 , p_6 , ... são facilmente obtidas. Basta substituir os valores de p_{000} , p_{001} , ..., p_{111} nas equações de equilíbrio dos respectivos estados S_4 , S_5 , S_6 , ..., conforme apresentado a seguir.

3.5 INSUFICIÊNCIA DAS EQUAÇÕES DE EQUILÍBRIO

Ao tentar resolver o sistema de equações determinado acima através da forma matricial $A \cdot x = b$, constata-se que as equações conduzem a um sistema com $b = 0$. Portanto, um sistema homogêneo, com solução trivial $p_{ijk} = 0$, para todo $i, j, k = 0, 1$. Observa-se também que trata-se de um sistema possível indeterminado pois, atribuindo-se valor a uma das probabilidades, determina-se as demais a partir desta condição. CHIYOSHI et al. (2000) explicam esta indefinição: as equações apenas impõem condições de equilíbrio para cada possível estado do sistema ((000), (001), ..., (111)), mas nada especifica sobre a forma como a massa total de probabilidades se distribui entre estes estados e os estados da cauda (S_4 , S_5 , S_6 , ...).

Também expõem que a forma natural de eliminar esta indeterminação é através da introdução de uma equação de normalização, ou seja, considerando que a soma das probabilidades de todos os possíveis estados do sistema deve ser igual a 1:

$$p_{000} + p_{001} + p_{010} + \dots + p_{111} + p_4 + p_5 + p_6 + \dots = 1 \quad (9)$$

A equação (9) pode ser simplificada quando trata-se de sistemas que permitem a formação de filas, e não impõe limitações sobre o número de usuários que podem aguardar pelo atendimento.

CHIYOSHI et al. (2000) apresentam tal simplificação:

Partindo-se da hipótese de que $\lambda \cdot p_{111} = \mu \cdot p_4$, e sendo a relação $\frac{\lambda}{\mu} = \rho$,

com $\rho < 1$, $\lambda = \lambda_1 + \lambda_2 + \lambda_3$ e $\mu = \mu_1 + \mu_2 + \mu_3$, tem-se $p_4 = \rho \cdot p_{111}$.

Aplicando a mesma condição de equilíbrio para os estados S_k e S_{k+1} , $k=4, 5, 6, \dots$, tem-se:

$$\begin{aligned} p_5 &= \tilde{n} \cdot p_4 = \tilde{n}^2 \cdot p_{111} \\ p_6 &= \tilde{n} \cdot p_5 = \tilde{n}^3 \cdot p_{111} \\ &\vdots \\ p_{k+1} &= \tilde{n} \cdot p_k = \tilde{n}^{k-N} \cdot p_{111} \\ &\vdots \end{aligned}$$

$$\text{Logo: } p_{111} + p_4 + p_5 + p_6 + \dots = p_{111} + \rho \cdot p_{111} + \rho^2 \cdot p_{111} + \dots = p_{111} \cdot \sum_{j=0}^{\infty} \tilde{n}^j.$$

$$\text{Como } \rho < 1, \sum_{j=0}^{\infty} \tilde{n}^j = \frac{1}{1-\rho} \text{ (série geométrica de razão igual a } \rho \text{)}.$$

$$\text{Portanto, } p_{111} + p_4 + p_5 + p_6 + \dots = \frac{p_{111}}{1-\rho}.$$

Logo, a equação (9) pode ser reescrita como sendo:

$$p_{000} + p_{001} + p_{010} + \dots + \frac{p_{111}}{1-\rho} = 1 \quad (10)$$

Substituindo-se qualquer uma das equações (1) – (8) por esta nova equação (10), o sistema torna-se determinado, com $2^3 = 8$ equações

linearmente independentes, gerando todas as probabilidades de equilíbrio dos respectivos estados.

3.5.1 Adaptações com relação às limitações da fila

Sistemas que não admitem formação de fila, ou que limitam o número de usuários aguardando atendimento, não podem utilizar o resultado apresentado acima.

3.5.1.1 Modelo hipercubo sem cauda (fila = 0)

Neste caso, para modelar a equação (9) deve-se desconsiderar a existência dos estados que formam a cauda, ou seja, para $N_A = 3$ e $N = 3$:

$$p_{000} + p_{001} + p_{010} + \dots + p_{111} = 1 \quad (11)$$

A solução do sistema de equações também se dá por meio da substituição de uma das equações de equilíbrio (1)-(8) pela equação (11), o que resulta em um sistema com $2^3 = 8$ equações linearmente independentes.

MENDONÇA (1999) utilizou tal adaptação para o Sistema Anjos do Asfalto cuja política operacional não permite a formação de filas de espera, isto é, a probabilidade de existência de fila é identicamente nula ($p_Q=0$).

Esta variação provoca alterações nas medidas de desempenho descritas, em especial nas equações das frequências de despachos, uma vez que $f_{nj}^{[2]} = 0$, e $f_{nj}^{[1]}$ passa a ser condicionada à condição de que o chamado não provém de uma fila.

Maiores detalhes podem ser encontrados em MENDONÇA (1999) e MENDONÇA & MORABITO (2000).

3.5.1.2 Modelo hipercubo com cauda limitada

Neste caso, deve-se considerar um número finito de estados na cauda, tantos quanto o número permitido de usuários que podem aguardar em fila imposto pela política operacional do sistema.

Esta situação aplica-se na representação da política operacional do estudo de caso realizado no SAMU-192 de Campinas, que limita a fila em exatamente o número de ambulâncias em operação.

Considerando o sistema com $N_A = 3$ átomos e $N = 3$ servidores, e admitindo a política operacional do SAMU-192 de Campinas, onde a fila permitida é no máximo igual ao número de ambulâncias disponíveis na operação, a cauda representada no espaço de estado da Figura 3.3 passa a ser limitada no estado S_6 , que representa o sistema com 3 servidores ocupados e 3 usuários aguardando em fila.

O sistema de equações de equilíbrio para este caso passa a ter 3 novas equações, a partir das equações (1)-(8), que representam a cauda do hipercubo.

$$\begin{cases} \ddot{e} \cdot p_{111} = \dot{i} \cdot p_4 \\ \ddot{e} \cdot p_4 = \dot{i} \cdot p_5 \\ \ddot{e} \cdot p_5 = \dot{i} \cdot p_6 \end{cases} \Rightarrow \begin{cases} p_4 = \rho \cdot p_{111} \\ p_5 = \rho^2 \cdot p_{111} \\ p_6 = \rho^3 \cdot p_{111} \end{cases}$$

onde $\lambda = \lambda_1 + \lambda_2 + \lambda_3$, $\mu = \mu_1 + \mu_2 + \mu_3$ e $\rho = \frac{\lambda}{\mu}$.

A equação (9) passa então a ser escrita como:

$$p_{000} + p_{001} + p_{010} + \dots + p_{111} + \rho \cdot p_{111} + \rho^2 \cdot p_{111} + \rho^3 \cdot p_{111} = 1$$

ou

$$p_{000} + p_{001} + p_{010} + \dots + \sum_{n=0}^3 \rho^n \cdot p_{111} = 1 \quad (12)$$

O sistema de equações é solucionado de maneira análoga, substituindo-se uma das equações (1)–(8) pela equação (12), o que resulta em $2^3=8$ equações linearmente independentes, conforme o sistema anterior.

Esta variação provoca alterações nas medidas de desempenho em consequência das variações dos valores das probabilidades de equilíbrio do sistema. Neste caso, $p_Q \neq 0$ e pode-se calcular o tempo médio de espera de um usuário qualquer, $\overline{T_Q}$.

Em ambos os casos associa-se a probabilidade de perda de chamados para o sistema. Para o caso em que não se permite a formação de fila, as perdas geralmente se dão quando se atinge o estado (111), ou devem ser determinadas de acordo com a política de despachos admitida, como é o caso do sistema Anjos do Asfalto. No segundo caso, as perdas se dão em S_6 , ou seja, quando todos os servidores estão ocupados e existem 3 usuários esperando para receber atendimento. Em particular, para o SAMU-192 de Campinas, que opera com 10 ambulâncias em sua configuração original, estas perdas se dão quando o sistema encontra-se no estado S_{20} (todas as 10 ambulâncias ocupadas e 10 chamados aguardando por atendimento).

3.5.2 Relação com o modelo clássico M/M/N

LARSON & ODoni (1981) mostraram que, se os servidores são idênticos e têm a mesma taxa de atendimento ($\mu_1 = \mu_2 = \dots = \mu_N$), o modelo estabelece uma relação com o modelo clássico M/M/N, e isto permite que suas equações sejam incorporadas ao sistema da seguinte forma:

$$\begin{aligned} p_{00\dots0} &= P_{M/M/N}(S_0) \\ p_{00\dots01} + p_{00\dots010} + p_{10\dots00} &= P_{M/M/N}(S_1) \\ &\vdots \\ p_{11\dots1} &= P_{M/M/N}(S_N) \end{aligned}$$

onde $P_{M/M/N}(S_n)$ representa a probabilidade de um sistema M/M/N encontrar-se no estado S_n , $n=1,2,...N$. Estas equações são denominadas *equações dos hiperplanos* definidos pelo modelo.

Caso o sistema apresente servidores heterogêneos, ou seja, com tempos médios de atendimento diferentes, esta relação deixa de ser válida. As variações nas taxas médias de atendimento de cada servidor, além de alterar as o sistema de equações de equilíbrio, provocam mudanças em algumas medidas de desempenho para o sistema. Como será visto no Capítulo 5, o SAMU-192 de Campinas apresenta servidores distintos. Neste sentido, as medidas de desempenho seguintes já foram determinadas considerando-se as diferenças entre os tempos médios de atendimento de cada servidor.

3.6 DESEMPENHO DO SISTEMA

Satisfeitas as hipóteses 1 a 9, o modelo fornece medidas de desempenho interessantes para o gerenciamento do sistema, tanto em nível de servidor, individualmente, quanto para um átomo específico ou uma região de cobertura, ou para o sistema como um todo.

As medidas apresentadas a seguir foram calculadas para sistemas com N servidores distintos. Sistemas em que os servidores podem ser considerados idênticos são tratados como um caso particular.

3.6.1 *Workload*

A carga de trabalho, também chamada *workload* de um servidor (ρ_n), representa a fração de tempo em que o servidor permanece ocupado, e pode ser facilmente obtida através da soma das probabilidades dos estados em que o servidor está ocupado, ou seja:

$$\begin{cases} \rho_1 = p_{00\dots01} + p_{00\dots011} + \dots + p_{11\dots11} + p_Q \\ \rho_2 = p_{00\dots010} + p_{00\dots011} + \dots + p_{11\dots11} + p_Q \\ \vdots \\ \rho_N = p_{100\dots000} + p_{100\dots001} + \dots + p_{11\dots11} + p_Q \end{cases} \quad (13)$$

onde $p_Q = p_4 + p_5 + p_6 + \dots = 1 - (p_{00\dots00} + p_{00\dots10} + \dots + p_{11\dots11})$ é a probabilidade de existência de fila no sistema.

3.6.2 Frequências de despachos

Outra medida bastante útil é a frequência de despachos de determinado servidor para determinado átomo, ou seja, a fração de todos os despachos do servidor n ao átomo j , f_{nj} .

Esta medida é decomposta em duas parcelas: a primeira, $f_{nj}^{[1]}$, correspondente à fração de todos os despachos do servidor n para o átomo j que não implica em tempo de espera para o usuário; a segunda parcela, $f_{nj}^{[2]}$, representa a fração de todos os despachos do servidor n para o átomo j sujeitos a esperas pelo atendimento.

$$f_{nj} = f_{nj}^{[1]} + f_{nj}^{[2]} = \underbrace{\frac{\lambda_j}{\lambda} \cdot \sum_{B \in E_{nj}} p_B}_{f_{nj}^{[1]}} + \underbrace{\frac{\lambda_j}{\lambda} \cdot p_s \cdot \frac{i_n}{i}}_{f_{nj}^{[2]}} \quad (14)$$

onde E_{nj} é o conjunto dos estados nos quais um chamado do átomo j é atendido pelo servidor n , e $p_s = p_Q + p_{11\dots11}$, que é chamada probabilidade de saturação do sistema.

Em um sistema particular onde os tempos médios de atendimento de todos os servidores são idênticos, iguais a 1 unidade de tempo, com taxas médias de atendimento $\mu_1 = \mu_2 = \dots = \mu_N = 1$, f_{nj} escreve-se simplesmente como:

$$f_{nj} = \frac{\lambda_j}{\lambda} \cdot \sum_{B \in E_{nj}} p_B + \frac{1}{N} \cdot \frac{\lambda_j}{\lambda} \cdot p_s$$

LARSON & ODoni (1981) apresentaram outras medidas como a fração de despachos em que o servidor n atuou como *backup*, a fração de despachos de *backup* para o átomo j , a fração total de despachos do tipo *backup* no sistema, dentre outras.

3.6.3 Tempos médios de viagem

Estes valores são obtidos a partir da matriz origem-destino dos tempos de viagem entre todos os pares de átomos (τ_{ij}). Em geral, refletem a influência de fatores como configurações do tráfego local, presença de barreiras, horário, etc. Desta forma, o tempo gasto por um servidor qualquer para ir do átomo i até o átomo j , não necessariamente coincide com o tempo gasto para ir de j para i ; isto é, nem sempre $\tau_{ij} = \tau_{ji}$. Caso estes valores não possam ser medidos no próprio sistema, eles devem ser estimados utilizando-se conceitos de probabilidade geométrica (LARSON & ODoni, 1981).

3.6.3.1 Tempo médio de viagem para o sistema

Para calcular o tempo médio de viagem para o sistema ($\bar{T}$), é necessário conhecer a localização de um servidor, quando despachado para atender a um chamado, o tempo médio necessário para um servidor n , quando disponível, viajar até o átomo j , e o tempo médio de espera de um chamado que está em fila.

a) matriz de localizações

A representação destas localizações é feita de forma bastante genérica: ela se apresenta de forma matricial ($L = [l_{nj}]$), onde os elementos da matriz representam a probabilidade de um servidor estar localizado em um

determinado átomo, quando disponível. L é, necessariamente, uma matriz estocástica, ou seja $\sum_{j=1}^{N_A} l_{nj} = 1$, com $l_{nj} = 1$ se o servidor está localizado no átomo j , e $l_{nk} = 0$ para todo $k \neq j$.

b) *tempo médio de viagem para um servidor deslocar-se até determinado átomo*

$$t_{nj} = \sum_{k=1}^{N_A} l_{nk} \cdot \tau_{kj} \quad n=1,2,\dots,N \text{ e } j=1,2,\dots,N_A \quad (15)$$

c) *tempo médio de espera para chamados em fila*

$$\overline{T_Q} = \sum_{i=1}^{N_A} \sum_{j=1}^{N_A} \frac{\lambda_i \cdot \lambda_j}{\lambda^2} \cdot \tau_{ij} \quad (16)$$

onde as razões $\frac{\lambda_j}{\lambda}$ e $\frac{\lambda_i}{\lambda}$ correspondem à probabilidade de um chamado que está em fila ter sido gerado no átomo j , e à probabilidade deste chamado ser atendido por um servidor localizado no átomo i , respectivamente.

Considerando também a probabilidade de saturação do sistema (p_s), o tempo médio de viagem para o sistema pode ser obtido como:

$$\overline{T} = \sum_{n=1}^N \sum_{j=1}^{N_A} f_{nj}^{[1]} \cdot t_{nj} + p_s \cdot \overline{T_Q} \quad (17)$$

3.6.3.2 *Tempo médio de viagem para cada átomo*

Outra medida importante e que reflete o nível de serviço oferecido pelo sistema, é o tempo médio de viagem para cada átomo j ($\overline{T_j}$), chamado de tempo médio de resposta (TR) nos sistemas reais.

$$\overline{T}_j = \frac{\sum_{n=1}^N f_{nj}^{[1]} \cdot t_{nj}}{\sum_{n=1}^N f_{nj}^{[1]}} \cdot (1 - p_s) + \sum_{i=1}^{N_A} \left(\frac{\lambda_i}{\lambda} \right) \cdot \tau_{ij} \cdot p_s \quad (18)$$

3.6.3.3 Tempo médio de viagem para cada servidor

Analogamente, os tempos médios de viagem de cada servidor são:

$$\overline{TU}_n = \frac{\sum_{j=1}^{N_A} f_{nj}^{[1]} \cdot t_{nj} + \left(\frac{\overline{T_Q} \cdot p_s}{N} \right)}{\sum_{j=1}^{N_A} f_{nj}^{[1]} + \left(\frac{p_s}{N} \right)} \quad (19)$$

De maneira semelhante, outras medidas podem ser calculadas. Alguns exemplos estão em LARSON (1974).

3.7 CALIBRAÇÃO DOS TEMPOS DE ATENDIMENTO

É comum observar, em determinados serviços de emergência, que a parcela relativa ao tempo de viagem dos servidores representa uma fração considerável do tempo total de atendimento, como é o caso dos serviços de ambulância, onde este tempo é composto pelo tempo de viagem até o local da ocorrência, tempo de viagem do local da ocorrência até um hospital ou mesmo domicílio das vítimas e tempo de viagem de retorno à base. Em casos como este é necessário que o tempo de atendimento de cada servidor seja ajustado, a fim de refletir os vários fatores que o influenciam. Isto é feito através do processo de calibração de μ_n^{-1} , utilizando para isto, a expressão do tempo médio de viagem do servidor n , $\overline{TU}_n$.

O processo consiste em verificar se existe diferença significativa entre os tempos médios de atendimento admitidos como parâmetros e os tempos gerados pelo modelo, ou seja, se a soma de todos os tempos que compõem o

tempo médio de atendimento do servidor n for diferente de μ_n^{-1} , admitido no início da modelagem, então, o modelo deve ser executado novamente, utilizando como parâmetro este novo tempo médio de atendimento. Este procedimento deve seguir até que a diferença entre os tempos médios de atendimento admitidos como parâmetros e os tempos médios de atendimento gerados pelo modelo sejam suficientemente próximos.

Experiências de diversos autores mostram que este procedimento costuma convergir em duas ou três iterações, para uma precisão razoável dos valores de μ_n^{-1} (LARSON & ODoni, 1981).

3.8 A SOLUÇÃO DO MODELO

Encontrar a solução do modelo hipercubo significa obter os elementos necessários para determinar suas características operacionais básicas. O método exato de solução consiste em obter as probabilidades associadas aos estados descritos pelos vértices do hipercubo. Sistemas que operam com N servidores, possuem um sistema linear com 2^N equações cujo resultado são os valores das probabilidades de equilíbrio do sistema. Portanto, mesmo em casos onde N é moderado, é essencial assegurar a viabilidade computacional de se resolver o problema. Já o método aproximado descrito em LARSON (1975), envolve um sistema não linear de equações que tem como incógnitas as N *workloads* dos servidores, e não mais as probabilidades de estado.

3.8.1 Método exato

Neste caso, o sistema linear de equações pode ser resolvido, por exemplo, pelo método de Gauss-Jordan, ou por um método iterativo como Gauss-Seidel ou Gauss-Jacobi.

Os métodos iterativos de solução de sistemas lineares de equações são métodos aproximados por natureza, pois sua implementação exige um critério de parada baseado no erro absoluto máximo tolerável. Além disso, existe a possibilidade de não convergirem. No caso do método de Gauss-Seidel, em particular, sabe-se que a condição suficiente de convergência é de que a matriz dos coeficientes seja *diagonalmente dominante*, ou seja, que o elemento diagonal de uma linha seja, em módulo, maior que a soma dos módulos dos demais elementos da mesma linha. A matriz dos coeficientes do sistema de equações do modelo hipercubo não satisfaz esta condição, de modo que não existe garantia formal de convergência da solução.

Embora as equações de equilíbrio do sistema caracterizem a aplicação de um método iterativo (como pode ser visto em detalhes em CHIYOSHI et al., 2000), no presente trabalho, dado que serão analisadas apenas configurações com $N \leq 13$ servidores, com cauda limitada (conforme pode ser constatado nos Capítulos 5 e 6), será utilizado o método de Gauss-Jordan, que, apesar de ser mais trabalhoso computacionalmente se comparado com o método de Gauss-Seidel, não envolve condições adicionais para a obtenção da solução.

O método de Gauss-Jordan baseia-se na técnica de eliminação de Gauss. Trata-se de um método direto para determinar a inversa de matrizes, gerando soluções exatas para sistemas de equações lineares do tipo $A \cdot x = b$.

No sistema exemplo ($N_A = 3$ e $N = 3$), considerando o hipercubo com cauda limitada, o método pode ser aplicado para as equações (1)-(7), sendo a equação (8) substituída por (12).

Maiores detalhes sobre métodos de solução de sistemas lineares podem ser encontrados em ATKINSON (1978) e CHAPRA & CANALE (1990).

3.8.2 Método aproximado

Segundo CHIYOSHI et al. (2000) são dois os aspectos do desenvolvimento do método que lhe conferem o caráter de ser considerado aproximado. O primeiro é a hipótese de que todos os servidores são homogêneos, e a segunda, é o fato de que no método aproximado, a escolha do servidor a ser despachado para realizar o atendimento se dá de forma aleatória em um sistema M/M/N, enquanto que no hipercubo esta seleção se dá na seqüência estabelecida pela política de despachos do sistema.

Resultados apresentados por LARSON (1975) com o método aproximado mostram desvios da ordem de 1 a 2% em relação aos resultados do método exato, para problemas com servidores homogêneos, comprovando a utilidade do método aproximado. Maiores detalhes sobre o método aproximado de solução para o hipercubo podem ser encontrados em LARSON (1975), LARSON & ODoni (1981) e CHIYOSHI et al. (2000).

No capítulo seguinte será introduzido o estudo de caso deste trabalho. A pesquisa de campo realizada será apresentada em suas diversas fases, identificando o porque da escolha do SAMU-192 como cenário para análise.

4 A PESQUISA DE CAMPO

Este capítulo trata do estudo realizado em alguns sistemas de atendimento médico-emergencial do interior do Estado de São Paulo buscando caracterizar cenários de cidades de pequeno, médio e grande porte. Também será apresentado o método utilizado para a coleta das informações necessárias para a aplicação do modelo hipercubo de filas como ferramenta de avaliação de desempenho para estes sistemas.

Um dos problemas freqüentemente encontrados por planejadores, especialmente do setor público de transportes, é adquirir informações seguras a respeito dos sistemas. Em muitos casos, a falta de uma base de dados consistente dificulta o desenvolvimento de estudos para o melhoramento do nível de serviço que está sendo oferecido. Nestes sistemas, as decisões dependem não somente da escassez de recursos disponíveis, como também do alto nível de interesses políticos envolvidos (LARSON, 1974b, 1983).

A necessidade de informações sobre os diversos aspectos da atividade, assim como sobre os fatores que a condicionam e os impactos que provocam, são requisitos essenciais para o planejamento de sistemas de transportes.

Estudos bem conduzidos, com tomadas de decisão acertadas têm conseqüências diretas sobre o desempenho atual e futuro do sistema (STRAMBI, 1997). Segundo RICHARDSON et al. (1995), para o planejamento de sistemas de transportes, é preciso diagnosticar o problema, identificar suas causas principais, e por fim, compreender o seu comportamento através de técnicas conhecidas.

Todo processo de avaliação de um sistema passa por duas fases de fundamental importância. A primeira, consiste em diagnosticar o problema; uma descrição fiel do fenômeno refletirá diretamente nos resultados da análise. Questões como: o que é importante ser medido, ou quais são as variáveis significativas, a existência de fontes de variabilidade e sua amplitude devem ser respondidas nesta etapa, definida como *planejamento experimental*. Na segunda fase, denominada *análise estatística*, o analista procura fazer inferências sobre os fatores influentes, bem como sugerir uma técnica de modelagem apropriada; ainda dentro desta fase encontra-se a avaliação dos resultados produzidos (BOX et al., 1978).

Os serviços de atendimento médico de urgência, no Brasil, possuem bases de dados com informações precisas a respeito da operação realizada; entretanto, estas informações ainda são muito pouco exploradas, e em alguns casos, a importância de seu conteúdo é desconhecida por parte dos profissionais que gerenciam o sistema.

A pesquisa de campo realizada mostrou características particulares para cenários de cidades de pequeno, médio e grande porte do interior do Estado de São Paulo. As diferenças se dão nos campos operacional, administrativo e de recursos humanos, mas de uma forma geral, é a carência de investimentos financeiros no setor que mais prejudica o desempenho destes sistemas.

4.1 BREVE HISTÓRICO SOBRE O SERVIÇO

A Organização Mundial de Saúde estabelece, como função principal de um serviço de atendimento médico-emergencial, dar suporte básico de vida a toda situação de risco envolvendo pessoas e bens (WORLD HEALTH ORGANIZATION, 2000).

O tempo decorrido entre o instante da ocorrência do chamado emergencial e o instante em que as vítimas começam a receber algum tipo de atendimento, denominado tempo de resposta, é um dos fatores que caracterizam a eficiência do sistema.

O atendimento de urgências é um desafio para todos os países modernos. Quer se trate de urgência traumática ou clínica, apenas por meio de uma organização rigorosa é possível oferecer um atendimento de boa qualidade.

Quase sempre cabe ao serviço público assumir a responsabilidade do atendimento às vítimas, principalmente nos casos muito graves que necessitam, além do meio de transporte, cuidados hospitalares especializados, que não são procedimentos rentáveis para serviços privados. Por este motivo, está claro que o atendimento médico de urgência torna-se a imagem mais visível da saúde pública de um país. Permitir que a população tenha um socorro de alta qualidade é um ponto fundamental que todo governo deve se preocupar.

No Brasil, desde o início dos anos 90, esforços têm sido direcionados no sentido de impulsionar a organização dos atendimentos médicos de urgências. Por se tratar de um serviço público de atendimento (sem fins lucrativos), o sistema brasileiro assemelha-se ao modelo francês, em operação há mais de 20 anos. Graças à ativa cooperação entre algumas secretarias municipais de saúde e o governo francês, as populações de Porto Alegre,

Campinas, Araraquara, Ribeirão Preto, Limeira, Vale do Ribeira, Belo Horizonte, Goiânia, Salvador, Recife e Fortaleza estão recebendo um atendimento de qualidade, podendo ser comparado aos oferecidos internacionalmente, com um melhor aproveitamento dos recursos disponíveis.

Um serviço de atendimento médico de urgência – SAMU deve responder, de forma organizada, a toda situação de urgência que necessite de meios médicos, desde o primeiro contato telefônico, a fim de evitar o uso excessivo de recurso, até a liberação das vítimas ou o seu encaminhamento hospitalar. O sistema deve determinar e desencadear a resposta mais adequada para o caso, assegurar a disponibilidade dos meios hospitalares, determinar o tipo de transporte exigido e preparar o acolhimento dos pacientes.

Operacionalmente falando, SAMU's são sistemas que englobam uma determinada cidade ou até uma região constituída de várias cidades, e contém uma central telefônica de recepção de chamados (no Brasil, o número mais comum é o 192), equipes de socorristas centralizados ou distribuídos pela área de atuação do sistema, veículos (ambulâncias) com equipamentos de suporte básico de vida, e em alguns casos, unidades de reanimação pré-hospitalar (unidades de tratamento intensivo móveis – UTI's móveis) com equipe qualificada. Esta organização está coordenada com uma rede hospitalar municipal ou regional em que os hospitais e as estruturas de atendimento são escolhidos em função de suas especialidades. A esta organização local deve-se somar uma coordenação nacional que, em caso de catástrofe por exemplo, possibilita o auxílio mútuo e um reforço médico especializado rápido.

Um dos componentes absolutamente necessários em um SAMU é a regulação, que é um ato médico constituído por duas fases: diagnóstica e de decisão. De maneira geral, existe um técnico de regulação que identifica o motivo do chamado, estima o nível de gravidade do caso, avalia as circunstâncias e local da ocorrência para acionar meios apropriados de

socorro, e um médico regulador que, através da escuta telefônica simultânea, determina a resposta mais adequada para cada caso. Esta sincronia no sistema só é possível se existir um bom gerenciamento e adequação entre as necessidades dos usuários e as possibilidades dos hospitais.

Neste sentido, as iniciativas pioneiras de Porto Alegre e Campinas são de fundamental importância pois buscam explorar as experiências e potencialidades nacionais para a implantação de uma rede brasileira de atenção às urgências, a fim de atender com qualidade o maior número possível de pessoas.

4.2 CARACTERIZAÇÃO DOS SISTEMAS

O primeiro passo para um bom planejamento de sistema é investigar o problema e identificar as variáveis significativas para a análise.

Para caracterizar o serviço, inicialmente foram consultadas cinquenta cidades do interior paulista sobre a sua disponibilidade de colaborar com o desenvolvimento da pesquisa, expondo os principais objetivos e justificativas para a realização da tese. Depois de várias tentativas, apenas oito se dispuseram a mostrar o serviço em operação e divulgar informações sobre os atendimentos realizados. A Tabela 4.1 apresenta estas cidades, e suas populações, segundo a Contagem da População – 1996 realizada pelo IBGE.

TABELA 4.1 - Cidades participantes da pesquisa e suas populações

Cidade	População (habitantes)
Jaú	103.601
Rio Claro	153.389
Araraquara	168.468
São Carlos	175.517
Limeira	230.348
Piracicaba	302.886
Ribeirão Preto	456.252
Campinas	908.906

4.2.1 Descrição dos oito cenários observados

Para se utilizar o modelo hipercubo de filas como ferramenta de avaliação de desempenho para os serviços de transporte por ambulância, é necessário conhecer características relacionadas aos chamados, às ambulâncias e ao gerenciamento dos sistemas.

Por este motivo, foi elaborado um questionário (apresentado no Anexo B) onde se procurou investigar o processo de chegadas (chamados), o processo de atendimento (ambulâncias) e as características administrativas dos sistemas.

Neste primeiro contato foi possível compreender os serviços que operam em pequenas, médias e grandes cidades, todos dotados de uma central telefônica para onde convergiam os chamados e trabalhando durante as 24 horas do dia. Foram identificados alguns problemas operacionais presentes nos sistemas e que, na maioria das vezes, ocorre devido à falta de experiência e capacitação profissional da equipe de gerenciamento.

Todas as cidades observadas possuíam uma base manuscrita de dados com registros das principais informações referentes à operação do sistema, como por exemplo, instantes de ocorrência de chamados, localização dos chamados, instantes de despacho das ambulâncias, tempos de realização do atendimento junto às vítimas e instantes de retorno à base.

Foi constatado que não havia um bom aproveitamento de todas estas informações de forma a avaliar periodicamente o desempenho dos sistemas. Algumas cidades, entretanto, utilizavam o crescimento do número de chamados ao longo do ano como medida para justificar o pedido de equipamentos e novos veículos para compor a frota.

Em nenhum caso notou-se a existência de um sistema digitalizado de dados. Mesmo assim, foi levantada a questão da sua disponibilidade para

efeitos de análise, com o compromisso de não se revelar informações pessoais envolvidas como nomes e endereços.

A resposta foi positiva em três casos, dentre os quais apenas a coordenação do sistema da cidade de Campinas mostrou interesse em participar da pesquisa, permitindo inclusive o uso das fichas de regulação médica dos atendimentos para qualquer tipo de esclarecimento.

Quanto ao gerenciamento, foi possível concluir que sistemas pequenos são administrados por profissionais com capacitação técnica inadequada para o desempenho da função, o que acaba interferindo no nível de serviço oferecido, pois as decisões nem sempre atendem às reais necessidades do sistema. À medida que cresce o porte da cidade, aumenta a complexidade dos sistemas e decisões operacionais acertadas são, sem dúvida, exigidas. Assim, surgem sistemas organizados onde as decisões partem de uma equipe de profissionais experientes, que buscam atender aos anseios do sistema e da comunidade de usuários como um todo.

4.2.2 Principais diferenças encontradas

Segundo SILVA (1993), cidades com população de até cem mil habitantes são consideradas de pequeno porte; cidades com população entre cem e quinhentos mil habitantes são consideradas de médio porte, e acima de quinhentos mil habitantes, cidades de grande porte.

a) Cenário da cidade pequena amostrada

Por apresentar uma estimativa populacional em torno de cem mil habitantes, o serviço de atendimento médico de urgência da cidade de Jaú representou o cenário de cidade de pequeno porte.

O sistema possuía uma configuração simples, operando com três ambulâncias, das quais apenas uma era dedicada aos atendimentos de emergência, sendo as demais destinadas a remoções internas (entre hospitais do município) e externas. Em consequência desta distribuição, a carga média de trabalho da ambulância dedicada aos atendimentos de urgência apresentava-se alta (em torno de 0,70). A política operacional não admitia a formação de filas de espera, e assim, todos chamados que ocorriam enquanto a ambulância estivesse ocupada eram perdidos, recomendando-se aos solicitantes que buscassem outros meios de atendimento (como por exemplo, o RESGATE do Corpo de Bombeiros). A central telefônica e a base (também chamada garagem) dos veículos localizavam-se nas mesmas instalações da secretaria municipal de saúde, na região central da cidade. No período da pesquisa, os responsáveis avaliavam a viabilização de implantação de um sistema de comunicação via rádio entre veículo e base.

b) Cenário das cidades médias amostradas

Os resultados a seguir foram reunidos com base nas informações obtidas nas cidades de Rio Claro, Araraquara, São Carlos, Limeira e Piracicaba.

Nestes casos, observou-se sistemas com um número médio de cinco ambulâncias destinadas a atendimentos de emergências e sete para o atendimento de remoções, equipadas com um sistema de comunicação via rádio. Em geral, os veículos dedicados às emergências localizavam-se na base central do sistema, junto à central de chamados, enquanto que os demais tinham como bases os pronto-socorros municipais (origens de grande parte dos casos de remoções). Também foi observada uma carga média de trabalho alta, de cerca de 0,60, e possibilidade de formação de fila. A complexidade operacional passa a ser notada. Os sistemas passam a ter um operador de frota, cuja função é despachar os veículos para atender os chamados, segundo o grau

de urgência apresentado, ou seja, determinando critérios de prioridade para os atendimentos.

c) Cenário das cidades grandes amostradas

Segundo as informações disponíveis no IBGE, das oito cidades visitadas, apenas Campinas poderia representar o cenário de grandes cidades. Porém, segundo orientação da secretaria municipal de saúde do município de Ribeirão Preto, a população de possíveis usuários do seu sistema de atendimento médico de urgência ultrapassa os limites apresentados pelo IBGE. O sistema tem atuação “regional”, oferecendo socorro aos pequenos distritos do entorno da cidade (fato também observado em Campinas). Neste sentido, os resultados abaixo são baseados nas informações obtidas em Ribeirão Preto e Campinas.

Operacionalmente, os sistemas das grandes cidades são muito semelhantes aos das cidades médias. Diferenciam-se, basicamente, no número de atendimentos diários realizados, número de ambulâncias e a presença de uma central de regulação médica, assegurando um atendimento adequado às vítimas, evitando a utilização excessiva de recursos.

O número médio de ambulâncias dedicadas às urgências é dez, tendo também outras oito reservadas para as remoções, todas equipadas com sistema de comunicação via rádio. Em todos os casos permite-se a formação de fila, além de se obedecer rigorosos critérios de prioridade para os atendimentos, devido à ação dos reguladores dos chamados.

As Tabelas 4.2 e 4.3 apresentam outros resultados médios fornecidos pelos administradores dos sistemas.

TABELA 4.2 - Resultados médios relativos aos *chamados de emergência*

	Cidades		
	Pequena	Médias	Grandes
número de atendimentos/dia	30	50	80
formação de filas	não	sim	sim
critério de prioridade	não	sim	sim
dias da semana mais solicitados	segunda a sexta	segunda a sexta	segunda a sexta
períodos de maior solicitação	10-12h e 18-20h	10-12h e 17-20h	10-14h e 16-21h
regiões de maior solicitação	periferia / regiões centrais	periferia / regiões industriais e escolares	periferia/regiões industriais e escolares/centros comerciais

TABELA 4.3 - Resultados médios relativos aos *atendimentos de emergência*

	Cidades – resultados médios		
	Pequena	Médias	Grandes
número de veículos	1	5	10
localização	base central	base central	base central
tipos de veículos (*)	VRS	VRS e VSB	VSB e VSA
tempo de resposta (**)	7	10	12
workload média	0,70	0,60	0,60
sistema de comunicação	ausente	presente	presente

(*) VRS – veículo de remoção simples
VSB – veículo de suporte básico (idêntico ao RESGATE)
VSA – veículo de suporte avançado (UTI móvel)

(**) valores em minutos

A terceira parte do questionário é um tanto subjetiva; mesmo assim foi possível reunir as seguintes opiniões:

- um bom serviço de atendimento médico de urgência deve oferecer suporte básico de vida às vítimas, com veículos e equipamentos adaptados para tal, operados por equipes devidamente treinadas para o atendimento às urgências e traumas;
- os parâmetros que melhor indicam o desempenho de um sistema são: tempo de resposta a um chamado, capacitação profissional da equipe, localização estratégica da(s) base(s), a existência de regulação médica, a conscientização da população de usuários, e um eficiente gerenciamento da operação de transporte;

- em uma escala de 0 a 10, foi solicitado que os sistemas se auto-avaliassem, e os resultados causaram surpresas: a administração do sistema que opera na cidade de pequeno porte atribuiu nota 8 justificada pelo fato de que atendem de maneira satisfatória aos seus usuários, mesmo com toda a carência de recursos existente. Já os casos de sistemas de cidades médias atribuíram valores entre 6 e 8, pois acreditam que é necessário realizar melhorias para se chegar a um nível de serviço mais elevado. As cidades grandes, embora com toda a infra-estrutura e organização até então conquistadas, atribuíram nota 7.

Em todos os cenários, notou-se o empenho dos administradores em buscar condições que elevem o nível do atendimento oferecido, justificando a necessidade de investimentos no setor a fim de garantir a disponibilidade de equipamentos e a qualidade técnica da equipe.

4.2.3 Escolha de um cenário para a coleta de dados

Avaliando a caracterização dos oito sistemas investigados e considerando a tendência crescente de uma reestruturação dos serviços que operam em todo o país, o SAMU-192 de Campinas foi eleito como cenário para estudo, uma vez que, além de já estar trabalhando com esta forma operacional organizada, possuía as informações exigidas para iniciar a aplicação da técnica de modelagem. É importante ressaltar o interesse demonstrado pela coordenação do SAMU de Campinas em participar da pesquisa, e o reconhecimento da importância de se construir uma maior interação entre a universidade e a realidade dos problemas sociais na busca de soluções mais eficazes.

4.3 O SAMU-192 DE CAMPINAS

A cidade de Campinas passou por uma reorganização do seu sistema de atendimento médico de urgência em 1994, implantando o SAMU-192 com a função de ser o centro regulador das urgências médicas do município.

Em janeiro de 1998, data da coleta de dados, o sistema operava com 18 ambulâncias, sendo 2 veículos de suporte avançado (VSA), 4 veículos de suporte básico (VSB), 11 veículos de remoção simples (VRS) e 1 veículo psiquiátrico (PSQ). Dentre os 11 VRS's, 7 eram dedicados às operações de remoções simples (pacientes agendados) e 4 operavam como sendo VSB's, totalizando 10 veículos dedicados, de fato, às operações de urgência.

Portanto, para efeitos de análise, foram considerados apenas estas 10 ambulâncias, já que o caso de pacientes agendados e transferências psiquiátricas não atendem às características de um sistema de transporte médico-emergencial.

Todos os veículos básicos e avançados, quando disponíveis, permanecem centralizados na base operacional do sistema, local onde se encontra a central telefônica 192 para onde convergem todos os chamados. O sistema também permite que seus usuários entrem em uma fila de espera (limitada, em média, por 1 usuário para cada veículo em operação), caso todas as ambulâncias estejam ocupadas no instante da ocorrência do evento.

A principal diferença entre os veículos está na equipe técnica e equipamentos de socorro que o acompanham.

Todos os motoristas são também socorristas, isto é, resumem a função de conduzir o veículo e são treinados a desempenhar atendimentos de primeiros socorros, auxiliando a equipe técnica do veículo.

A Figura 4.1 ilustra o tipo de veículo utilizado pelo SAMU de Campinas para os básicos e avançados.

FIGURA 4.1 – Visão geral dos veículos utilizados pelo SAMU-192 de Campinas

Os *veículos de remoção simples* (VRS) são utilizados para o transporte de pacientes com casos clínicos simples, e principalmente para serviços agendados; são dotados de equipamentos mínimos de primeiros socorros, como oxigênio e maca, além do sistema de comunicação via rádio. Nestes casos. No veículo há apenas um auxiliar de enfermagem acompanhando o motorista. Esta classe representa em torno de 35% de todos os chamados.

Os *veículos de suporte básico* (VSB) são muito importantes, pois representam 50% dos atendimentos diários do sistema. Estes veículos possuem equipamentos básicos para o atendimento domiciliar de urgência. Idênticos aos veículos RESGATE do Corpo de Bombeiros, trazem materiais imobilizadores, oxigênio, aspiradores e medicamentos básicos. A equipe é composta por um enfermeiro e um auxiliar que se comunicam com a base, para receberem as ordens de atendimento da equipe de regulação médica. Realizam atendimentos com certo grau de gravidade clínica e traumática, como é o caso de acidentes.

Os *veículos de suporte avançado* (VSA) atendem 10% dos chamados do sistema e, operacionalmente, são mais aprimorados que os veículos básicos no que diz respeito aos equipamentos a bordo, além da equipe acompanhante que sempre é composta por um médico, um enfermeiro e um auxiliar de enfermagem. São veículos do tipo UTI (unidade de tratamento intensivo), realizam atendimentos mais complexos, não só de ordem clínica, mas também certos procedimentos cirúrgicos.

Já o *veículo psiquiátrico* (PSQ) é um caso especial da frota, pois é totalmente adaptado para suportar tratamentos de pacientes com comportamento agressivo, pacientes alcoolizados e drogados. O condutor deste veículo, assim como o enfermeiro que o acompanha são treinados, também de forma específica. Os atendimentos psiquiátricos representam 5% do total de chamados.

A equipe gerenciadora mostrou preocupação quanto aos elevados índices de utilização das ambulâncias em determinados períodos do dia, bem como dos excessivos tempos de espera das vítimas. Então, surgiram discussões sobre uma possível descentralização do serviço, ou seja, quanto uma localização estratégica das ambulâncias poderia interferir no nível de serviço oferecido, além do número ideal de veículos para realizar a operação.

A Figura 4.2 representa a distribuição populacional do município de Campinas sob a área de atuação do SAMU-192. A variação de cores do branco ao preto indica a densidade demográfica em cada região (a cor branca representa áreas com densidade entre 0 e 5 habitantes/hectare, e a cor preta indica áreas com densidade entre 340 e 965,71 habitantes por hectare – valor máximo obtido). Esta distribuição foi obtida segundo informações do Censo Demográfico realizado pelo IBGE, em 1990.

FIGURA 4.2 – Densidade demográfica na área de atuação do SAMU em Campinas

4.3.1 Coleta de dados

Para implementar o modelo hipercubo no SAMU-192 é necessário conhecer as seguintes variáveis:

- taxa média de chegada de chamados no sistema e em cada região particular (átomo geográfico);
- tempo médio de viagem entre os átomos;
- tempo médio de resposta de cada ambulância, e a fração correspondente do tempo total de atendimento (tempo de preparação da equipe, tempo junto ao paciente e tempo de retorno à base); e
- localização das ambulâncias.

Para obter estas informações, foi preciso estudar cuidadosamente o comportamento do sistema em Campinas. Neste sentido, a coleta de dados foi separada em duas etapas: a primeira, com a descrição dos tipos e quantidade

de atendimento realizados pelo SAMU, e a segunda com a seleção de uma amostra significativa para determinar as respectivas estatísticas e validar as hipóteses críticas da modelagem.

4.3.1.1 Primeira etapa: descrição dos atendimentos

O sistema disponibilizou as fichas de regulação médica dos atendimentos realizados desde sua implantação em 1994, para auxiliar no processo de coleta dos dados. Exemplos desta ficha e da planilha de coleta de dados estão apresentados nos Apêndices II e III.

a) Observações gerais dos meses ao longo do ano

A coordenação do sistema informou que notava algumas diferenças no número de solicitações de serviço ao longo do ano, atentando para situações como períodos de férias escolares de verão. Então, para verificar as variações no comportamento dos chamados, foi construída a distribuição do número de atendimentos realizados desde a implantação do SAMU de Campinas. A Figura 4.3 comprova este fato.

FIGURA 4.3 – Atendimentos emergenciais realizados pelo SAMU
(período: junho/1996 – dezembro/1997)

A equipe também informou que os dias úteis (segunda a sexta-feira) são os mais problemáticos para o sistema. Segundo observações, notaram que os sábados, domingos e feriados apresentavam, via de regra, uma redução média de 30% no número de solicitações pelo serviço.

Outra importante observação foi com relação aos períodos do dia em que foi observado o maior índice de chamados no sistema. Períodos entre 10 e 14 horas e em seguida das 16 às 21 horas foram considerados críticos em todos os dias da semana.

Com base neste fato, tomou-se então uma amostra aleatória de um dia de cada mês, neste período, desconsiderando, para o sorteio, os sábados, domingos e feriados, e foram observados todos os chamados emergenciais ocorridos durante as 24 horas dos dias observados. A Figura 4.4 apresenta a distribuição dos atendimentos ao longo das 24 horas de operação para esta amostra.

FIGURA 4.4 – Distribuição do número de atendimentos ao longo do dia (período: junho/1996 - dezembro/1997)

Outras conclusões importantes desta análise são:

- i. dos 35.000 telefonemas que a central 192 recebe, em média, ao mês, apenas 16% correspondem a chamados de urgência, 33% representam trotes, e o restante divide-se em enganos, informações, agendamentos, e ligações administrativas;
- ii. em média, 36% dos chamados de urgência têm origem na via pública, 33% são solicitações dos pronto-atendimentos, centros de saúde e hospitais, 28% são domiciliares, e apenas 3% provém de ambientes de trabalho e escolas;
- iii. do total de atendimentos realizados, 70% são casos clínicos, como atendimento a gestantes e pacientes cardíaco-respiratórios, os outros 30% são casos traumáticos como acidentes, ferimentos por arma de fogo, tentativas de suicídio, etc.

Ainda nesta primeira parte, foi possível obter uma estimativa para alguns indicadores de desempenho para o sistema, como por exemplo: intervalos entre chegadas sucessivas ao sistema, tempo de resposta, tempo junto ao paciente (em cena), etc. Os resultados estão na Tabela 4.4.

TABELA 4.4 – Medidas de desempenho obtidas para o SAMU-192
(amostra relativa ao período junho/1996 a dezembro/1997)

	Média		Desvio-padrão	
	VSB	VSA	VSB	VSA
número de chamados por dia	59	9	10	3
intervalos entre chegadas sucessivas ^(*)	17	110	18	128
tempo de resposta ^(*) ^(**)	12	10	5	3
tempo em cena ^(*)	36	39	26	30
tempo total de atendimento ^(*)	63	62	48	50

^(*) valores em minutos

^(**) tempo de viagem + possível tempo de preparação da equipe (*setup*)

O objetivo desta primeira análise foi verificar se existiam diferenças significativas durante os meses do ano, e períodos característicos ao longo do dia. Conforme observado nas Figuras 4.3 e 4.4, os meses de dezembro, janeiro

e fevereiro não representam relativamente bem o comportamento médio do sistema no restante do ano. Também confirmou-se, graficamente, a hipótese de que os períodos das 10 às 14 horas e das 16 às 21 horas são críticos para a operação, uma vez que apresentam picos bem caracterizados.

Sendo assim, para completar esta primeira etapa de descrição do sistema, é necessário observar se existem diferenças no comportamento do sistema durante os dias da semana e semanas ao longo do mês.

b) observação ao longo do mês

Dentre os meses disponíveis para informações (junho/1996 a dezembro/1997), foram desconsiderados os meses de dezembro de 1996, janeiro, fevereiro e dezembro de 1997, e então foi realizado um sorteio entre os meses restantes a fim de investigar o comportamento do sistema durante os dias da semana, ao longo do mês. O mês sorteado foi novembro de 1997.

O mês de novembro de 1997 contou com vinte dias úteis, e outros dez dias entre sábados, domingos e feriados. Dos vinte dias úteis foi extraída uma amostra aleatória de 10 datas independentes, que forneceu os resultados que seguem. As datas sorteadas e dias da semana respectivos estão apresentados na Tabela 4.5.

TABELA 4.5 – Dias observados durante o mês de novembro de 1997

dia		dia	
03 - segunda-feira		17 - segunda-feira	
05 - quarta-feira		18 - terça-feira	
07 - sexta-feira		19 - quarta-feira	
11 - terça-feira		21 - sexta-feira	
13 - quinta-feira		27 - quinta-feira	

Nesta amostra, foram observados um total de 814 chamados, dos quais 716 foram atendidos por equipes de veículos de suporte básico, e 97 por

equipes de veículos de suporte avançado. A Figura 4.5 ilustra a distribuição dos chamados durante os dez dias de observação, e a Figura 4.6 apresenta a distribuição do total de chamados durante 24 horas para todos os dias de observação.

FIGURA 4.5 – Atendimentos emergenciais realizados pelo SAMU (nov./1997)

FIGURA 4.6 – Distribuição dos atendimentos ao longo de 24 horas (nov./1997)

Do total de chamados, 17,1% tiveram origem na região Norte, 16,5% na região Sul, 17,6% na região Leste, 32,2% na região Oeste e 16,7% na região Central. O número efetivo de atendimentos e a distribuição percentual dos mesmos estão representados nas Figuras 4.7 e 4.8.

FIGURA 4.7 – Número total de chamados nas regiões (nov./1997)

FIGURA 4.8 – Distribuição percentual do total de chamados (nov./1997)

Também foram estimados alguns indicadores de desempenho para esta amostra. Os valores estão apresentados na Tabela 4.6, e podem ser comparados aos da Tabela 4.4, onde estão presentes os valores médios obtidos para a amostra total do período de junho de 1997 a dezembro de 1997.

TABELA 4.6 – Medidas de desempenho para o SAMU (nov./1997)

	Média		Desvio-padrão	
	VSB	VSA	VSB	VSA
número de chamados por dia	70	10	5	3
intervalos entre chegadas sucessivas ^(*)	14	106	15	82
tempo de resposta ^(*) ^(**)	12	10	6	3
tempo em cena ^(*)	37	39	28	30
tempo total de atendimento ^(*)	64	62	48	50

^(*) valores em minutos

^(**) tempo de viagem + possível tempo de preparação da equipe (*setup*)

Verificado o comportamento do sistema para os dados amostrados, concluiu-se que não há diferenças significativas durante os dias e semanas, ao longo do mês (desconsiderando-se sábados, domingos e feriados). Também foi constatada a presença de períodos críticos durante as 24 horas de operação do sistema para todos os dias observados (conforme mostra a Figura 4.6). Então, passou-se à investigação da presença de picos no sistema – períodos em que o nível de serviço oferecido é, em geral, deteriorado.

4.3.1.2 Segunda etapa: identificação de períodos de pico

Assim como todos os sistemas de emergência citados no Capítulo 2, o transporte emergencial de saúde apresenta variações aleatórias ao longo dos períodos de observação, e ainda mais: possui um período de pico bem caracterizado, onde o número de solicitações por atendimento aumenta consideravelmente. Por se tratar de um serviço de emergência, onde a rapidez na realização do atendimento é uma das maiores exigências, é recomendável que os indicadores de desempenho produzidos pela análise que poderão direcionar o tomador de decisões satisfaça tanto as necessidades do período típico (fora do pico) quanto do período atípico do sistema (pico).

Por este motivo, optou-se por determinar possíveis períodos atípicos para o sistema e suas características descritivas como, duração dos intervalos, taxas médias de chegada de chamados e desvios-padrão destas taxas como medida de variabilidade.

A Figura 4.6 mostra a distribuição do número total de chamados em 24 horas. Esta figura é a composição da distribuição dos chamados durante 24 horas, para os dez dias de observação no mês de novembro de 1997. Os gráficos relativos a cada dia estão apresentados no Anexo C.

Nela, é possível visualizar o período entre 9 e 14 horas, onde o número de solicitações de serviço aumenta, e isto também se verifica nos gráficos individuais, salvo pequenas variações. Porém, é preciso realizar uma análise mais consistente para determinar, de fato, a amplitude do intervalo que define o pico.

Então, foi realizado um estudo baseado nas taxas de chegada de chamados dentro de possíveis combinações de períodos entre 8 e 15 horas e 16 e 22 horas, para verificar qual é o intervalo de tempo que, para todos os dias, apresenta taxa média de chegadas significativa, com menor variabilidade.

A Tabela 4.7 mostra os valores das taxas médias de alguns períodos avaliados e suas respectivas variações (a última coluna corresponde ao desvio-padrão respectivo).

TABELA 4.7 – Taxas médias de chegadas para possíveis períodos de pico
(chegadas/minuto)

Período	DIA										média	D-P
	1	2	3	4	5	6	7	8	9	10		
9-14	0,11	0,08	0,06	0,10	0,07	0,08	0,10	0,09	0,08	0,07	0,09	0,02
10-14	0,11	0,08	0,07	0,10	0,08	0,08	0,10	0,09	0,09	0,07	0,09	0,01
9-13	0,13	0,08	0,06	0,11	0,07	0,08	0,10	0,10	0,08	0,08	0,09	0,02
10-13	0,13	0,08	0,07	0,11	0,07	0,07	0,12	0,09	0,09	0,08	0,09	0,02

Os demais intervalos analisados apresentaram taxas menores, não caracterizando, de fato, o período de congestionamento do sistema. O intervalo escolhido, portanto, inicia-se às 10 e termina às 14 horas, pois foi o que apresentou maiores frequências durante os dez dias amostrados, com menor variação em torno do valor médio.

Das 209 observações entre 10 e 14 horas, 190 foram atendidas por equipes de VSB e 19 por equipes de VSA. A Figura 4.9 exibe a distribuição de todos os chamados durante o período de pico.

FIGURA 4.9 – Número de chamados no período das 10 às 14 horas

A Figura 4.10 a seguir apresenta o número de chamados relativos a cada região.

FIGURA 4.10 – Número total de chamados em cada região no período de pico

Para o pico também foram calculados alguns indicadores de desempenho (Tabela 4.8), dentre eles, o intervalo médio entre chegadas sucessivas de chamados, tempo médio de resposta a um chamado, tempo médio em cena e tempo médio de atendimento. Como previsto, por se tratar de

um período de alto índice de congestionamento para o sistema, algumas medidas diferenciam-se dos resultados médios para as 24 horas apresentados nas Tabelas 4.4 e 4.6.

TABELA 4.8 – Medidas de desempenho obtidas para o período de pico

	Média		Desvio-padrão	
	VSB	VSA	VSB	VSA
número de chamados/4 horas	20	2	3	1
Intervalos entre chegadas sucessivas ^(*)	10	92	8	98
tempo de resposta ^{(*)(**)}	13	11	5	3
tempo em cena ^(*)	37	39	28	30
tempo total de atendimento ^(*)	66	63	53	47

^(*) valores em minutos

^(**) tempo de viagem + possível tempo de preparação da equipe (*setup*)

O próximo passo é validar as hipóteses exigidas pela modelagem e então aplicar a técnica do hipercubo para uma análise crítica do desempenho atual do sistema.

5 MODELAGEM DO SAMU-192 E ANÁLISE DOS RESULTADOS

Este capítulo trata do estudo de caso realizado para validar a técnica do hipercubo como ferramenta para modelar serviços de transporte emergencial de saúde em cidades brasileiras. As hipóteses críticas da modelagem são verificadas com dados obtidos em uma grande cidade brasileira. Foram determinados os possíveis estados e equações de balanço do sistema. Os resultados obtidos sobre o desempenho atual do sistema são comparados com valores reais observados e assim determinados os possíveis desvios produzidos com a aplicação da modelagem.

5.1 HIPÓTESES DA MODELAGEM

Para se aplicar o modelo hipercubo, o sistema deve atender às nove hipóteses descritas em 3.2., considerando todas as características do sistema apresentadas no Capítulo 4.

5.1.1 Átomos geográficos

Há diversas maneiras de se representar os átomos geográficos de um sistema: através da divisão política da cidade, ou dos setores policiais, bairros, etc. Em Campinas, o SAMU obedece as regiões correspondentes às áreas de cobertura dos Centros de Saúde: Norte, Sul, Leste e Oeste, sendo também considerada uma região Central no entorno de sua base, como mostra a Figura 5.1.

FIGURA 5.1 - Distribuição espacial do sistema

5.1.1.1 Diferenciando classes de usuários na modelagem

Sabendo-se que o SAMU-192 de Campinas possui classes diferenciadas de usuários, LARSON & ODONI (1981) recomendam que seja utilizado o processo de camadas, conforme exposto em 3.1.

Neste estudo, foram consideradas apenas duas classes emergenciais:

- básica – definida pelos chamados atendidos por VSB's, e
- avançada – definida pelos chamados atendidos por VSA's.

Assim, todas as regiões mostradas na Figura 5.1 foram biparticionadas em básicas (B) e avançadas (A), gerando um total de 10 átomos geográficos

no sistema: Norte B (N_B), Norte A (N_A), Sul B (S_B), Sul A (S_A), Leste B (L_B), Leste A (L_A), Oeste B (O_B), Oeste A (O_A), Centro B (C_B), Centro A (C_A). Esta redistribuição pode ser observada na Figura 5.2.

FIGURA 5.2 – Átomos geográficos biparticionados do SAMU-192

5.1.2 Chegadas Poissonianas independentes

Para verificar esta hipótese, é preciso concluir que as chegadas, em cada átomo do sistema, constituem processos de Poisson independentes.

5.1.2.1 Processo de Poisson

Um processo de Poisson é um processo estocástico $\{A(t) \mid t \geq 0\}$, onde $A(t)$ pode ser interpretado como sendo o número de eventos ocorridos no intervalo $[0, t]$ (por este motivo também ser chamado de processo de contagem). Matematicamente, um processo de Poisson de parâmetro $\lambda > 0$ é um processo de contagem $\{A(t)\}$ com as seguintes propriedades:

a) $\{A(t)\}$ tem incrementos independentes, e para $\lambda > 0$,

$$b) \quad P(A(t) = k) = \frac{(\lambda \cdot t)^k \cdot e^{-\lambda t}}{k!}, \quad k = 0, 1, 2, \dots$$

O número de chegadas no intervalo $(t, t + h)$, ou seja, $A(t + h) - A(t)$ para $h > 0$, é chamado *incremento* no processo $\{A(t)\}$ entre os instantes t e $t+h$. Um processo de chegadas tem *incrementos independentes* se o número de chegadas em intervalos disjuntos são variáveis aleatórias independentes. Em particular, se $\{A(t)\}$ apresenta esta propriedade, $A(t)$ e $[A(t + h) - A(t)]$ são variáveis aleatórias independentes.

Em um processo de Poisson, a probabilidade de ocorrer um evento em um pequeno intervalo de tempo $h \rightarrow 0$ é proporcional a λ , e a probabilidade de ocorrerem dois ou mais eventos neste curto intervalo de tempo é igual a 0.

Com base nesta definição, pode-se facilmente demonstrar que: “Se o número de chegadas em um determinado intervalo de tempo tem distribuição de Poisson de parâmetro $\lambda > 0$, os intervalos entre chegadas sucessivas ao sistema tem distribuição exponencial de parâmetro $1/\lambda$ ”. A recíproca deste resultado também é verdadeira, e a prova completa pode ser encontrada em WOLFF (1989).

5.1.2.2 Identificando os processos de chegadas

As chegadas de chamados em todos os átomos do sistema constituem processos de contagem com incrementos independentes. Para completar a análise, é preciso verificar se o número de chegadas em cada átomo segue padrão Poissoniano.

Para isto, é necessário determinar as taxas médias de solicitação pelo serviço durante o período de observação ou, equivalentemente, os intervalos médios entre chegadas sucessivas para os dez átomos do sistema.

Ao considerar cada átomo individualmente, o número de chamados neste período de quatro horas é relativamente pequeno (a Figura 4.10 ilustra

este fato), dificultando assim a aplicação de ferramentas estatísticas que validem a hipótese de chegadas Poissonianas independentes.

Por outro lado, os chamados ocorrem independentemente átomo a átomo, e o número de chamados de cada átomo representa uma proporção do total do sistema. Assim, são determinadas as probabilidades p_j que reproduzem as chances de ocorrência de um chamado qualquer no sistema ter origem no átomo j .

A Tabela 5.1 descreve o número de chamados em cada átomo e a proporção com relação ao total observado no sistema.

TABELA 5.1 – Proporção de chamados em cada átomo do sistema

Átomo	Nº de chamados	Proporção - p_j
1 - N_B	33	0,1579
2 - N_A	4	0,0191
3 - S_B	33	0,1579
4 - S_A	3	0,0144
5 - L_B	33	0,1579
6 - L_A	3	0,0144
7 - O_B	58	0,2750
8 - O_A	5	0,0239
9 - C_B	32	0,1531
10 - C_A	5	0,0239
Total	209	1,0000

Com tais características no sistema, a hipótese poderá ser facilmente testada utilizando a propriedade de decomposição de processo de Poisson (WOLFF, 1989): “Seja $Y \sim P_o(\lambda)$ um processo de Poisson de parâmetro $\lambda > 0$ particionado aleatoriamente em k classes, onde p_j ($j = 1, 2, \dots, k$) é a probabilidade de ocorrência da classe j , e Y_j é a frequência da classe j , com $\sum p_j = 1$ e $\sum Y_j = Y$. Então, as variáveis aleatórias $Y_1, Y_2, \dots, Y_k$ são mutuamente independentes, com $Y_j \sim P_o(\lambda \cdot p_j)$, $j = 1, 2, \dots, k$.”

Este resultado pode ser melhor compreendido no esquema abaixo:

FIGURA 5.3 – Decomposição de um Processo de Poisson de parâmetro $\lambda > 0$

Desta forma, basta testar a hipótese de que o número de chegadas ao sistema por intervalo de tempo constitui um Processo de Poisson, com parâmetro $\lambda > 0$, ou equivalentemente, que os intervalos médios entre chegadas sucessivas ao sistema têm distribuição exponencial, com parâmetro $1/\lambda$.

5.1.2.3 Cálculo de $\mathbf{I}$

Para determinar a taxa média de chegadas ao sistema, foram considerados os intervalos médios entre chegadas sucessivas para todos os dias de observação, no período das 10 às 14 horas, considerando todos os átomos do sistema. A Tabela 5.2 apresenta o tempo médio, desvio-padrão e coeficiente de variação para cada dia.

Nota-se que os desvios-padrão são, em geral, da ordem de grandeza das médias, ou equivalentemente, que os coeficientes de variação estão próximos de 1. Isto sugere que a variável que define o intervalo de tempo entre chegadas sucessivas ao sistema é exponencialmente distribuída.

TABELA 5.2 – Intervalos médios entre chegadas sucessivas ao sistema

Dia	Intervalo médio (min)	Desvio-padrão	Coefficiente de variação
1	9	7	1,2857
2	11	12	0,9167
3	14	12	1,1667
4	9	9	1,0000
5	11	17	0,6471
6	13	15	0,8667
7	9	7	1,2857
8	10	14	0,7143
9	10	9	1,1111
10	15	19	0,7895
Média	11,10 minutos (0,1850 hora)	12,10	0,9783
$\lambda = 1/0,1850 = 5,4054$ chamados/hora			

5.1.2.4 Teste de aderência

Para verificar estatisticamente a forma da distribuição populacional dos intervalos entre chegadas sucessivas ao sistema, foi aplicado o *Método de Kolmogorov-Smirnov (K-S)*, ao nível de significância $\alpha = 0,05$.

A Figura 5.4 mostra a distribuição dos dados amostrais e os valores esperados obtidos através da aplicação do modelo exponencial para os dez dias observados em novembro de 1997, no período das 10 às 14 horas.

FIGURA 5.4 - Distribuição dos intervalos entre chegadas sucessivas ao sistema.

A diferença máxima entre as frequências relativas acumuladas observadas e produzidas pela distribuição exponencial foi, em módulo, de 0,0902 e o valor crítico tabelado para o Método *K-S*, ao nível de significância $\alpha = 5\%$, considerando os 209 valores da amostra é de 0,0941 $(1,36/\sqrt{209})$. Logo, não se pode rejeitar a hipótese de que os intervalos entre chegadas sucessivas ao sistema têm distribuição exponencial, e portanto, o número de chegadas por intervalo de tempo tem distribuição de Poisson. A tabela com a aplicação do método está apresentada no Anexo D. Isto valida a hipótese 2 do modelo hipercubo (seção 3.2).

5.1.2.5 Cálculo dos $\mathbf{l}_j$

Com este resultado, e considerando as devidas proporções que representam as chegadas em cada átomo do sistema apresentadas na Tabela 5.1, admite-se que os chamados chegam em cada átomo segundo Processos de Poisson independentes, com taxas médias $\lambda_j = \lambda \cdot p_j$ ($j = 1, 2, \dots, 10$).

TABELA 5.3 – Taxas médias de chegadas para cada átomo

Átomo	Nº de chamados	p_j	$\mathbf{l}_j$ (chamados/hora)
1 - N_B	33	0,1579	0,8535
2 - N_A	4	0,0191	0,1035
3 - S_B	33	0,1579	0,8535
4 - S_A	3	0,0144	0,0776
5 - L_B	33	0,1579	0,8535
6 - L_A	3	0,0144	0,0776
7 - O_B	58	0,2750	1,5001
8 - O_A	5	0,0239	0,1293
9 - C_B	32	0,1531	0,8276
10 - C_A	5	0,0239	0,1293
Total	209	1,0000	5,4054

5.1.3 Tempos de viagem

A matriz dos tempos médios de viagem foi fornecida pelo próprio sistema, admitidas as dez regiões geradoras de chamados.

TABELA 5.4 – Tempos médios de viagem (minutos)

Átomo		N _B	N _A	S _B	S _A	L _B	L _A	O _B	O _A	C _B	C _A	Média
1 -	N _B	6	6	28	28	16	16	24	24	16	16	18
2 -	N _A	4	4	24	24	12	12	20	20	12	12	14
3 -	S _B	28	28	6	6	24	24	16	16	16	16	18
4 -	S _A	24	24	4	4	20	20	12	12	12	12	14
5 -	L _B	16	16	24	24	6	6	24	24	16	16	17
6 -	L _A	12	12	20	20	4	4	20	20	12	12	14
7 -	O _B	24	24	16	16	24	24	6	6	16	16	17
8 -	O _A	20	20	12	12	20	20	4	4	12	12	14
9 -	C _B	16	16	16	16	16	16	16	16	5	5	14
10 -	C _A	12	12	12	12	12	12	12	12	3	3	10
Média do sistema = 15												

Os veículos de suporte avançado, geralmente desempenham tempos médios de viagem menores que os básicos, devido ao grau de urgência envolvido nos casos atendidos. Daí explicar-se a diferença entre os tempos médios de viagem com origens nos átomos Norte A, Sul A, Leste A, Oeste A e Central A para os demais átomos do sistema.

Os tempos de viagem dentro do próprio átomo são distintos de zero, pois é preciso representar as reais distâncias percorridas por uma ambulância e os tempos necessários para tal, dentro de uma determinada área.

5.1.4 Servidores

No cenário considerado, o SAMU trabalha com uma frota de 10 veículos distintos: 8 VSB's e 2 VSA's.

Todas as ambulâncias permanecem centralizados na base, quando disponíveis, podendo deslocar-se para qualquer um dos átomos para realizar um atendimento (conforme hipótese 4 do modelo hipercubo, seção 3.2)

5.1.5 Localização dos servidores

A matriz de localização (L) dos servidores é facilmente determinada para a configuração original do sistema, pois todos permanecem na base central quando disponíveis (conforme hipótese 5 do modelo hipercubo, seção 3.2).

Para generalizar a nomenclatura, os veículos de suporte avançado passam a ser chamados de veículos 1 e 2, e os veículos de suporte básico de 3, 4, 5, 6, 7, 8, 9 e 10.

$$L = \begin{array}{c} \begin{array}{cccccccccc} \mathbf{1} & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} & \mathbf{6} & \mathbf{7} & \mathbf{8} & \mathbf{9} & \mathbf{10} \\ \mathbf{N_B} & \mathbf{N_A} & \mathbf{S_B} & \mathbf{S_A} & \mathbf{L_B} & \mathbf{L_A} & \mathbf{O_B} & \mathbf{O_A} & \mathbf{C_B} & \mathbf{C_A} \end{array} \\ \left[\begin{array}{cccccccccc|c} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \mathbf{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathbf{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathbf{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathbf{5} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathbf{6} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathbf{7} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathbf{8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathbf{9} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathbf{10} \end{array} \right] \end{array}$$

5.1.6 Despacho dos servidores

O SAMU admite a política de enviar apenas uma equipe para atender a um chamado emergencial, o que está de acordo com a hipótese 6 do modelo hipercubo (seção 3.2). Casos de grandes acidentes ou catástrofes envolvem outros sistemas de apoio como a defesa civil, corpo de bombeiros ou polícia.

Ainda dentro da política de atendimento do sistema, é permitida a formação de fila de espera, onde usuários que telefonam solicitando o atendimento enquanto todas as ambulâncias estão ocupadas podem aguardar

até que um veículo desocupe para atendê-los. Esta fila é limitada em dez, o que corresponde a um usuário por ambulância.

5.1.7 Política de despacho dos servidores

De acordo com a distribuição espacial da cidade de Campinas através do processo de camadas mostrado na Figura 5.2 e, considerando que na configuração original do sistema todos os veículos permanecem centralizados na base quando disponíveis, a lista de preferência de despacho tem, nas linhas correspondentes aos átomos geradores de chamados da classe avançada, os veículos 1 e 2 como servidores primários. As demais posições de *backup* são assumidas pelos oito veículos básicos, aleatoriamente. Para os casos de átomos geradores de chamados básicos, os oito veículos assumem as oito primeiras posições da lista, aleatoriamente, restando os veículos 1 e 2 como sendo as últimas opções de envio.

Dos procedimentos citados na descrição da hipótese 7, seção 3.2, optou-se, no presente trabalho, pelo método de geração de matrizes de despachos aleatórias (BURWELL et al., 1993), segundo a política operacional do sistema (por exemplo, no caso original, todas as ambulâncias permanecem posicionadas na base central). Foram realizadas experiências computacionais gerando até vinte e cinco matrizes aleatórias, e observadas as variações sobre as medidas de desempenho produzidas pelo modelo.

Verificou-se que os resultados produzidos resolvendo-se o modelo acima dez vezes para a mesma configuração operacional não apresentaram diferenças significativas nas probabilidades de equilíbrio, e em consequência, nas medidas de desempenho produzidas, sendo o erro absoluto da ordem de 10^{-6} . Em termos práticos, os reflexos desta diferença não provocam interferências sobre as decisões a serem tomadas no sistema, e portanto,

decidiu-se por resolver o modelo computacionalmente dez vezes, correspondentes a dez matrizes aleatórias, para cada cenário investigado.

As Tabela 5.5 e 5.6 apresentam exemplos de lista de preferência de despachos para o SAMU (geradas aleatoriamente) para a configuração original do sistema.

TABELA 5.5 –Exemplo 1: Matriz de preferência de despachos (cenário original)

Átomo	Preferência de despacho									
	1 ^a	2 ^a	3 ^a	4 ^a	5 ^a	6 ^a	7 ^a	8 ^a	9 ^a	10 ^a
1 - N_B	9	10	3	4	5	6	7	8	2	1
2 - N_A	1	2	10	7	3	4	9	5	8	6
3 - S_B	10	3	4	5	6	7	8	9	1	2
4 - S_A	2	1	5	3	10	9	4	6	7	8
5 - L_B	3	4	7	10	8	5	6	9	2	1
6 - L_A	1	2	6	9	4	10	5	3	8	7
7 - O_B	4	5	8	6	7	3	9	10	2	1
8 - O_A	2	1	9	8	3	5	10	4	6	7
9 - C_B	5	6	8	10	9	4	3	7	1	2
10 - C_A	2	1	10	5	6	8	7	9	3	4

veículos

TABELA 5.6 – Exemplo2: Matriz de preferência de despachos (cenário original)

Átomo	Preferência de despacho									
	1 ^a	2 ^a	3 ^a	4 ^a	5 ^a	6 ^a	7 ^a	8 ^a	9 ^a	10 ^a
1 - N_B	4	3	7	8	9	10	6	5	1	2
2 - N_A	1	2	3	8	4	5	6	7	9	10
3 - S_B	3	7	8	9	10	6	5	4	1	2
4 - S_A	2	1	6	4	3	10	5	7	8	9
5 - L_B	7	8	6	3	5	9	10	4	2	1
6 - L_A	2	1	10	4	8	3	9	7	6	5
7 - O_B	5	6	9	7	8	3	9	7	6	5
8 - O_A	2	1	4	5	7	9	3	8	10	6
9 - C_B	5	6	8	10	9	4	3	7	1	2
10 - C_A	2	1	3	6	7	9	8	10	4	5

5.1.8 Tempos de atendimento

Os tempos de atendimento para cada ambulância podem ser determinados como sendo o intervalo de tempo entre os instantes de saída e retorno à base.

Com as informações contidas na ficha de regulação médica, estes valores foram calculados como sendo a soma do tempos de preparo da equipe, viagem até o local da ocorrência (ida) , em cena e viagem de retorno à base (volta).

a) tempo de preparação da equipe

O valor do tempo de preparação tanto para equipes de VSA's quanto para equipes de VSB's são pequenos, da ordem de 30 segundos.

b) tempo de viagem de ida

No período em análise, o tempo médio de viagem para os VSA's foi de 11 minutos, enquanto que para os VSB's, este valor foi de 13 minutos.

c) tempo em cena

Na Tabela 4.8 estão os tempos que cada equipe passa junto às vítimas para realizar o atendimento. No caso dos VSA's, o tempo médio observado foi de 39 minutos, e para os VSB's, 37 minutos. Esta pequena diferença se dá principalmente devido ao tipo de atendimento realizado por cada equipe.

d) tempo de viagem de volta

Os tempos médios de viagem de retorno à base, são, em geral diferentes do tempo de viagem de ida pois, em sua grande maioria, este tempo é medido a partir do local para onde as vítimas tenham sido removidas (hospital, pronto-

socorro ou mesmo domicílio) e não a partir do local do incidente, e também porque o veículo não retorna com os equipamentos de alerta acionados.

Para o conjunto das 209 observações, este tempo foi em média de 13 minutos para VSA's e de 16 minutos para VSB's.

Portanto, o tempo médio de atendimento para um chamado de urgência é de 63 minutos para chamados atendidos por VSA's e de 66 minutos para chamados atendidos por VSB's.

A Tabela 5.7 mostra os tempos médios de atendimento de cada ambulância, seus desvios-padrão, coeficientes de variação e as taxas médias correspondentes.

TABELA 5.7 – Taxas médias de atendimento para cada ambulância

Ambulância	Tempo médio de atendimento (*)	Desvio padrão (*)	Coeficiente de variação	χ^2 (**)	Tempo médio de viagem (*)	Desvio padrão (*)
1	61	40	0,7	0,9836	10	2
2	64	44	0,7	0,9375	10	2
3	63	52	0,8	0,9524	12	3
4	70	48	0,7	0,8571	13	3
5	61	44	0,7	0,9836	12	2
6	66	58	0,9	0,9091	12	4
7	64	50	0,8	0,9375	13	3
8	61	40	0,7	0,9836	12	3
9	70	58	0,8	0,8571	13	2
10	70	56	0,8	0,8571	13	4
VSA	63	42	0,7	0,9606	10	2
VSB	66	51	0,8	0,9172	13	3
Total	65	49	0,8	0,9259	12	3

(*) minutos

(**) chamados/hora

Observa-se que os desvios-padrão dos tempos médios de atendimento são da ordem de grandeza das médias (isto é, os coeficientes de variação são razoavelmente próximos de 1). Estes valores são, portanto, indicações de que os tempos de atendimento têm padrão exponencial.

De fato, verificou-se, através do *Método de Kolmogorov-Smirnov*, que não se pode rejeitar a hipótese de que os tempos de atendimento de cada

ambulância têm distribuição exponencial, com nível de significância $\alpha = 0,05$. As tabelas com os resultados dos teste para cada ambulância estão apresentados no Anexo D. Isto valida a hipótese 8 do modelo hipercubo (seção 3.2).

Com base na Tabela 5.7 foram testadas hipóteses de igualdade entre os tempos médios de atendimento de VSA's e VSB's. Para isto foi utilizada a técnica de análise de variância (ANOVA), a fim de identificar diferenças entre as médias populacionais devidas às possíveis causas atuando simultaneamente sobre os elementos da população.

O teste verificou, ao nível de significância $\alpha = 0,05$ diferenças entre os tempos médios de atendimento para os dez veículos. O quadro de ANOVA está exposto no Anexo D. Para efeito de modelagem, optou-se então por considerar servidores distintos no sistema, devido à variabilidade observada nos tempos de atendimento de cada ambulância, em particular. Assim, o desenvolvimento do modelo e os resultados que a seguir serão obtidos são baseados na modelagem de sistemas com taxas médias de atendimento (μ_n) distintas, conforme apresentada no Capítulo 3.

5.1.9 Relação entre o tempo de atendimento e o tempo de viagem

Pode-se observar na Tabela 5.7 que os tempos médios de viagem são pequenos em relação aos tempos médios de atendimento, para cada ambulância (em média, cerca de 18%), e que variações nos tempos médios de viagem contribuem pouco para as variações nos tempos médios de atendimento (basta comparar os valores dos desvios-padrão apresentados na Tabela 5.7). Isto valida a hipótese 9 do modelo hipercubo (seção 3.2).

Validadas as nove hipóteses mais críticas do modelo, conforme seção 3.2, o próximo passo é determinar os possíveis estados para então encontrar o conjunto de equações de equilíbrio e, conseqüentemente, os valores das probabilidades de equilíbrio para o sistema.

5.2 POSSÍVEIS ESTADOS DO SISTEMA

O SAMU-192 de Campinas possui 10 átomos e 10 ambulâncias. Portanto, o modelo será aplicado para uma situação em que $N_A = 10$ e $N = 10$.

O serviço em operação na cidade de Campinas permite que seus usuários entrem em uma fila de espera (limitada, em média por 1 usuário por veículo em operação, ou seja, fila máxima igual a 10), caso todas as ambulâncias estejam ocupadas no instante da ocorrência do evento.

Cada estado particular do sistema será representado por um vetor com $N=10$ elementos, onde cada um pode assumir valores 0 ou 1, representando o estado de cada ambulância (0 se estiver disponível e 1, caso contrário). Logo, existem 2^{10} estados possíveis no sistema, além dos 10 estados, $S_{11}, S_{12}, \dots, S_{20}$, que definem a fila do sistema.

É importante observar que neste modelo não há transições de passos maiores que 1. Na prática, não se despacha mais que uma equipe para atender a um chamado e apenas uma é liberada após a conclusão de um atendimento.

As Figuras 5.5 e 5.6 mostram o diagrama de fluxos para os estados (0000000000) – sistema vazio e (1111111111) – sistema totalmente ocupado com as respectivas taxas de transição, considerando a matriz de preferências de despachos apresentada na Tabela 5.5.

Sistema vazio

FIGURA 5.5 – Diagrama de fluxos para o estado (0000000000)

Sistema cheio

FIGURA 5.6 – Diagrama de fluxos para o estado (1111111111)

Exemplificando um estado intermediário do sistema, a Figura 5.7 apresenta o diagrama de fluxos para o estado (1000000101) que representa o sistema com a ambulância 1 (VSA) ocupada, ambulância 2 (VSA) livre, ambulância 3 (VSB) ocupada, ambulâncias 4, 5, 6, 7, 8 e 9 (VSB's) livres e ambulância 10 (VSB) ocupada. Neste caso, as possíveis transições se dão com os estados (1000000100), (1000000001), (0000000101), (1000000111), (1000001101), (1000010101), (1000100101), (1001000101), (1010000101) e (1100000101).

FIGURA 5.7 - Diagrama de fluxos para o estado (1000000101)

Para ilustrar os efeitos da política de despachos aleatória, a Figura 5.8 mostra o diagrama de fluxos para o estado (1000000101) considerando a Tabela 5.6 como nova matriz de preferências de despachos gerada aleatoriamente para o cenário original.

FIGURA 5.8 – Aleatoriedade da política de despachos: estado (1000000101)

Raciocínio análogo deve ser repetido para os 1024 possíveis estados do sistema.

5.3 EQUAÇÕES DE BALANÇO

Para se descrever as equações de balanço dos estados, é preciso admitir a condição de equilíbrio para o sistema, ou seja, para todos os possíveis estados, o fluxo com que o sistema entra em tal estado deve ser igual ao fluxo com que ele sai deste estado.

É indispensável também considerar a matriz de preferência de despacho dos servidores, uma vez que os veículos são despachados na ordem de preferência que aparecem nesta lista, pois as taxas de transição dependem da ordem em que os veículos são despachados para os átomos.

O sistema de equações produzido para este estudo de caso resulta em $2^{10} = 1024$ equações lineares cujas incógnitas são as probabilidades de equilíbrio do sistema, conforme discutido na seção 3.5.1.2.

Devido à insuficiência destas equações, é necessário introduzir a equação de normalização das probabilidades, como apresentado na seção 3.5, com o objetivo de explicar a forma como a massa total de probabilidades se distribui entre todos os estados do sistema.

Abaixo seguem as equações de balanço dos estados (0000000000), (1111111111) e (1000000101) representados, graficamente, nas Figuras 5.5, 5.6 e 5.7. Analogamente pode-se escrever as demais equações de balanço para o sistema.

- Estado 0000000000

$$\begin{aligned}
 &[(\lambda_2 + \lambda_6) + (\lambda_4 + \lambda_8 + \lambda_{10}) + (\lambda_5) + (\lambda_7) + (\lambda_9) + (\lambda_1) + (\lambda_3)] \cdot p_{0000000000} = \\
 &\mu_1 \cdot p_{0000000001} + \mu_2 \cdot p_{0000000010} + \mu_3 \cdot p_{0000000100} + \mu_4 \cdot p_{0000001000} + \mu_5 \cdot p_{0000010000} + \\
 &\mu_6 \cdot p_{0000100000} + \mu_7 \cdot p_{0001000000} + \mu_8 \cdot p_{0010000000} + \mu_9 \cdot p_{0100000000} + \mu_{10} \cdot p_{1000000000}
 \end{aligned}$$

- Estado 1111111111

$$[\mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5 + \mu_6 + \mu_7 + \mu_8 + \mu_9 + \mu_{10}]p_{1111111111} = \lambda p_{1111111110} + \lambda p_{1111111101} + \lambda p_{1111111011} + \lambda p_{1111110111} + \lambda p_{1111011111} + \lambda p_{1110111111} + \lambda p_{1101111111} + \lambda p_{1011111111} + \lambda p_{0111111111}$$

$$\text{onde } \lambda = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10}$$

- Estado 1000000101

$$[(\lambda_2 + \lambda_4 + \lambda_6 + \lambda_8 + \lambda_{10}) + (\lambda_7 + \lambda_3 + \lambda_5) + (\lambda_9) + (\lambda_1) + (\mu_1) + (\mu_3) + (\mu_{10})]p_{1000000101} = \mu_2 p_{1000000111} + \mu_4 p_{1000001101} + \mu_5 p_{1000010101} + \mu_6 p_{1000100101} + \mu_7 p_{1001000101} + \mu_8 p_{1010000101} + \mu_9 p_{1100000101} + (\lambda_2 + \lambda_6) p_{1000000100} + (\lambda_3 + \lambda_5) p_{1000000001} + \lambda_3 p_{0000000101}$$

Para o caso do estado (1000000101) gerado a partir da matriz representada na Tabela 5.6, a equação de balanço escreve-se como:

$$[(\lambda_2 + \lambda_4 + \lambda_6 + \lambda_8 + \lambda_{10}) + (\lambda_1) + (\lambda_7 + \lambda_9) + (\lambda_5 + \lambda_3) + (\mu_1) + (\mu_3) + (\mu_{10})]p_{1000000101} = \mu_2 p_{1000000111} + \mu_4 p_{1000001101} + \mu_5 p_{1000010101} + \mu_6 p_{1000100101} + \mu_7 p_{1001000101} + \mu_8 p_{1010000101} + \mu_9 p_{1100000101} + (\lambda_2) p_{1000000100} + (\lambda_3) p_{1000000001}$$

Devido a quantidade de equações de balanço para cada cenário investigado (1024 equações), no Apêndice IV somente estão listadas as 100 primeiras equações de balanço geradas para a política de despachos da Tabela 5.5, a fim de ilustrar o sistema de equações de equilíbrio para este estudo de caso, em cada cenário investigado.

Com base na Figura 5.6, pode-se escrever as equações dos estados que representam a fila no sistema:

$$\begin{cases} \ddot{e} \cdot p_{1111111111} = \dot{\lambda} \cdot p_{11} \\ \ddot{e} \cdot p_{11} = \dot{\lambda} \cdot p_{12} \\ \ddot{e} \cdot p_{12} = \dot{\lambda} \cdot p_{13} \\ \vdots \\ \ddot{e} \cdot p_{19} = \dot{\lambda} \cdot p_{20} \end{cases}$$

onde $p_{11} = P(S_{11})$, $p_{12} = P(S_{12})$, ..., $p_{20} = P(S_{20})$, $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_{10}$ e $\mu = \mu_1 + \mu_2 + \dots + \mu_{10}$.

E portanto, a equação de normalização das probabilidades passa a ser escrita como:

$$p_{0000000000} + p_{0000000001} + \dots + \sum_{n=0}^{10} \rho^n \cdot p_{111} = 1$$

onde $\rho = \lambda/\mu$.

5.4 RESULTADOS

O modelo foi implementado computacionalmente em linguagem Pascal, e executado em uma estação de trabalho IBM 3CT, com 128 MB de memória e sistema operacional IBM AIX.

Os resultados obtidos através da solução do sistema de equações de balanço para os estados mostraram a eficiência do modelo hipercubo para avaliar o serviço de transporte médico-emergencial de Campinas, apresentando desvios pouco significativos com relação aos dados reais para as várias medidas de desempenho calculadas.

Para esta configuração, a probabilidade de se encontrar o sistema vazio ($p_{0000000000}$), ou seja, com todas as ambulâncias disponíveis é muito pequena, em torno de 0,0028, e a probabilidade do sistema estar saturado (p_s) é de 0,0877, isto é, cerca de 9% dos chamados encontram o sistema saturado, ou

seja, com todos os servidores ocupados e com 10 usuários aguardando por atendimento. A probabilidade de perda (p_{s20}) foi calculada como sendo 0,0001, ou seja, 0,01% dos chamados são redirecionados para outros serviços pois encontram o sistema com todas as ambulâncias ocupadas e dez usuários aguardando para receber o atendimento.

5.4.1 Workloads

Uma comparação entre as *workloads* das ambulâncias está apresentada na Figura 5.9.

FIGURA 5.9 – *Workloads* das ambulâncias para o cenário original

Esta figura confirma as observações iniciais feitas pelos gerentes do sistema de que os fatores de utilização de cada ambulância são relativamente altos, em se tratando de um serviço emergencial. Uma das vantagens deste cenário, onde todos os veículos encontram-se centralizados, é o fato de existir um bom balanceamento das *workloads* dentro de cada classe (VSB's passam, em média, 63% do tempo ocupados; já VSA's passam apenas 39% do tempo realizando atendimentos). Quanto ao sistema, o valor obtido foi de 0,58, compatível com o valor fornecido pela coordenação do sistema, no início da pesquisa de campo (em torno de 0,6). Os valores das *workloads* estão expressos na Tabela 5.8.

5.4.2 Tempos médios de resposta

Considerando que o tempo de preparação da equipe é praticamente desprezível (1/2 minuto), os tempos médios de resposta serão obtidos diretamente por meio das expressões dos tempos médios de viagem (expressões 17, 18 e 19).

5.4.2.1 Ambulância

Um indicador de desempenho importante de ser medido no sistema é o tempo médio de resposta de cada ambulância, ou seja, quanto tempo, em média, cada equipe gasta para chegar ao local do chamado e iniciar o atendimento junto às vítimas, independentemente da região que solicitou o serviço. A Tabela 5.8 ilustra as *workloads* e os tempos médios de resposta de cada ambulância.

TABELA 5.8 – Tempos médios de resposta das ambulâncias (cenário original)

Ambulância	Workload	Tempo médio de resposta (minutos)		Desvio	
		Modelo	Amostra	minutos	%
1	0,39	10,88	10,38	0,50	4,82
2	0,39	10,90	10,70	0,20	1,87
3	0,63	14,22	13,39	0,83	6,20
4	0,64	14,24	13,54	0,70	5,17
5	0,62	14,23	14,30	-0,07	-0,49
6	0,64	14,24	13,64	0,60	4,40
7	0,63	14,23	13,17	1,06	8,05
8	0,62	14,22	13,27	0,95	7,16
9	0,64	14,24	13,23	1,01	7,63
10	0,64	14,24	13,76	0,48	3,49
Média VSA	0,39	10,89	10,54	0,35	3,32
Média VSB	0,63	14,23	13,54	0,70	5,13

Os valores apontam para respostas mais rápidas quando o chamado é classificado como avançado. Isto se explica devido ao grau de urgência envolvido. Os desvios com relação aos tempos de resposta observados no sistema são da ordem de 3,32% para os VSA's e de 5,13% para os VSB's,

com um desvio médio de 4,84% se consideradas todas as ambulâncias. Note que estes desvios são relativamente pequenos, o que valida a utilização do modelo hipercubo para analisar o presente sistema.

5.4.2.2 Átomo

Ainda com relação aos tempos médios de resposta, uma outra maneira de olhar para o sistema é com relação ao tempo médio de resposta para cada átomo, ou seja, em função da região onde estão as vítimas, isto é, em média, quanto tempo uma equipe qualquer do sistema gasta para chegar e iniciar os procedimentos médicos em cada região. Estes valores são obtidos através da aplicação da expressão 3.15, e estão apresentados na Tabela 5.9.

TABELA 5.9 – Tempos médios de resposta para cada átomo (cenário original)

Átomo	Tempo médio de resposta (minutos)		Desvio	
	Modelo	Amostra	minutos	%
1 - N _B	16,01	15,45	0,56	3,62
2 - N _A	13,09	13,60	-0,51	-3,75
3 - S _B	15,93	14,52	1,41	9,71
4 - S _A	13,00	11,67	1,33	11,40
5 - L _B	15,96	14,78	1,18	7,98
6 - L _A	13,03	12,50	0,53	4,24
7 - O _B	15,76	14,17	1,59	11,22
8 - O _A	12,83	14,00	-1,17	-8,36
9 - C _B	5,50	6,63	-1,13	-17,04
10 - C _A	4,33	4,50	-0,17	-3,78
Média A	11,25	11,24	0,01	0,08
Média B	13,83	13,11	0,72	5,49

Desconsiderando-se a região central, onde se localiza a base das ambulâncias, observa-se que as regiões geradoras de chamados básicos têm tempos médios de resposta muito semelhantes; o mesmo acontece para as regiões geradoras de chamados avançados. Isto se deve ao fato da localização da base do SAMU, na região Central, na prática, ser equidistante de todos os demais átomos do sistema. Os desvios relativos dos tempos médios de resposta foram pouco significativos, em termos práticos, da ordem de 0,08%

no caso dos átomos da classe avançada e 5,49% para a classe básica, e um desvio médio para o sistema de 2,97% (os desvios absolutos foram da ordem de 6,59% para os átomos da classe avançada, 8,95% para os átomos da classe básica e 7,89% para o sistema).

As Figuras 5.10 e 5.11 mostram a distribuição dos tempos de resposta para as ambulâncias e para os átomos, respectivamente. Novamente é possível visualizar os efeitos de uma configuração centralizada para o sistema.

FIGURA 5.10 – Tempos médios de resposta para cada átomo (cenário original)

FIGURA 5.11 – Tempos médios de resposta das ambulância (cenário original)

5.4.2.3 Sistema

Obtido através da expressão 17, o tempo médio de resposta, independente do átomo gerador do chamado e da equipe despachada para realizar o atendimento, reflete uma medida global de desempenho para o sistema. Portanto, ao ocorrer um chamado qualquer no sistema, sabe-se, em média, em quanto tempo as vítimas começarão a receber algum tipo de intervenção clínico-cirúrgica.

Para o SAMU, este tempo foi calculado em 13,82 minutos, e o valor observado na amostra foi de 14,47 minutos (diferença de 0,65 minutos). O desvio com relação ao valor amostrado é de 4,49%, o que é aceitável para validar o modelo hipercubo nesta aplicação, considerando, por exemplo, possíveis imprecisões nos dados de entrada.

Os resultados apresentados mostram que o modelo está razoavelmente calibrado, pois produziu desvios pouco significativos com relação aos dados reais, validando a hipótese inicial de que o hipercubo constitui uma ferramenta eficaz para auxiliar nas análises de planejamento e operação deste sistema.

Finalmente, no próximo capítulo, serão realizadas avaliações sobre os efeitos de se implantar configurações operacionais alternativas no serviço, e também uma análise comparativa entre os vários cenários investigados.

6 AVALIAÇÃO DE CENÁRIOS ALTERNATIVOS

Este capítulo apresenta análises de configurações operacionais alternativas para o SAMU-192 de Campinas, e confirma a técnica do hipercubo como ferramenta de apoio às decisões de planejamento e operação de sistemas de transporte emergencial de saúde em cidades brasileiras.

Uma das maiores preocupações da coordenação do sistema é com relação aos tempos médios de resposta aos chamados. Este tempo é precioso do ponto de vista social, pois quanto mais rápido as vítimas começam a receber atendimento, maiores são suas chances de recuperação, e menores as probabilidades de agravamento dos casos. O tempo médio de resposta de um sistema de transporte emergencial de saúde é um dos principais indicadores do nível de serviço que está sendo oferecido.

Vários são os fatores que podem influenciar no tempo médio de resposta a um chamado, dentre eles: dia da semana, período do dia, condições do tráfego local, etc, que muitas vezes fogem do controle do sistema, e outros como capacitação profissional da equipe, número de veículos disponíveis e

suas localizações, que podem ser controlados e redimensionados pela equipe de gerência.

Uma das alternativas a ser avaliada é com relação aos efeitos causados devido a um aumento no número de ambulâncias para realizar as operações. Esta opção nem sempre é a mais viável para ser implantada, pois envolve a disponibilidade de recursos financeiros para a aquisição de novos veículos e contratação de equipe especializada, o que na maioria das cidades brasileiras é de responsabilidade do poder público municipal.

Por outro lado, opções mais simples, como por exemplo uma reorganização da estrutura existente, podem causar altos efeitos importantes sobre o nível de serviço oferecido, sem exigir que sejam feitos grandes investimentos.

Neste sentido, uma alternativa operacional a ser investigada e que pode melhorar a medida do tempo médio de resposta para o SAMU é a descentralização dos veículos, de tal forma que as distâncias percorridas até o local da ocorrência sejam minimizadas. Intuitivamente, as distâncias médias de viagem tornam-se menores; logo, os tempos médios de viagem também diminuem, implicando em respostas mais rápidas.

As análises que seguem apresentam opções de descentralização das ambulâncias e seus reflexos sobre alguns indicadores de desempenho para o SAMU, comparados aos produzidos com a configuração original do sistema. Também serão estudados os impactos de um possível aumento no número de veículos da operação.

Finalmente, serão feitas análises de sensibilidade com relação às possíveis alterações na estrutura do sistema e as conseqüências sobre o nível de serviço oferecido à população.

Por se tratar de investigação de cenários alternativos para o SAMU de Campinas, não é possível obter as taxas médias de atendimento reais para cada configuração operacional a seguir estudada. Por isso, todos os resultados que seguem foram obtidos utilizando-se o processo de calibração dos tempos médios de atendimento das ambulâncias, conforme discutido no Capítulo 3. Os valores iniciais considerados foram os tempos médios de atendimento observados no sistema, em sua configuração original. A precisão admitida foi da ordem de 10^{-6} , e a solução obtida, em geral, na quarta iteração.

6.1 Descentralização das ambulâncias

Segundo informações da coordenação do sistema, não existe uma legislação para os serviços de transporte emergencial de saúde brasileiros que exija a localização de uma base junto a um hospital, pronto-socorro, ou qualquer outro tipo de unidade de saúde do município. Entretanto, ao se concentrar as equipes em uma unidade deste tipo, há o aproveitamento da infraestrutura existente, gerando uma economia que pode vir a ser aplicada para a manutenção e desenvolvimento do sistema, além da possível colaboração de profissionais da unidade com as equipes de atendimento de emergência.

Neste sentido, os cenários descentralizados apresentados a seguir consideram como bases os principais centros de referência do sistema de urgência do município localizados em cada uma das regiões consideradas (Norte, Sul, Leste, Oeste e Central).

6.1.1 Cenário 1: descentralização de uma ambulância

FIGURA 6.1 - Primeiro cenário (1 VSB em O_B)

Este primeiro cenário investigado desloca o VSB número 5 para a região de maior índice de demanda do sistema (região Oeste), e considera como base desta equipe a unidade de atendimento de urgência local (Pronto-Socorro Ouro Verde). Os demais veículos permanecem centralizados na base do SAMU (região Central), ficando responsáveis por dar atendimento às outras regiões, sendo os VSB restantes *backup* para a região Oeste B.

No caso das regiões geradoras de chamados avançados, os seus servidores primários continuam sendo os veículos 1 e 2, e para a região Oeste A, em particular, o primeiro *backup* básico, caso os dois VSA estejam ocupados, passa a ser o VSB número 5.

Devido à nova organização do sistema, as hipóteses (5) e (7) da modelagem são modificadas, sendo que as demais permanecem como apresentadas no Capítulo 5. Para exemplificar, serão apresentadas as modificações ocorridas nestas hipóteses para o primeiro cenário.

6.1.1.3 Espaço de estados

A Figura 6.2 exemplifica o diagrama de fluxos considerando o estado (10000000101) nesta nova configuração operacional.

FIGURA 6.2 – Diagrama de fluxos para o estado (10000000101): cenário 1

Para os demais estados, os diagramas podem ser obtidos de maneira análoga.

6.1.1.4 Equações de balanço

Como um exemplo de equação de balanço, tomou-se o estado (10000000101) representado no diagrama de fluxos da Figura 6.3.

$$[(\lambda_2 + \lambda_4 + \lambda_6 + \lambda_8 + \lambda_{10}) + (\lambda_3 + \lambda_9) + (\lambda_7) + (\lambda_5) + (\lambda_1) + (\mu_1) + (\mu_3) + (\mu_{10})] \cdot p_{1000000101} = \mu_2 \cdot p_{1000000111} + \mu_4 \cdot p_{1000001101} + \mu_5 \cdot p_{1000010101} + \mu_6 \cdot p_{1000100101} + \mu_7 \cdot p_{1001000101} + \mu_8 \cdot p_{1010000101} + \mu_9 \cdot p_{1100000101} + (\lambda_4 + \lambda_8) \cdot p_{1000000100} + (\lambda_9) \cdot p_{1000000001}$$

As outras equações de balanço, bem como as equações dos estados que formam a cauda do hipercubo (fila), podem ser obtidas conforme apresentado no Capítulo 5.

6.1.1.5 Resultados

A Tabela 6.1 seleciona alguns indicadores de desempenho para esta configuração, comparados aos valores obtidos pelo modelo para o cenário original do sistema. Os desvios são calculados para os tempos médios de resposta.

TABELA 6.1 – Medidas de desempenho para o primeiro cenário

Ambulância	Cenário original		1º cenário		Desvio	
	Workload	Tempo médio de resposta (min)	Workload	Tempo médio de resposta (min)	minutos	%
1	0,39	10,88	0,39	11,34	0,46	4,19
2	0,39	10,90	0,41	10,58	-0,32	-2,92
3	0,63	14,22	0,64	14,16	-0,06	-0,42
4	0,64	14,24	0,63	14,17	-0,08	-0,54
5	0,62	14,23	0,66	8,80	-5,43	-38,15
6	0,64	14,24	0,64	14,16	-0,08	-0,53
7	0,63	14,23	0,65	14,16	-0,07	-0,46
8	0,62	14,22	0,65	14,17	-0,06	-0,41
9	0,64	14,24	0,64	14,17	-0,07	-0,51
10	0,64	14,24	0,63	14,16	-0,07	-0,50
Média VSA	0,39	10,89	0,40	10,96	0,07	0,64
Média VSB	0,63	14,23	0,64	13,49	-0,74	-5,19
D-P VSA	0,00	0,01	0,01	0,54	0,55	-
D-P VSB	0,01	0,01	0,01	1,90	1,90	-

(*) D-P: desvio-padrão

Neste cenário, observa-se, valores semelhantes aos originais para as *workloads* dos veículos, sendo que o veículo 5 apresentou um fator de utilização um pouco mais elevado que os demais. Isto explica-se pelo fato de que este veículo passou a ser o servidor primário da região Oeste, que

representa o maior índice de solicitações pelo serviço em todo o sistema. O tempo médio de resposta das ambulâncias é menor, em especial no caso da ambulância 5, localizada na região Oeste B, cuja redução é significativa (38,15%), melhorando sensivelmente a qualidade do serviço oferecido à população desta região específica. Para as demais regiões, ainda observa-se longas esperas. O tempo médio de resposta do sistema obtido através da expressão (17) sofreu uma redução média de 5,86%, de 13,82 para 13,01 minutos. O desvio absoluto para o sistema foi de 4,93%, sendo de 5,19 para os VSB's e 3,58 para os VSA's.

Esta primeira avaliação confirma a hipótese de que, sem alterar o número de ambulâncias, uma possível descentralização pode elevar o nível de serviço oferecido pelo sistema.

6.1.2 Cenário 2: descentralização de duas ambulâncias

FIGURA 6.3 – Segundo cenário (2 VSB em O_B)

Este cenário desloca os veículos números 3 e 5 para a região Oeste B, e também considera como base das equipes o Pronto-Socorro Ouro Verde. Os demais veículos permanecem centralizados no SAMU, ficando responsáveis por dar atendimento às demais regiões, funcionando como *backup* para a

região Oeste. A fim de ser comparado com o cenário 1, a Tabela 6.2 apresenta os indicadores de desempenho para este caso.

As *workloads* dos veículos continuam balanceadas, e com valores altos, sendo que os veículos que foram deslocados para a região Oeste B apresentaram uma suave redução neste valor, causada por sua dedicação ao atendimento de chamados desta região apenas (funcionando como *backup* para as demais regiões). O tempo médio de resposta das ambulâncias é menor, em especial nos casos 3 e 5, cuja redução também é significativa (32,41%, em média), elevando a qualidade do serviço oferecido às vítimas nesta região. Observa-se maior redução no tempo médio de resposta do sistema (expressão 17): de 13,82 (cenário original) para 12,53 minutos. Os desvios absolutos observados foram de 8,75% para o sistema, 9,70% para os VSB's e 3,81 para os VSA's).

TABELA 6.2 – Medidas de desempenho para o segundo cenário

Ambulância	Cenário original		2º cenário		Desvio	
	<i>Workload</i>	Tempo médio de resposta (min)	<i>Workload</i>	Tempo médio de resposta (min)	minutos	%
1	0,39	10,88	0,39	11,35	0,47	4,32
2	0,39	10,90	0,41	10,54	-0,36	-3,30
3	0,63	14,22	0,61	10,02	-4,20	-29,54
4	0,64	14,24	0,64	13,86	-0,38	-2,67
5	0,62	14,23	0,62	9,21	-5,02	-35,28
6	0,64	14,24	0,64	13,96	-0,28	-1,97
7	0,63	14,23	0,65	14,10	-0,13	-0,91
8	0,62	14,22	0,65	14,01	-0,21	-1,48
9	0,64	14,24	0,65	13,83	-0,41	-2,88
10	0,64	14,24	0,64	13,83	-0,41	-2,88
Média VSA	0,39	10,89	0,40	10,95	0,05	0,51
Média VSB	0,63	14,23	0,64	12,85	-1,38	-9,70
D-P VSA	0,00	0,01	0,01	0,57	0,59	-
D-P VSB	0,01	0,01	0,01	2,01	2,01	-

(*) D-P: desvio-padrão

Do ponto de vista prático, este cenário tem desempenho semelhante ao anterior. O indicador de desempenho que auxilia na escolha entre eles é tempo de resposta que, neste segundo cenário, apresenta uma redução média de 8% (cerca de 1 minuto), porém, com um desvio-padrão um pouco maior (1,96 ao invés de 1,41).

6.1.3 Cenário 3: descentralização de quatro ambulâncias

Esta alternativa operacional considera um sistema ainda mais descentralizado. Neste caso, cada unidade de atendimento de urgência das regiões passam a ser bases para as equipes (*região Norte: Pronto Socorro Vila Padre Anchieta; região Sul: Pronto Socorro São José; região Leste: Centro de Saúde Souza; região Oeste: Pronto Socorro Ouro Verde; região Central: base operacional do SAMU*).

FIGURA 6.4 – Terceiro cenário(1 VSB em N_B , S_B , L_B e O_B)

A lista de preferência de despachos deste cenário aloca, como servidores primários das regiões básicas, os veículos nelas posicionados. Os *backup* do sistema, neste caso, são escolhidos segundo o critério da menor distância entre o átomo gerador do chamado e a base do veículo a ser despachado. É importante observar que os VSA continuam centralizados no SAMU, bem como sendo servidores primários dos átomos geradores de

chamados da classe avançada, tendo como primeiros *backup*, quando os dois VSA estão ocupados, os veículos posicionados nos átomos básicos correspondentes.

A Tabela 6.3 mostra as medidas de desempenho do sistema para este cenário.

TABELA 6.3 – Indicadores de desempenho para o terceiro cenário

Ambulância	Cenário original		3º cenário		Desvio	
	<i>Workload</i>	Tempo médio de resposta (min)	<i>Workload</i>	Tempo médio de resposta (min)	minutos	%
1	0,39	10,88	0,39	10,89	0,01	0,06
2	0,39	10,90	0,39	10,89	-0,01	-0,11
3	0,63	14,22	0,64	11,71	-2,52	-17,68
4	0,64	14,24	0,60	12,44	-1,81	-12,68
5	0,62	14,23	0,59	10,69	-3,54	-24,91
6	0,64	14,24	0,70	9,96	-4,28	-30,08
7	0,63	14,23	0,64	12,03	-2,20	-15,47
8	0,62	14,22	0,60	12,68	-1,54	-10,82
9	0,64	14,24	0,62	13,00	-1,24	-8,72
10	0,64	14,24	0,61	12,92	-1,32	-9,24
Média VSA	0,39	10,89	0,39	10,89	0,00	-0,03
Média VSB	0,63	14,23	0,63	11,93	-2,31	-16,20
D-P VSA	0,00	0,01	0,00	0,00	0,01	-
D-P VSB	0,01	0,01	0,04	1,10	1,10	-

(*) D-P: desvio-padrão

Nota-se que as *workloads* sofrem um suave desbalanceamento, que é explicado pelas diferenças entre as demandas de cada região, especificamente. O veículo 6, por exemplo, que é posicionado na região Oeste B (maior índice de demanda do sistema), sobrecarrega-se com relação aos demais. No caso do veículo 5, posicionado na região Sul B, esta situação é inversa pois, além de ser uma das regiões de menor índice de demanda, sua localização faz com que os tempos médios de viagem para os demais átomos seja maior, colocando então o seu veículo nas últimas posições da lista de preferência de despachos para cada átomo (isto pode ser comprovado na Tabela 5.4).

Quanto aos tempos médios de resposta das ambulâncias, a redução é significativa em todos os casos; excessivas esperas deixam de ser observadas. O tempo médio de resposta do sistema a um chamado qualquer (expressão 17) também reflete esta característica, passando de 13,82 para 11,30 minutos (18.23%), com um desvio-padrão para o sistema um pouco menor (1,06 ao invés de 1,41). O desvio absoluto observado para o sistema foi de 13,60%, para os VSB's, 16,19% e para os VSA's, 0,09%.

Ainda com relação aos tempos médios de resposta, considerando-se agora a região de origem do chamado, os cenários 1 e 2 apresentaram diferenças significativas apenas para a região onde foram posicionados os veículos. Já no cenário 3 estas diferenças se dão em todo o sistema.

TABELA 6.4 – Tempos médios de resposta em cada átomo(cenários 1, 2 e 3)

Átomo	Tempo médio de resposta (minutos)				Desvio (%)		
	Original	Cenário 1	Cenário 2	Cenário 3	Cenário 1	Cenário 2	Cenário 3
1 - N_B	16,01	16,20	16,48	13,00	1,16	2,92	-18,85
2 - N_A	13,08	13,20	13,23	12,74	0,92	1,15	-2,61
3 - S_B	15,94	15,94	15,94	12,40	0,04	0,00	-22,18
4 - S_A	13,00	13,06	13,04	12,59	0,51	0,37	-3,09
5 - L_B	15,96	16,14	16,42	12,98	1,13	2,89	-18,65
6 - L_A	13,02	13,13	13,17	12,68	0,88	1,15	-2,63
7 - O_B	15,77	12,51	10,25	13,28	-20,66	-34,97	-15,83
8 - O_A	12,83	12,43	12,09	12,53	-3,13	-5,75	-2,29
9 - C_B	5,49	5,80	6,18	6,95	5,68	12,57	26,56
10 - C_A	4,31	4,48	4,52	4,64	3,76	4,73	7,51
Média A	12,54	11,26	11,21	11,04	0,58	0,33	-0,62
Média B	13,83	13,32	13,05	11,72	-2,53	-3,32	-9,79
D-P A	3,88	3,80	3,77	3,58	-	-	-
D-P B	4,67	4,48	4,65	2,69	-	-	-

(*) D-P: desvio-padrão

Nota-se uma ampla variabilidade nos valores dos tempos médios de resposta para cada átomo específico, nos três cenários estudados. Medidas mais agregadas como tempo médio de resposta de um servidor apresentam um grau de variabilidade bastante inferior. Este fato é consistente em se tratando

de resultados produzidos por um modelo probabilístico LARSON & ODoni (1981).

Portanto, considerando a medida de tempo médio de resposta, esta terceira opção mostra-se atraente, ficando um pouco prejudicada apenas no desbalanceamento das *workloads* das ambulâncias.

A Tabela 6.5 apresenta as frequências de despachos de cada ambulância: frequência total de atendimentos no sistema, frequência de atendimentos em sua área de cobertura primária e frequências de atendimentos *backup*.

TABELA 6.5 – Frequências de despachos para cada ambulância (cenário 3)

Ambulância	Frequências de despachos		
	área de cobertura	<i>backup</i>	total
1	0,0352	0,0304	0,0656
2	0,0412	0,0292	0,0704
3	0,0596	0,0498	0,1094
4	0,0339	0,0715	0,1054
5	0,0667	0,0363	0,1030
6	0,0860	0,0380	0,1240
7	0,0578	0,0530	0,1108
8	0,0548	0,0592	0,1140
9	0,0218	0,0690	0,0908
10	0,0105	0,0962	0,1067
Média VSA	0,0382	0,0298	0,0680
Média VSB	0,0489	0,0591	0,1080

Com estes valores, conclui-se que, nesta configuração operacional, 47% dos atendimentos realizados foram para as áreas de cobertura primária dos veículos, ou seja, onde estão posicionados, e 53% foram inter-áreas, ou seja, *backup*. Ao se considerar exclusivamente os VSA, 56% dos atendimentos realizados foram para suas áreas de cobertura primária.

Também foram verificadas as frequências de chamados que foram atendidos por seu veículo de preferência primária e por veículos *backup* para todos os átomos do sistema. A Tabela 6.6 apresenta estes valores.

Nota-se, nesta tabela, que o número de chamados atendidos por veículos *backup* é relativamente alto (em torno de 34%). Apenas 13% dos chamados da classe avançada receberam atendimento de equipes *backup*, ou seja, VSB.

TABELA 6.6 – Frequências de atendimentos para cada átomo (cenário 3)

Átomo	Frequências de atendimentos		
	veículo primário	<i>backup</i>	total
1 - N _B	0,0585	0,0994	0,1579
2 - N _A	0,0164	0,0029	0,0193
3 - S _B	0,0658	0,0921	0,1579
4 - S _A	0,0123	0,0021	0,0144
5 - L _B	0,0570	0,1009	0,1579
6 - L _A	0,0123	0,0021	0,0144
7 - O _B	0,0848	0,1927	0,2775
8 - O _A	0,0203	0,0036	0,0239
9 - C _B	0,1177	0,0353	0,1530
10 - C _A	0,0224	0,0015	0,0239
Média A	0,0167	0,0024	0,0192
Média B	0,0768	0,1041	0,1808

6.1.4 Cenário 4: descentralização de cinco ambulâncias

FIGURA 6.5 – Quarto cenário (1 VSB em N_B, S_B e L_B, 2 VSB em O_B)

O terceiro cenário mostrou que, embora estando o sistema descentralizado, o veículo posicionado na região Oeste B sofre uma sobrecarga de trabalho com esta configuração se comparado aos outros

veículos do sistema. Sendo assim, na tentativa de redimensionar a carga de trabalho dos veículos, este cenário toma uma variação do cenário 3, posicionando mais uma equipe na região de maior solicitação de serviço. A Tabela 6.7 apresenta os resultados obtidos.

TABELA 6.7 – Indicadores de desempenho para o quarto cenário

Ambulância	Cenário original		4º cenário		Desvio	
	<i>Workload</i>	Tempo médio de resposta (min)	<i>Workload</i>	Tempo médio de resposta (min)	minutos	%
1	0,39	10,88	0,39	10,87	-0,01	-0,09
2	0,39	10,90	0,39	10,87	-0,03	-0,28
3	0,63	14,22	0,64	11,69	-2,53	-17,79
4	0,64	14,24	0,60	12,76	-1,48	-10,39
5	0,62	14,23	0,60	10,96	-3,27	-22,98
6	0,64	14,24	0,65	10,45	-3,79	-26,62
7	0,63	14,23	0,64	12,03	-2,20	-15,46
8	0,62	14,22	0,61	12,00	-2,22	-15,61
9	0,64	14,24	0,65	10,45	-3,78	-26,62
10	0,64	14,24	0,60	12,14	-2,10	-14,75
Média VSA	0,39	10,89	0,39	10,87	-0,02	-0,18
Média VSB	0,63	14,23	0,63	11,56	-2,67	-18,78
D-P VSA	0,00	0,01	0,00	0,00	0,01	-
D-P VSB	0,01	0,01	0,02	0,85	0,85	-

(*) D-P: desvio-padrão

Nota-se por estes resultados que as *workloads* dos veículos continuam altas e ligeiramente desbalanceadas devido aos diferentes índices de demandas das regiões, mas os tempos médios de resposta das equipes são menores que os observados no cenário 3; na prática, isto significa um aumento no nível de serviço oferecido.

O valor do tempo médio de resposta do sistema obtido foi de 11,41 minutos (17,43 % menor com relação à configuração original do sistema). Houve também uma redução nos tempos médios de resposta para cada átomo, e os valores estão apresentados na Tabela 6.8.

TABELA 6.8 – Tempos médios de resposta para cada átomo (cenário 4)

Átomo	Tempo médio de resposta (min)		Desvio	
	Original	Cenário 4	minutos	%
1 - N _B	16,01	13,26	-2,75	-17,22
2 - N _A	13,08	12,78	-0,31	-2,36
3 - S _B	15,94	12,49	-3,44	-21,58
4 - S _A	13,00	12,59	-0,41	-3,15
5 - L _B	15,96	13,25	-2,71	-16,95
6 - L _A	13,02	12,72	-0,30	-2,33
7 - O _B	15,77	11,08	-4,68	-29,71
8 - O _A	12,83	12,22	-0,62	-4,82
9 - C _B	5,49	5,62	0,12	2,21
10 - C _A	4,31	4,43	0,10	2,37
Média A	12,54	10,95	-0,31	-
Média B	13,83	11,14	-2,69	-
D-P A	3,88	3,65	0,26	-
D-P B	4,67	3,21	1,76	-

(*) D-P: desvio-padrão

Com relação às frequências de despachos, a Tabela 6.9 apresenta os valores obtidos.

TABELA 6.9 – Frequências de despachos para cada ambulância (cenário 4)

Ambulância	Frequências de despachos		
	área de cobertura	backup	total
1	0,0352	0,0296	0,0648
2	0,0413	0,0285	0,0698
3	0,0593	0,0506	0,1099
4	0,0292	0,0767	0,1059
5	0,0663	0,0371	0,1034
6	0,0615	0,0464	0,1079
7	0,0564	0,0556	0,1120
8	0,0606	0,0489	0,1095
9	0,0910	0,0303	0,1213
10	0,0250	0,0751	0,0956
Média VSA	0,0383	0,0291	0,0673
Média VSB	0,0556	0,0526	0,1082

Nesta configuração operacional, 53% dos atendimentos realizados foram para as áreas de cobertura primária dos veículos, ou seja, onde estão posicionados, e 47% foram inter-áreas. No caso dos VSA, 57% dos atendimentos realizados foram para suas áreas de cobertura primária.

Com relação aos átomos do sistema, os resultados estão apresentados na Tabela 6.10.

TABELA 6.10 – Frequências de atendimentos para cada átomo (cenário 4)

Átomo	Frequências de atendimentos		
	veículo primário	<i>backup</i>	total
1 - N _B	0,0582	0,0998	0,1580
2 - N _A	0,0164	0,0029	0,0193
3 - S _B	0,0654	0,0925	0,1579
4 - S _A	0,0124	0,0019	0,0143
5 - L _B	0,0556	0,1023	0,1579
6 - L _A	0,0122	0,0021	0,0143
7 - O _B	0,1504	0,1272	0,2276
8 - O _A	0,0212	0,0026	0,0238
9 - C _B	0,1074	0,0455	0,1529
10 - C _A	0,0221	0,0020	0,0241
Média A	0,0169	0,0023	0,0192
Média B	0,0874	0,0935	0,1809

O número de chamados atendidos por veículos *backup* ainda é considerado alto (em torno de 32%). Dos chamados da classe avançada, 12% receberam atendimento de equipes básicas, ou seja, *backup*.

6.1.5 Cenário 5: descentralização de seis ambulâncias

Completando as avaliações sobre uma possível descentralização do sistema, este cenário considera a configuração do cenário 4 e aloca um dos veículos da região Central para também realizar o atendimento na região Norte que, dentre as três regiões que são bases de apenas um veículo, apresentou o maior índice de demanda no período observado.

As medidas de desempenho para esta configuração são praticamente idênticas às obtidas com a configuração 4, a menos das relativas aos veículos posicionados na região Norte B (veículos 3 e 10) que diminuiram com relação ao cenário anterior, elevando a qualidade do serviço oferecido. O tempo médio de resposta do sistema para qualquer chamado é de 11,48 minutos, 16,94% inferior ao observado na configuração original onde os veículos permanecem centralizados na base do SAMU.

FIGURA 6.6 – Quinto cenário (1 VSB em S_B e L_B , 2 VSB em N_B e O_B)

TABELA 6.11 – Indicadores de desempenho para o quinto cenário

Ambulância	Cenário original		5º cenário		Desvio	
	Workload	Tempo médio de resposta (min)	Workload	Tempo médio de resposta (min)	minutos	%
1	0,39	10,88	0,39	10,96	0,07	0,66
2	0,39	10,90	0,39	10,81	-0,09	-0,86
3	0,63	14,22	0,58	12,68	-1,54	-10,85
4	0,65	14,24	0,65	12,30	-1,95	-13,66
5	0,62	14,23	0,62	11,39	-2,84	-19,98
6	0,64	14,24	0,65	10,49	-3,75	-26,34
7	0,63	14,23	0,67	10,43	-3,80	-26,69
8	0,62	14,22	0,64	11,55	-2,67	-18,76
9	0,64	14,24	0,62	11,64	-2,60	-18,24
10	0,65	14,24	0,61	12,57	-1,67	-11,71
Média VSA	0,39	10,89	0,39	10,88	-0,01	-0,18
Média VSB	0,63	14,23	0,63	11,63	-2,60	-18,78
D-P VSA	0,00	0,01	0,00	0,11	0,12	-
D-P VSB	0,01	0,01	0,03	0,87	0,87	-

(*) D-P: desvio-padrão

O maior reflexo deste cenário está no tempo médio de resposta para cada átomo.

TABELA 6.12 – Tempos médios de resposta para cada átomo (cenário 5)

Átomo	Tempo médio de resposta (min)		Desvio	
	Original	Cenário 5	minutos	%
1 - N _B	16,01	10,98	-5,03	-31,41
2 - N _A	13,09	12,46	-0,62	-4,75
3 - S _B	15,93	13,24	-2,69	-16,87
4 - S _A	13,00	12,72	-0,29	-2,20
5 - L _B	15,96	13,14	-2,82	-17,66
6 - L _A	13,03	12,71	-0,32	-2,43
7 - O _B	15,76	11,47	-4,29	-27,21
8 - O _A	12,83	12,29	-0,54	-4,23
9 - C _B	5,50	6,02	0,52	9,49
10 - C _A	4,33	4,36	0,03	0,75
Média A	12,54	10,91	-0,31	-
Média B	13,83	10,97	-3,15	-
D-P A	3,88	3,66	0,26	-
D-P B	4,67	2,94	2,05	-

(*) D-P: desvio-padrão

As frequências de despacho de cada veículo e frequências de atendimentos inter-átomos estão mostradas nas tabelas 6.13 e 6.14.

TABELA 6.13 – Frequências de despachos para cada ambulância (cenário 5)

Ambulância	Frequências de despachos		
	área de cobertura	backup	total
1	0,0412	0,0288	0,0700
2	0,0353	0,0299	0,0652
3	0,0416	0,0479	0,0895
4	0,0589	0,0503	0,1092
5	0,0607	0,0477	0,1084
6	0,0683	0,0372	0,1055
7	0,0832	0,0415	0,1247
8	0,0386	0,0595	0,0981
9	0,0504	0,0673	0,1177
10	0,0564	0,0550	0,1114
Média VSA	0,0383	0,0294	0,0676
Média VSB	0,0573	0,0508	0,1081

Nesta configuração, 54% dos atendimentos realizados foram para as áreas de cobertura primária dos veículos, e 46% foram inter-áreas. No caso dos VSA, 60% dos atendimentos realizados se deram em suas áreas de cobertura primária.

TABELA 6.14 – Frequências de atendimentos para cada átomo (cenário 5)

Átomo	Frequências de atendimentos		
	veículo primário	<i>backup</i>	total
1 - N _B	0,0961	0,0618	0,1579
2 - N _A	0,0172	0,0020	0,0192
3 - S _B	0,0599	0,0979	0,1578
4 - S _A	0,0123	0,0020	0,0143
5 - L _B	0,0581	0,0998	0,1579
6 - L _A	0,0123	0,0020	0,0143
7 - O _B	0,1494	0,1282	0,2776
8 - O _A	0,0212	0,0026	0,0238
9 - C _B	0,0868	0,662	0,1530
10 - C _A	0,0213	0,0026	0,0239
Média A	0,0169	0,0022	0,0191
Média B	0,0901	0,0908	0,1808

A proporção de chamados atendidos por veículos *backup* é de aproximadamente 31%. Dos chamados da classe avançada, 10% receberam atendimento de equipes básicas.

Os resultados destes cinco cenários levam a concluir que a operação descentralizada é uma boa alternativa para o SAMU-192, provocando uma sensível redução no tempo médio de resposta a um chamado emergencial (também chamado tempo de salvamento), conseqüentemente aumentando as chances de sobrevivência das vítimas.

Quanto às frequências de despachos inter-átomos, ou seja, atendimentos realizados por veículos *backup*, os valores permaneceram altos, mesmo com as variações de descentralização presentes nos cenários 3, 4 e 5. Este fato é facilmente compreendido se considerar o baixo número de ambulâncias em operação para atender o alto índice de solicitações do sistema.

Algumas outras alternativas foram examinadas, como por exemplo, posicionando dois veículos nas Sul B e Leste B e não na região Norte B, mas os resultados não foram tão satisfatórios, pois tratam-se de regiões de menores taxas de solicitação pelo serviço.

Questões sobre o reposicionamento de VSA também foram levantadas. Entretanto, são condições inviáveis para o sistema, pois todo atendimento realizado por uma equipe avançada possui um médico acompanhando, e devido à escassez de recursos, este profissional também desempenha a função de médico regulador das urgências. Logo, ao se deslocar a base dos VSA da central de regulação médica (base do SAMU), automaticamente deslocam-se os profissionais que compõem as equipes de atendimento, prejudicando o desempenho do serviço como um todo.

6.2 Aumentando o número de ambulâncias no sistema

Os próximos cenários investigados têm por objetivo avaliar os impactos causados nas medidas de desempenho do sistema devido ao aumento do número de veículos na operação.

6.2.1 Cenário 6: sistema operando com um novo VSB

Este cenário considera a configuração do cenário 5, e aloca um novo veículo de suporte básico na região Leste, já que apresentou maior *workload* se comparada à região Sul. O mesmo teste foi realizado posicionando este novo veículo na região Central, que é *backup* para todas as demais regiões e os resultados foram satisfatórios.

A Tabela 6.15 mostra as diferenças nas medidas de desempenho provocadas pelo aumento de uma ambulância básica no SAMU-192, associadas à descentralização das operações. As *workloads* diminuem, em média, 17% (tanto para os casos básicos, quanto para os casos avançados). Os

tempos médios de resposta das ambulâncias reduzem em torno de 14% para os veículos básicos e 9% para os veículos avançados. Os tempos médios de resposta a cada átomo reduzem, em média, 4% com relação aos observados no cenário 5, e o tempo médio de resposta a um chamado qualquer é de 10,52 minutos.

FIGURA 6.7 - Sexto cenário: aquisição de um novo VSB

TABELA 6.15 – Indicadores de desempenho para o sexto cenário

Ambulância	VSB – região Central		VSB – Região Leste	
	Workload	Tempo médio de resposta (min)	Workload	Tempo médio de resposta (min)
1	0,31	10,06	0,31	10,02
2	0,34	9,72	0,32	9,42
3	0,57	11,17	0,56	11,00
4	0,58	10,57	0,56	9,93
5	0,56	9,61	0,55	9,50
6	0,62	9,45	0,62	9,63
7	0,62	9,45	0,62	9,25
8	0,55	10,28	0,52	11,36
9	0,51	13,21	0,51	11,21
10	0,50	11,44	0,50	11,16
11	0,55	10,59	0,55	10,51
Média VSA	0,32	10,45	0,31	10,21
Média VSB	0,56	10,58	0,56	10,32
D-P VSA	0,02	1,03	0,01	1,12
D-P VSB	0,04	1,27	0,04	0,85

(*) D-P: desvio-padrão

As frequências de despacho de cada veículo estão apresentadas na Tabela 6.16, e foram calculadas para a opção de posicionamento do novo veículo na região Central, já que esta produziu melhores tempos médios de resposta a um chamado.

Nesta configuração, 61% dos atendimentos realizados foram para as áreas de cobertura primária dos veículos, e 39% foram inter-áreas. No caso dos VSA, 72% dos atendimentos realizados se deram em suas áreas de cobertura primária.

Em todos os átomos do sistema, a proporção de chamados atendidos por veículos *backup* é de aproximadamente 27%. Dos chamados da classe avançada, apenas 7% receberam atendimento de equipes básicas.

TABELA 6.16 – Frequências de despachos para cada ambulância (cenário 6)

Ambulância	Frequências de despachos		
	área de cobertura	<i>backup</i>	total
1	0,0375	0,0158	0,0533
2	0,0445	0,0150	0,0595
3	0,0627	0,0437	0,1064
4	0,0671	0,0349	0,1020
5	0,0713	0,0275	0,0988
6	0,0711	0,0290	0,1001
7	0,900	0,0318	0,1218
8	0,0663	0,0375	0,1038
9	0,0181	0,0678	0,0859
10	0,0437	0,0361	0,0798
11	0,0353	0,0530	0,0883
Média VSA	0,0410	0,0154	0,0564
Média VSB	0,0584	0,0401	0,0985

Este cenário permite concluir, portanto, que o aumento do número de veículos pode elevar significativamente o desempenho do sistema.

6.2.2 Cenário 7: sistema operando com 2 novos VSB

FIGURA 6.8 - Sétimo cenário: aquisição de dois novos VSB

A construção deste cenário foi incentivada pelos resultados obtidos com o cenário 6. Aqui, um outro veículo de suporte básico é adquirido e posicionado na região Leste. Com exceção da região Sul que apresenta a menor taxa de solicitação pelo serviço, e por ser a mais distante dos demais átomos, é escolhida para as últimas posições de *backup* na lista de preferência de despachos, todas as demais regiões são atendidas por ao menos duas ambulâncias. Os resultados estão apresentados a seguir.

Observa-se, com duas novas ambulâncias básicas uma expressiva redução e um certo balanceamento das *workloads* (com exceção no caso dos veículos posicionados na região Oeste). Comparadas ao cenário original com 8 VSB e 2 VSA centralizados na base do SAMU, esta configuração apresenta *workloads* 30% inferiores, resultando também na redução das chances de um chamado ocorrer enquanto todas as ambulâncias estão ocupadas (0,12 na configuração original e 0,05 para esta descentralização).

TABELA 6.17 – Indicadores de desempenho para o sétimo cenário

Ambulância	2 novos VSB	
	<i>Workload</i>	Tempo médio de resposta (min)
1	0,27	7,67
2	0,27	7,58
3	0,44	8,88
4	0,49	9,22
5	0,50	9,11
6	0,54	8,92
7	0,54	8,87
8	0,43	8,09
9	0,45	8,15
10	0,45	8,80
11	0,44	8,53
12	0,48	8,75
Média VSA	0,27	7,63
Média VSB	0,48	8,73
D-P VSA	0,00	0,06
D-P VSB	0,04	0,37

(*) D-P: desvio-padrão

Os tempos médios de resposta das ambulâncias reduzem em torno de 25% para os veículos básicos e 30% para os veículos avançados. Os tempos médios de resposta a cada átomo reduzem, em média, 25% com relação à configuração original do sistema, e o tempo médio de resposta a um chamado qualquer é de 8,67 minutos, uma diferença de 3,8 minutos com relação ao cenário original. Em se tratando de um sistema de atendimento de urgência, estas reduções nos tempos médios de resposta contribuem diretamente na definição das condições de sobrevivência das vítimas, em situações de risco.

As frequências de despacho *backup* de cada veículo também diminuíram: 70% dos atendimentos realizados foram para as áreas de cobertura primária dos veículos. No caso dos VSA, 82% dos atendimentos realizados se deram em suas áreas de cobertura primária.

Em todos os átomos do sistema, a proporção de chamados atendidos por veículos *backup* é de aproximadamente 25%. Dos chamados da classe avançada, apenas 5% receberam atendimento de equipes básicas (*backup*).

Com base nestes resultados, foi estudada a situação em que o sistema opera ‘também com um novo veículo avançado. Os resultados estão apresentados no cenário a seguir.

6.2.3 Cenário 8: sistema operando com 2 novos VSB e 1 novo VSA

FIGURA 6.9 – Oitavo cenário: aquisição de dois novos VSB e um novo VSA

Este cenário apresenta uma forma não usual para o sistema, pois aloca o novo VSA na região Oeste (que não coincide com a central de regulação do sistema). Neste caso, como no cenário 6, outras possibilidades operacionais foram avaliadas, porém esta em particular mostrou-se a melhor alternativa de operação para o SAMU dentre as investigadas. Note que ela produz resultados comparáveis aos padrões internacionais de excelência (BRANDEAU & LARSON, 1986), como poderá ser constatado a seguir.

Nesta configuração, as frequências de despacho *backup* de cada ambulância são muito pequenas, apenas 18% dos atendimentos realizados não foram para as áreas de cobertura primária dos veículos. No caso dos VSA, 93% dos atendimentos realizados se deram em suas áreas de cobertura primária.

Em todos os átomos do sistema, a proporção de chamados atendidos por veículos *backup* é de aproximadamente 15%. Dos chamados da classe avançada, 3% receberam atendimento de equipes básicas (*backup*).

TABELA 6.18 – Indicadores de desempenho para o sétimo cenário

Ambulância	2 novos VSB e 1 novo VSA	
	Workload	Tempo médio de resposta (min)
1	0,17	7,45
2	0,18	7,13
3	0,41	8,75
4	0,42	8,36
5	0,42	7,74
6	0,45	7,87
7	0,44	7,82
8	0,40	8,09
9	0,42	8,15
10	0,42	8,67
11	0,41	8,12
12	0,42	8,18
13	0,17	6,56
Média VSA	0,17	7,05
Média VSB	0,42	8,18
Média do sistema	0,36	7,92
Padrões internacionais	0,3090	7,41

As avaliações de cenários alternativos realizadas construíram relações entre as medidas de desempenho produzidas no sistema e possíveis alterações em sua estrutura operacional, constituindo ferramental de apoio às decisões de caráter operacional e de planejamento para este tipo de sistema.

6.3 Comparações entre as alternativas operacionais

A seguir, para sintetizar as conclusões, será apresentada uma análise comparativa dos resultados obtidos para os vários cenários estudados.

6.3.1 Workload

A Figura 6.10 mostra a relação entre as *workloads* médias de cada tipo de veículo para todos os cenários estudados. Observa-se que, à medida que

crece a descentralização do sistema, há uma redução dos fatores de utilização dos veículos.

FIGURA 6.10 – Comparação entre as *workloads* médias do sistema

Também pode-se concluir, comparando-se os cenários 6, 7 e 8 com os demais, que o aumento no número de veículos na operação associado à descentralização do sistema diminui significativamente as *workloads* (e assim, diminui a probabilidade de um chamado encontrar todos as equipes ocupadas). Os cenários (1)-(5), que consideram a operação com os dez veículos do sistema, apresenta *workloads* médias de 0,39, no caso de VSA's e 0,63, no caso de VSB's. Comparando-se os cenários 6 e 5, conclui-se que uma nova ambulância básica na operação provoca uma reduções médias das *workloads* em torno de 20% no caso dos VSA's, e de 11% no caso de VSB's, porém, os desvios-padrão aumentam (de 0,00 para 0,02 e de 0,01 para 0,04, para VSA's e VSB's, respectivamente). Este raciocínio pode ser estendido para os cenários 7 e 8, gerando conclusões idênticas.

Sistemas com altos fatores de utilização dos servidores produzem medidas mais balanceadas se centralizado. Conforme o número de servidores cresce, as *workloads* decrescem, e então, um possível reposicionamento pode gerar um certo equilíbrio entre as medidas.

6.3.2 Probabilidades

Estas análises têm reflexos diretos sobre as probabilidades de sistema vazio, isto é todos os veículos disponíveis, probabilidade de sistema cheio, onde todos os veículos encontram-se ocupados, probabilidade de formação de fila e probabilidade de saturação do sistema. As Figuras 6.11 e 6.12 mostram algumas destas relações.

FIGURA 6.11 – Impactos do aumento do número de servidores sobre p_Q e p_S

Nota-se que as probabilidades de saturação e formação de fila são sensíveis ao aumento do número de ambulâncias no sistema. A redução média, no caso do aumento de um veículo é praticamente de 39%, com relação ao cenário original e de 30% com relação ao cenário totalmente descentralizado (cenário 5).

O mesmo acontece com a probabilidade de se encontrar o sistema vazio, apresentada na Figura 6.12. O aumento no número de ambulância provoca aumentos da ordem de 40% no valor desta probabilidade.

FIGURA 6.12 – Impactos do aumento do número de ambulâncias sobre $p_{00...0}$

Com relação à probabilidade de perda do sistema, o cenário original, bem como os cenários 1 a 5, que operavam com dez ambulâncias, apresentaram valor de 0,0001. Os cenários 6, 7 e 8 apresentaram, respectivamente, os valores 0,00001, 0,000001 e 0,0000001, o que na prática representa mínima chance de, no instante da ocorrência de um chamado, o sistema encontrar-se com todos os veículos ocupados e com dez chamados em fila.

6.3.3 Tempos médios de resposta

As curvas das Figuras 6.13, 6.14 mostram que sistemas descentralizados produzem tempos de resposta menores, e na prática, isto representa níveis de serviços mais elevados para os usuários, pois é um dos principais fatores que influenciam nas condições de sobrevivência das vítimas. Estas observações também são verificadas se o número veículos na operação cresce.

FIGURA 6.13 – Tempos médios de resposta de cada ambulância

FIGURA 6.14 – Comparação entre os tempos médios de resposta a cada átomo

A descentralização do sistema com dez ambulâncias provoca uma redução média de 18% no tempo de resposta das ambulâncias (em torno de 2,5 minutos). Os desvios-observados aumentaram (por exemplo, de 0,01 para até 1,96 no caso de veículos básicos, e de 0,00 a 0,11 nos casos de veículos avançados). O aumento de um veículo básico produz reduções de 25% nos tempos de resposta dos VSB's e 4% dos VSA's. Ao se aumentar um VSA e dois VSB's, considerando a descentralização, esta redução aumenta de 25% para 42%, e de 4% para 35%, respectivamente.

6.3.4 Tempo total de atendimento

O tempo total de atendimento de uma equipe é a composição entre os tempos de preparação da equipe, viagem, em cena e retorno. Através do processo de calibração das taxas médias de atendimento, o modelo gera novos tempos médios para a configuração estudada. A Figura 6.15 mostra a redução dos tempos médios de atendimento para cada tipo de veículo do sistema.

FIGURA 6.15 – Tempos de atendimento para os veículos do sistema

Como era esperado, as reduções no tempo de atendimento acabam refletindo as reduções do tempo médio de resposta. Somente a descentralização das ambulâncias provoca reduções médias de 10% no tempo médio de atendimento (cerca de 7 minutos). Aumentando-se o número de ambulâncias em 1, esta redução é de 17%. O cenário 8, onde o número de VSB's cresce em 2 e o número de VSA's em 1, a redução é de 21%.

6.3.5 Atendimentos *backup*

A proporção dos atendimentos realizados fora da área de cobertura primária dos veículos decresce à medida que o número de veículos aumenta (implicitamente, aumenta o número de veículos disponíveis para realizar a

cobertura dos átomos). O mesmo acontece com a proporção de chamados atendidos por um veículo *backup*.

FIGURA 6.16 – Despachos inter-átomos e chamados atendidos por *backup*

Um importante indicador que reflete as variações no nível de serviço oferecido em sistemas com classes diferenciadas de usuários, como é o caso do SAMU de Campinas, é a fração de chamados da classe “avançada” que recebe atendimento de uma equipe de suporte básico (*backup*). Esta medida indica o nível de deterioração do serviço. Nestes casos, o usuário está recebendo um atendimento inferior às suas reais necessidades.

Este valor é determinado pela frequência de despachos de veículos de suporte básico para atender chamados das regiões avançadas (N_A , S_A , L_A , O_A e C_A). Para o cenário original, a frequência de despachos de um veículo básico qualquer para atender um chamado da classe avançada é de 0,019 (19%). Ao aumentar o número de VSB's em 1, este valor cai para 13,5%. No caso do cenário 8, quando o número de VSA's é acrescido em 1, este valor reduz-se sensivelmente, 5%. Ao se pensar do ponto de vista do usuário, eleva-se sensivelmente a qualidade do serviço recebido.

7 CONCLUSÕES E PERSPECTIVAS

Este trabalho apresentou a análise de um sistema urbano de atendimento de urgência de uma cidade do interior paulista, por meio do modelo hipercubo de filas. Tal modelo foi desenvolvido na década de 70 para auxiliar nas avaliações do serviço de patrulhamento policial da cidade de Nova Iorque, mostrando-se uma ferramenta eficaz para avaliar serviços de transporte por ambulância, quando aplicado para o redimensionamento do serviço oferecido na cidade de Boston, na década de 80.

O modelo foi utilizado para representar as características brasileiras de atendimento e produzir medidas de desempenho que avaliem o nível de serviço oferecido, em condições típicas de uma cidade de grande porte.

O processo de planejamento de sistemas de transporte por ambulância deve ser realizado considerando-se as características da demanda local, levando-se em conta fatores como configuração geográfica do município, índices de demanda diferenciados para as diversas regiões, e classes de usuários diferenciadas para os casos de serviços em que os tipos de chamados exigem equipes também diferenciadas de atendimento.

7.1 CONCLUSÕES

Os cenários brasileiros estudados mostraram sistemas com altos índices de demanda e um número reduzido de veículos para realizar a operação. Em consequência desta distribuição, a carga média de trabalho dos veículos apresentou níveis altos para sistemas de atendimento de urgência. Isto justificou investigações sobre a operação atual e possíveis alterações nos sistemas buscando elevar o nível de serviço oferecido

Os resultados da verificação das hipóteses críticas da modelagem mostram que operações do tipo do SAMU de Campinas são bem representadas analiticamente por esta ferramenta, produzindo desvios pouco significativos dos valores gerados pelo modelo com relação aos observados no sistema real.

Da aplicação do modelo às condições da cidade de Campinas, concluiu-se que várias são as alternativas operacionais que podem elevar o nível de serviço oferecido, dentre elas a descentralização e o aumento do número de ambulâncias para realizar a operação. Se, por um lado, é óbvio que melhorias são obtidas com o aumento do número de ambulâncias no sistema, por outro lado, quantificar tais melhorias por meio de índices de desempenho, não é um processo tão evidente, e o modelo hipercubo de filas é um instrumento eficaz para isto.

As opções de descentralização das ambulâncias que atuam hoje no SAMU são promissoras, e pode-se concluir, no estudo de caso, que:

- i) Partindo de um caso simples, onde se desloca um dos veículos centralizados para a região de maior índice de demanda, ocorre uma ligeira redução nas medidas de tempos médios de resposta. Para o sistema, este tempo diminui cerca de 5,86%, o que equivale a uma redução média de 1 minuto. Porém, os impactos são sentidos

especialmente pelos chamados da região que passa a ser base da ambulância, diminuindo em torno de 38% (5 minutos);

- ii) Aumentando-se o nível de descentralização considerando o mesmo número de ambulâncias, os resultados mostram-se ainda melhores; as reduções nos tempos médios de resposta refletem-se para todos os atendimentos do sistema, independente da sua região de origem. Alternativas simples de reposicionamento dos servidores implicam em reduções médias de 18% do tempo de resposta a um chamado (cerca de 3 minutos), com desvio-padrão de 0,82 minutos, e isto pode representar diferenças significativas no possível estado futuro das vítimas.

Associado à descentralização, o aumento do número de ambulâncias no sistema provoca uma melhora considerável nas medidas de desempenho obtidas para o sistema, podendo ser comparadas aos padrões internacionais de atendimento.

Análises comparativas mostram que o aumento de um único veículo básico no sistema, provoca uma redução média de 17% nas *workloads* das ambulâncias, e o tempo de resposta a um chamado qualquer diminui cerca de 23% com relação ao cenário original. As probabilidades de saturação do sistema e formação de fila sofrem uma redução média de 30% com o aumento de um veículo.

Quanto aos atendimentos *backup*, a fração de chamados da classe avançada que são atendidos por equipes básicas também é sensível às variações do número de ambulâncias na operação, e os valores são ainda mais acentuados quando se trata do aumento de um veículo avançado (onde 5% dos casos apenas, recebem um atendimento não adequado, isto é, de uma equipe básica, em comparação com 19%, no cenário original). Este fato representa, na prática uma sensível elevação do nível de serviço oferecido. A situação

inversa (ou seja, um chamado básico ser atendido por uma equipe avançada) também existe, porém em diferentes proporções (7% no cenário original e 2% no cenário descentralizado, com dois novos VSB e um novo VSA).

Todos os esforços despendidos no trabalho, juntamente com as conclusões obtidas através da aplicação do modelo hipercubo ao SAMU de Campinas, contribuíram para difundir, para a Rede Brasileira de Cooperação em Emergências (RBCE), a importância de conciliar a experiência administrativa da classe médica, com o uso de ferramentas analíticas para auxiliar nas decisões estratégicas do atendimento emergencial de saúde à população brasileira.

7.2 PERSPECTIVAS PARA FUTURAS PESQUISAS

Da análise das informações fornecidas, dos resultados e do contato constante com a coordenação do sistema, várias sugestões para estudos futuros foram surgindo.

Uma alternativa interessante a ser pesquisada é a remodelagem das regiões do sistema, uma vez que a subdivisão utilizada pela Secretaria Municipal de Saúde como sendo de cobertura dos Centros de Saúde pode não refletir as reais necessidades de demanda de um serviço de atendimento de urgência.

Ainda neste sentido, um estudo especial sobre a partição da região Oeste (Oeste 1 e Oeste 2, por exemplo), não avaliada neste trabalho por ausência de informações, também seria interessante, uma vez que esta representa o maior índice de solicitação de chamados. O objetivo deste estudo é buscar o balanceamento das taxas médias de solicitação de serviço em cada região do sistema, visando equilibrar as *workloads* das ambulâncias.

Estudos híbridos envolvendo o modelo hipercubo e técnicas de otimização para resolver problemas como localização estratégica de bases com restrições de cobertura, tanto para áreas urbanas quanto rural, ou ainda através da utilização de um sistema de informações geográficas para auxiliar o redimensionamento das regiões também são atraentes.

Outra importante contribuição nesta direção, ainda não explorada na literatura, é o estudo de variabilidade, visto que cada indicador de desempenho produzido pelo modelo é caracterizado apenas pela média da distribuição de equilíbrio.

A dificuldade do usuário em contatar a central 192, tendo em vista o volume de chamados recebidos (16% de chamados de urgência, 33% de trotes e o restante divide-se em enganos, informações, etc.), torna necessário o estudo do dimensionamento do tronco telefônico a fim de racionalizar o atendimento.

Um estudo procurando estabelecer a programação diária das equipes (definição dos turnos de trabalho e pessoal escalado para cada turno), respeitando-se as restrições legais e de recursos do sistema, também se faz necessário.

Pesquisas com a aplicação da modelagem em outros estudos de casos de SAMU's brasileiros, se possível em outros estados, considerando cenários de médias e pequenas cidades, a fim de analisar se os resultados aqui encontrados de fato refletem a realidade de outras regiões brasileiras, também seriam interessantes.

ANEXO A

Caracterização de um sistema de filas

DEFINIÇÃO

A fila é caracterizada por um processo de chegadas a um sistema de atendimento formado por uma ou mais unidades de serviço, no qual os usuários podem sofrer esperas ou não pelo recebimento deste atendimento.

Sistemas geradores de filas constituem parte da ampla classe de sistemas dinâmicos denominados sistemas de fluxos, uma vez que a chegada de usuários, seu processamento e suas respectivas partidas provocam uma movimentação dentro do sistema, dando origem, assim, aos chamados fluxos de entrada e de saída do sistema. Estes fluxos de usuários podem ser discretos ou contínuos.

COMPONENTES DE UM SISTEMA DE FILAS

Os principais componentes de um sistema de filas são: a população de usuários (ou fonte) e os servidores (ou canais de atendimento). Quanto à população pode ser finita ou infinita. Sistemas com população infinita permitem representações analíticas simples; sistemas com população finita, em geral são mais complexos.

Os canais de atendimento representam as facilidades de serviço (servidores) que o sistema oferece. Existem sistemas simples, com apenas um canal de atendimento e sistemas com múltiplos canais, que podem estar dispostos paralelamente ou em série (de forma

seqüencial), ou ainda de forma combinada - série/paralelo, o que dá origem às chamadas redes de filas (*queueing networks*).

Outro componente importante é a sala de espera de um sistema, cuja capacidade pode ser limitada ou não. Alguns sistemas de filas são fisicamente limitados; sendo assim, um usuário que chega ao sistema e encontra todos os servidores ocupados, somente espera pelo atendimento quando a sala de espera não está cheia; outros, porém, possuem capacidade nula, e portanto, quando um usuário chega e encontra todos os servidores ocupados não consegue entrar para esperar pelo atendimento. Existem ainda sistemas em que esta capacidade é infinita; nestes casos, todo usuário que chega ao sistema pode esperar para receber o atendimento.

CARACTERÍSTICAS BÁSICAS

Destacam-se, como características básicas de um sistema de filas, o padrão das chegadas, o padrão de atendimento e a disciplina da fila, além do número de canais de atendimento e também o tamanho da população.

O padrão de chegadas reflete a maneira com a qual os usuários entram no sistema. Quantitativamente pode ser representado pelo número médio de chegadas por unidade de tempo ou pelo tempo médio entre chegadas sucessivas. Nos sistemas em que o fluxo de chegadas é de natureza determinística, isto é, não ocorrem oscilações aleatórias no entorno do valor médio, o padrão de chegadas é totalmente descrito por estes valores. Por outro lado, se o sistema é de natureza estocástica (o que ocorre em grande parte dos fenômenos reais), estes valores fornecem somente medidas de tendência central do processo de chegadas, e para melhor caracterizá-lo é necessário trabalhar com a distribuição de probabilidades associada ao processo (ALLEN, 1978).

O padrão de atendimento é em geral expresso em função do número médio de usuários atendidos pelo sistema por unidade de tempo, ou equivalentemente, pelo tempo médio de atendimento dos usuários. Aqui também diferenciam-se sistemas cujo processo

de atendimento é de natureza determinística e sistemas cujos atendimentos constituem um processo estocástico que necessitam, além do valor médio de usuários atendidos por unidade de tempo, a distribuição de probabilidades que rege o processo.

A disciplina da fila é a regra que seleciona o próximo usuário a ser atendido no sistema. As mais comuns são: FCFS (*first-come, first-served*), ou seja, o primeiro usuário a entrar no sistema será o primeiro a receber o atendimento; LCFS (*last-come, first-served*), isto é, o último a entrar será o primeiro a receber o atendimento; SIRO (*service in random order*), sistema com atendimento aleatório, e PRI (*priority service*), ou atendimento com prioridade.

NOTAÇÃO USUAL

David Kendall (1953) propôs uma notação para facilitar a descrição de um sistema de filas considerando algumas de suas características:

$$X / Y / c$$

onde:

X representa as características do processo das chegadas;

Y representa as características do processo de atendimento; e

c o número de canais de atendimento oferecidos pelo sistema (ou servidores, ou ainda, facilidades de serviço).

Dentre os valores que X e Y podem assumir, destacam-se: D - processo determinístico, M - processo de Markov, E_k - distribuição dos tempos de chegadas ou atendimentos seguindo o padrão da distribuição de Erlang de ordem $k > 0$ e G - processo genérico.

Em 1971, a notação de Kendall foi estendida a fim de incorporar os demais parâmetros que caracterizam um sistema:

$$X^\alpha / Y^{\beta,\gamma} / c / m / N / \text{disciplina}$$

onde:

m representa a capacidade máxima da sala de espera;

N representa o tamanho da população;

X^α descreve o processo de chegadas em grupos de no máximo α usuários; e

$Y^{\beta,\gamma}$ representa processo de atendimento em grupos de no máximo β e no mínimo γ usuários.

PRINCIPAIS MODELOS

A literatura apresenta modelos analíticos para representar sistemas geradores de filas. Maiores detalhes sobre o assunto podem ser encontrados em GROSS & HARRIS (1974) e KLEINROCK (1975). Para que a modelagem seja válida, admite-se que os sistemas estejam em equilíbrio.

Esta breve revisão apresenta apenas os principais modelos que fornecem informações para a modelagem do hipercubo apresentada no Capítulo 3.

O modelo M/M/1

Este modelo representa um sistema de fila onde o processo de chegada é Markoviano, ou seja, o número de chegadas por intervalo de tempo tem distribuição de Poisson, o processo de atendimento também é Markoviano, isto é, a distribuição dos tempos de atendimento é exponencial negativa, e possui apenas 1 servidor. A sala de espera tem capacidade infinita, assim como a população de usuários também é infinita, e a disciplina é FCFS. Supondo que no equilíbrio a taxa média de chegadas seja igual a λ usuários por unidade de tempo, e a taxa média de atendimento seja igual a μ usuários por intervalo de tempo, determina-se o índice de congestionamento para o sistema, ρ ($\rho = \lambda/\mu$), que neste caso coincide com o fator de utilização do servidor, λ/μ . Este valor deve ser estritamente menor que a unidade. Caso contrário, diz-se que o sistema está congestionado.

O diagrama de fluxos para este caso está representado na Figura A.1.

FIGURA A.1 – DIAGRAMA DE FLUXOS PARA O MODELO M/M/1

Com isto, pode-se calcular a probabilidade de haver exatamente k usuários no sistema, p_k :

$$p_k = \frac{\tilde{\rho}^k}{k!} \cdot p_0 = \tilde{n}^k \cdot p_0$$

$$\text{onde } p_0 = \frac{1}{1 + \sum_{k=1}^{\infty} \rho^k}.$$

Medidas de desempenho como número médio de usuários em fila, tempo médio de espera, número médio de usuários no sistema, tempo médio de permanência no sistema, etc, podem ser calculadas com base em p_k , considerando $\sum_{k=0}^{\infty} p_k = 1$. Detalhes sobre estes indicadores podem ser obtidos em GROSS & HARRIS (1974) e KLEINROCK (1975).

O modelo M/M/c

Similar ao modelo M/M/1, o modelo M/M/c representa um sistema de filas com chegadas e atendimentos Markovianos, com c servidores idênticos e em paralelo. Nesta situação, o índice de congestionamento ρ passa a ser $\tilde{n} = \frac{\tilde{\rho}}{1 \cdot c}$, e mantém-se a restrição de que ρ seja estritamente menor que a unidade. O diagrama de fluxos para este caso pode ser visto na Figura A.2.

FIGURA A.2 – DIAGRAMA DE FLUXOS PARA O MODELO M/M/C

Neste caso, $\lambda_k = \begin{cases} k \cdot \lambda, & 1 \leq k < c \\ c \cdot \lambda, & k \geq c \end{cases}$ e p_k terá dois comportamentos:

$$p_k = \begin{cases} \frac{(c \cdot \tilde{n})^k}{k!} \cdot p_0, & 1 \leq k < c \\ \frac{(c \cdot \tilde{n})^c}{c!} \cdot p_0, & k \geq c \end{cases}$$

$$\text{onde } p_0 = \left[1 + \sum_{k=1}^{c-1} \frac{(c \cdot \tilde{n})^k}{k!} + \sum_{k=c}^{\infty} \frac{(c \cdot \tilde{n})^k}{c!} \cdot \frac{1}{c^{k-c}} \right]^{-1}.$$

Analogamente ao modelo M/M/1, as medidas de desempenho são obtidas com base em p_k (GROSS & HARRIS, 1974, KLEINROCK, 1975).

O modelo M/M/c/K

Este modelo representa um sistema com chegadas e atendimentos Markovianos, c servidores e sala de espera com capacidade limitada em K usuários ($K \geq c$).

Neste caso, $\lambda_k = \begin{cases} \lambda, & 0 \leq k < K \\ 0, & k \geq K \end{cases}$ e $\lambda_k = \begin{cases} k \cdot \lambda, & 1 \leq k < c \\ c \cdot \lambda, & c \leq k \leq K \end{cases}$.

$$\text{Assim, } p_k = \begin{cases} \frac{1}{k!} \cdot \left(\frac{\tilde{n}}{\lambda} \right)^k \cdot p_0, & 1 \leq k < c \\ \frac{1}{c^{k-c} \cdot c!} \cdot \left(\frac{\tilde{n}}{\lambda} \right)^k \cdot p_0, & c \leq k \leq K \end{cases}$$

ANEXO B

Questionário aplicado nos serviços de emergência

PRIMEIRA ETAPA DA PESQUISA DE CAMPO – SISTEMA DE TRANSPORTE EMERGENCIAL DE SAÚDE

Identificação do sistema
Nome do serviço: _____
Localização: _____
Rua/Avenida: _____ n.º _____
Bairro: _____ Cidade: _____
Fone/Fax: _____ CEP: _____

Chamados	
Número de chamados/dia: _____ Período de operação do sistema: _____	
1. Existe uma central telefônica para onde convergem todos os chamados?	
☐ S Especificar: _____ _____	☐ N Como chegam os chamados? _____ _____
2. Existe sistema de apoio?	
☐ S Qual? _____ Como funciona? _____ _____	☐ N Como ficam os chamados se todos os veículos estão atendendo? _____ _____
3. O sistema permite que uma vítima fique esperando (formação de fila)?	
☐ S Especificar (fila única central ou em cada base) _____ _____	☐ N Chamados são perdidos? _____ _____
4. Existem critérios de prioridade, de acordo com o tipo da urgência?	
☐ S Qual? _____ Por que este critério? _____	☐ N Por que não admitir um critério? _____ _____
5. Observa-se períodos e locais de maior solicitação pelo serviço?	
☐ S Qual (horário, dia da semana, região, etc.)? _____ _____	☐ N O serviço pode ser considerado homogêneo? _____ Por quê? _____

Atendimentos	
Número de veículos na frota: _____ Tempo médio que o veículo permanece ocupado: _____ Tempo médio de viagem: _____	Período de operação dos veículos: _____ Tempo médio de espera de uma vítima: _____ Tempo médio de atendimento: _____
1. Os veículos ficam localizados em uma base central?	
☐ S Todos os atendimentos partem desta central? _____	☐ N Onde se localizam? _____ _____
2. Existe sistema de comunicação entre veículo e base?	
☐ S Qual? _____	☐ N Por quê? _____
3. Existem diferenças entre as características operacionais e funcionais dos veículos?	
☐ S Especificar: _____ _____	☐ N Os veículos são idênticos? _____ _____
4. Presença do operador de frota...	
☐ S Descrever a função: _____ _____	☐ N Como funcionam as alocações de serviço? _____
5. Presença de regulação médica...	
☐ S Descrever a função: _____ _____	☐ N Existe algum outro serviço regulador? _____

Avaliação
1. Definição de serviço de assistência médica de urgência...

2. Parâmetros para avaliar o nível de serviço que está sendo oferecido...

3. Existência de avaliações periódicas a fim de encontrar os erros e acertos do sistema...

4. Existência de estudos para justificar a compra de veículos e equipamentos, e elevar a qualificação profissional da equipe...

Avaliação (atribuindo uma nota de 0 a 10) para o seu sistema...

ANEXO C

Definição do período de pico

ANEXO D

Testes de aderência e análise de variância

PROCESSO DE CHEGADAS

TABELA D.1 – Teste *K-S* para os intervalos entre chegadas sucessivas

Classe	x_i	# observações	$G(x_i)$	$F(x_i)$	$ F(x_i) - G(x_i) $
(0, 5]	2,5	61	0,2919	0,2017	0,0902
(5, 10]	7,5	57	0,5646	0,4912	0,0734
(10, 15]	12,5	40	0,7560	0,6757	0,0803
(15, 20]	17,5	16	0,8325	0,7933	0,0392
(20, 25]	22,5	10	0,8804	0,8683	0,0121
(25, 30]	27,5	7	0,9139	0,9160	0,0022
(30, 35]	32,5	5	0,9378	0,9465	0,0087
(35, 40]	37,5	4	0,9569	0,9659	0,0090
(40, 45]	42,5	3	0,9713	0,9783	0,0070
(45, 50]	47,5	2	0,9809	0,9861	0,0053
(50, 55]	52,5	3	0,9952	0,9912	0,0040
(55, 60]	57,5	1	1,0000	0,9944	0,0056

Onde $G(x_i)$ são as frequências acumuladas observadas e $F(x_i)$ a função distribuição de probabilidade exponencial; a média amostral dos intervalos entre chegadas sucessivas foi considerada igual a 1,11 minutos.

O valor crítico para o teste, ao nível de significância $\alpha = 5\%$ é de $1,36/\sqrt{209} \cong 0,0941 > 0,0902 = \max |F(x_i) - G(x_i)|$ (valor obtido na amostra).

PROCESSO DE ATENDIMENTOS

TABELA D.2 – Teste *K-S* para os tempos de atendimento ($\alpha = 5\%$)

Ambulância	Tempo médio de atendimento (*)	Desvio padrão (*)	# observações	Valor crítico	Valor obtido na amostra
1	61	40	8	0,454	0,402
2	64	44	11	0,391	0,390
3	63	52	26	0,259	0,208
4	70	48	28	0,251	0,216
5	61	44	23	0,276	0,224
6	66	58	25	0,264	0,205
7	64	50	24	0,270	0,233
8	61	40	22	0,282	0,277
9	70	58	21	0,288	0,256
10	70	56	21	0,288	0,243

(*) minutos

Em todos os casos, o valor obtido com a amostra apresentou-se menor que os valores críticos tabelados para o teste, validando a hipótese de tempos de atendimentos exponencialmente distribuídos.

ANÁLISE DE VARIÂNCIA

O objetivo da Análise de Variância é identificar diferenças significativas entre as médias populacionais devidas às várias causas atuando simultaneamente sobre os elementos de uma população.

O teste aplicado considerou amostras de tamanhos diferentes, e quantificou as principais fontes de fontes de variação para a variável tempo médio de atendimento do SAMU-192. Os resultados estão apresentados no quadro a seguir.

TABELA D.3 – Quadro de ANOVA, ao nível de significância $\alpha = 5\%$

Fonte da variação	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F</i>	<i>F crítico</i>
Entre amostras	5.756,54	9	639,62	2,27	1,92
Residual	56.058,82	199	281,70		
Total	61.815,36	208			

Note que $F > F_{crítico}$, representando um alto grau de variabilidade na amostra. Portanto, não se pode considerar tempos médios de atendimento das ambulâncias do SAMU são estatisticamente iguais.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, L. M.; GONÇALVES, M. B. (1996). Um estudo sobre modelos de localização e alocação e critérios de equidade para os serviços públicos. In: CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTES, 10, Brasília, 1996. *Anais*. Brasília, ANPET. v. 1, p. 189 – 97.
- ANDERSON, L. R.; FONTENOT, R. A. (1992). Optimal positioning of service units along a coordinate line. *Transportation Science*, v. 26, n. 4, p. 346 – 51.
- ALBINO, J. C. C. (1994). *Quantificação e locação de unidades móveis de atendimento de emergência a interrupções em redes de distribuição de energia elétrica: aplicação do modelo hipercubo*. Florianópolis. 66p. Dissertação (Mestrado) - Universidade Federal de Santa Catarina.
- ALLEN, A. O. (1978). *Probability, statistics, and queueing theory with computer science applications*. New York, Academic Press.
- ATKINSON (1978). *An introduction to numerical analysis*. New York, John Wiley.
- AUBIN, J. (1992). Scheduling ambulances. *Interfaces*, v. 22, n. 2, p. 1 – 10.
- BATTA, R.; DOLAN, J. M.; KRISHNAMURTHY, N. N. (1989). The maximal expected covering location problem: revisited. *Transportation Science*, v. 23, n. 4, p. 277 – 87.
- BATTA, R.; MANNUR, N. R. (1990). Covering-location models for emergency situations that require multiple response units. *Management Science*, v. 36, n. 1, p. 16 – 23.
- BEAUMONT, J. R. (1987). Location – allocation models and central place theory. In: GOSH, A.; RUSHTON, G., (eds). *Spatial Analysis and Location Allocation Models*, 6, Van Nostrand Reinhold Company, p. 24-54.
- BELARDO, S.; HARRALD, J.; WALLACE, W.; WARD, J. (1984). A partial covering approach to siting response resources for major maritime oil spills. *Management Science*, v. 30, p. 1184-96.

- BELL, C.; ALLEN, D. (1969). Optimal planning of an emergency ambulance service. *Journal of Socio-Economic Planning Science*, v. 3, n. 2, p. 95 – 101.
- BELTRAMI, E. J. (1977). *Models for public systems analysis*. New York, Academic Press.
- BERMAN, O.; LARSON, R. C. (1982). The median problem with congestion. *Computers and Operations Research*, v. 9, n. 2, p. 119 – 26.
- BERMAN, O.; LARSON, R. C.; PARKAN, C. (1987). The stochastic queue p-median problem. *Transportation Science*, v. 21, p. 207-16.
- BOX, G. E. P.; HUNTER, W. G.; HUNTER, J. S. (1978). *Statistics for experimenters*. New York, John Wiley & Sons.
- BRANDEAU, M.; LARSON, R. C. (1986). Extending and applying the hypercube queueing model to deploy ambulances in Boston. In: SWERSEY, A. J. and INGALL, E. J., (eds). *Delivery of Urban Services. TIMS Studies in the Management Science*, 22, Elsevier, p. 121 – 53.
- BURWELL, T. H., JARVIS J. P.; MCKNEW, M. A. (1993). Modeling co-located servers and dispatch ties in the hypercube model. *Journal of Computers and Operations Research*, v. 20, n. 2, p. 113 – 120.
- CHAIKEN, J. (1971). *The number of emergency units busy at alarms with require multiple servers*. Santa Monica, CA, The Rand Corporation (R-531-NYC/HUD).
- CHAIKEN, J. M.; DORMONT, P. (1978a). A patrol car allocation model: background. *Management Science*, v. 24, n. 12, p. 1280-90.
- CHAIKEN, J. M.; DORMONT, P. (1978b). A patrol car allocation model: capabilities and algorithms. *Management Science*, v. 24, p. 1291 – 300.
- CHAIKEN, J. M.; LARSON, R. C. (1972). Methods for allocating urban emergency units: a survey. *Management Science*, v. 19, n. 4, p. P-110 - P-130.
- CHAPRA, S. C.; CANALE, R. P. (1990). *Numerical methods for engineers*. New York, McGraw-Hill, Inc.
- CHELST, K. R.; BARLACH, Z. (1981). Multiple unit dispatches in emergency services: models to estimate system performance. *Management Science*, v. 27, n. 12, p. 1390 – 409.

- CHELST, K. R.; JARVIS, J. P. (1979). Estimating the probability distribution of travel times for urban emergency service systems. *Operations Research*, v. 27, n. 1, p. 199 – 204.
- CHIYOSHI, F.; GALVÃO, R. D.; MORABITO, R. (2000). O uso do modelo hipercubo na solução de problemas de localização probabilísticos. *Gestão & Produção*, v. 7, n. 2, p.146-74.
- CHRISTOFIDES, N. (1975). *Graph theory – an algorithm approach*. London, Academic Press.
- CHURCH, R.; REVELLE, C.(1974). The maximal covering location problem. *Papers Regional Science Association*, v. 32, p. 101-120.
- DASKIN, M. S. (1983). A maximum expected covering location model: formulation, properties and heuristic solution. *Transportation Science*, v. 17, n. 1, p. 48 – 70.
- DASKIN, M. S.; STERN, E. H. (1981). A hierarchical objective set covering model for emergency medical service vehicle deployment. *Transportation Science*, v. 15, p. 137 – 152.
- EATON, D. J.; DASKIN, M. S; SIMMONS, D.; BULLOCH, B.; JANSMA, G. (1985). Determining emergency medical service vehicle deployment in Austin, Texas. *Interfaces*, v. 15, n. 1, p. 96 - 108.
- FITZSIMMONS, J. A. (1973). A methodology for emergency ambulance deployment. *Management Science*, v. 19, n. 6, p. 627 – 636.
- FUJIWARA, O.; MAKJAMROEN, T.; GUPTA, K. K. (1987). Ambulance deployment analysis: a case study of Bangkok. *European Journal of Operational Research*, 31, p. 9 – 18.
- GALVÃO, R. D. (1980). A dual-bounded algorithm for the p-median problem. *Operations Research*, v. 28, n. 5, p. 1112-21.
- GALVÃO, R. D. (1993). The use of Lagrangean relaxation in the solution of uncapacitated facility location problems. *Location Science*, v. 1, p. 57-79.
- GALVÃO, R. D.; RAGGI, L. A. (1989). A method for solving to optimality uncapacitated location problems. *Annals of Operations Research*, v. 18, p. 225-244.
- GALVÃO, R. D.; REVELLE, C. S.(1996). A Lagrangean heuristic for the maximal covering location problem. *European Journal of Operational Research*, v. 88, p. 114-23.

- GOLDBERG, J.; DIETRICH, R.; CHEN, J.; MITWASI, M.; VALENZUELA, T.; CRISS, E. (1990). A simulation model for evaluating a set of emergency vehicle base locations: development validation and usage. *Journal of Socio-Economic Planning Science*, 24, p. 125 – 141.
- GONÇALVES, M. B. (1994). Métodos de P. O. em serviços emergenciais. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 26., Florianópolis, SC, 1994. *Anais*. Florianópolis, v. 1, p. 597 – 601.
- GONÇALVES, M. B.; NOVAES, A. G.; ALBINO, J. C. C. (1994). Modelos para localização de serviços emergenciais em rodovias. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 26., Florianópolis, SC, 1994. *Anais*. Florianópolis, v. 1, p. 591 – 596.
- GONÇALVES, M. B.; NOVAES, A. G.; SCHMITZ, R. (1995). Um modelo de otimização para localizar unidades de serviços emergenciais em rodovias. In: CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTES, 9., São Carlos, SP, 1995. *Anais*. São Carlos, v. 3, p. 962 – 972.
- GREEN, L. (1984). A multiple dispatches queueing model of police patrol operations. *Management Science*, v. 30, n. 6, p. 653 – 664.
- GREEN, L.; KOLESAR, P. (1984). A comparison of the multiple dispatch and M/M/C priority queueing models of police patrol. *Management Science*, v. 30, n. 6, p. 665 – 670.
- GROSS, D.; HARRIS, C. M. (1974). *Fundamentals of queueing theory*. New York, John Wiley & Sons.
- HALPERN, J. (1977). The accuracy of estimates for the performance criteria in certain emergency service queueing systems. *Transportation Science*, v. 11, n. 3, p. 223 – 242.
- HEMPEL, C. G. (1974). *Filosofia da ciência natural*. Rio de Janeiro, Zahar.
- HODGART, R. L. (1980). Optimizing access to public services: a review of problems, models and methods of locating central facilities. *Geographical Analysis*, v. 12, p. 17-48.
- HOGG, J. M. (1968). The siting of fire stations. *Operational Research Quarterly*, v. 19, n. 3, p. 275 – 287.
- INGALL, E. J.; KOLESAR, P.; SWERSEY, A.; WALKER, W.; BLUM, E.; CARTER, G.; BISHOP, H. (1975). Improving the deployment of New York City fire companies. *Interfaces*, v. 2, n. 2, p. 46-8.

- JARVINEN, P.; RAJALA, J.; SINERVO, H. (1972). A branch-and-bound algorithm for seeking the p-median. *Operations Research*, v. 20, n. 10, p. 173-8.
- JARVIS, J. (1975). Optimization in stochastic systems with distinguishable servers. *Technical Report*, p. 19-75.
- KAPLAN, E. H. (1979). Evaluating the effectiveness of one-officer versus two officer patrol units. *Journal of Criminal Justice*, v. 7, p. 325-55.
- KEENEY, R. L. (1972). A method for districting among facilities. *Operations Research*, v. 20, n. 3, p. 613 – 618.
- KLEINROCK, L. (1975). *Queueing systems, volume I: theory*. New York, John Wiley & Sons.
- KOLESAR, P.; SWERSEY, A. J. (1986). The deployment of urban emergency units: a survey. In: A. Swersey and E. Ingall (eds), *Delivery of Urban Services. TIMS Studies in the Management Science*, 22, Elsevier, p. 87 – 119.
- KOLESAR, P.; BLUM, E. (1973). Square root law for fire engines response distances. *Management Science*, v. 19, n. 12, p. 1368-1378.
- LARSON, R. C. (1972). *Urban police patrol analysis*. Cambridge, MIT Press.
- LARSON, R. C. (1974c). Illustrative police sector redesign in district 4 in Boston. *Urban Analysis*, 2, p. 51 – 91.
- LARSON, R. C. (1974b). Resource planning for urban public safety systems. *Technology Review*, v. 76, n. 7, p. 20 – 29.
- LARSON, R. C. (1974a). A hypercube queueing model for facility location and redistricting in urban emergency services. *Computers and Operations Research*, 1, p. 67 – 95.
- LARSON, R. C. (1975). Approximating the performance of urban emergency service systems. *Operations Research*, v. 23, n. 5, p. 845 – 868.
- LARSON, R. C. (1983). Social-systems and the public-sector. *Operations Research*, v. 31, n. 1, p. 22 – 22.
- LARSON, R. C.; ODoni, A. R. (1981). *Urban Operations Research*. New Jersey, Prentice-Hall.
- LARSON, R. C.; RICH, T. F. (1987). Travel-time analysis of New York City police patrol cars. *Interfaces*, v. 17, n. 2, p. 15 – 20.

- LARSON, R. C.; LI, V. O. K. (1981). Finding minimum rectilinear distance paths in the presence of barriers. *Networks*, v. 11, n. 3, p. 285 – 304.
- LIMA, F. S; GONÇALVES, M. B. (1999). Logística de serviços públicos: uma aplicação à distribuição espacial de serviços especializados de saúde. In: CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTES, 13., São Carlos, SP, 1999. *Anais*. São Carlos, v. 1, p. 482 – 493.
- LOUVEAUX, F. (1993). Stochastic location analysis. *Location Science*, v. 1, p. 127 – 154.
- MAYER, D. J. (1979). Paramedic response time and survival from cardiac arrest. *Social Science*, v. 13, n. 4, p. 267-71.
- MENDONÇA, F. C. (1999). *Aplicação do modelo hipercubo baseado em teoria de filas, para análise de um sistema médico-emergencial em rodovia*. São Carlos. 112p. Dissertação (Mestrado) - Universidade Federal de São Carlos.
- MENDONÇA, F. C.; MORABITO, R. (2000). Aplicação do modelo hipercubo para análise de um sistema médico-emergencial em rodovia. *Gestão & Produção*, v. 7, n. 1, p.73-91
- MIRCHANDANI, P. B.; ODONI, A. R. (1979). Locations of medians on stochastic network. *Transportation Science*, v. 13, p. 85-97.
- NOVAES, A. G. (1989). *Sistemas logísticos: transporte, armazenagem e distribuição física de produtos*. São Paulo, Edgard Blücher Ltda.
- OWEN, S. H.; DASKIN, M. S. (1998). Strategic facility location: a review. *European Journal of Operational Research*, v. III, n. 3, p. 423 – 447.
- PIRKUL, H.; SCHILLING, D. A. (1989). The capacitated maximal covering location problem with backup service. *Annals of Operations Research*, v. 18, p. 141-154.
- PLANE, D.; HENDRICK, T. (1977). Mathematical programming and the location of fire companies for the Denver fire department. *Operations Research*, v. 25, n. 4, p. 563-78.
- REVELLE, C.; HOGAN, K. (1989). The maximum availability location problem. *Transportation Science*, v. 23, n. 3, p. 192 – 200.
- REVELLE, C.; SWAIN, R. (1970). Central facilities location. *Geographical Analysis*, v. 2, p. 30 – 42.

- RICHARDSON, A. J.; AMPT, E. S.; MEYBURG, A. H. (1995). *Survey methods for transport planning*. Parkville, Eucalyptus Press.
- SAVAS, E. S. (1969). Simulation and cost-effectiveness analysis of New York's emergency ambulance service. *Management Science*, v. 15, n. 12, p. B-608 - B-627.
- SAYDAM, C.; REPEDE, J. F.; BURWELL, T. (1994). Accurate estimation of expected coverage – a comparative study. *Socio-Economic Planning Science*, v. 28, p. 113-20.
- SCHREUDER, J. A. M. (1981). Application of a location model to fire stations in Rotterdam. *European Journal of Operational Research*, v. 6, p. 212-9.
- SILVA, A. N. R. (1993). *O custo do solo urbano e uma nova sistemática de distribuição da propriedade*. São Carlos. 137p. Tese (Doutorado) - Escola de Engenharia de São Carlos, Universidade de São Paulo.
- SOUZA, J. C., NOVAES, A. G.; GONÇALVES, M. B. (1998). Localização de unidades de atendimento médico em rodovias. In: CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTES, 12., Fortaleza, CE, 1998. *Anais*. Fortaleza, v. 1, p. 387 – 398.
- STRAMBI, O. (1997). Survey methods for transport planning. *Transportes*, v. 5, n. 1, p. 105 – 111.
- SWAIN, R. (1971). *A decomposition algorithm for a class of facility location problems*. Ithaca,. PhD. Thesis, Cornell University.
- SWERSEY, A. (1982). A Markovian decision model for deciding how many fire companies to dispatch. *Management Science*, v. 28, n. 4, p. 352-65.
- SWERSEY, A. J. (1994). The deployment of police, fire, and emergency medical units. *Handbooks in OR & MS*, S. M. Pollock, et al., v. 6, p. 151 – 200.
- SWERSEY, A. J.; INGALL, E. J. (1986) *Delivery of urban services*. New York, Elsevier Science Publishing Company.
- SWOVELAND, C.; UYENO, D.; VERTINSKY, I.; VICKSON, R. (1973). Ambulance location: a probabilistic enumeration approach. *Management Science*, v. 20, n. 4, p. 686 – 698.
- TAYLOR, I. D. S.; TEMPLETON J. G. C. (1980). Waiting time in a multi-server cutoff-priority queue, and its application to an urban ambulance service. *Operations Research*, v. 28, n. 5, p. 1168 – 1188.

- TAYLOR, P. E.; HUXLEY, S. J. (1989). A break from tradition for San Francisco police: patrol officer scheduling using an optimization-based decision support system. *Interfaces*, v. 19, n. 1, p. 4-24.
- TIEN, J.; CHIU, S. (1985). Impact analysis of some cost reduction options for a metropolitan fire department. *Fire Technology*, v. 21, n. 1, p. 5-21.
- TORGAS, C., R. SWAIN, C. REVELLE; L. BERGMAN (1971). The location of emergency service facilities. *Operations Research*, v. 19, p. 1363 – 1373.
- TRUDEAU, P., J-M ROUSSEAU, J. A. FERLAN; J. CHOQUETTE (1989). An operations research approach for the planning and operation of an ambulance service. *Infor*, v. 27, n. 1, p. 95 – 113.
- VOLZ, R. A. (1971). Optimum ambulance location in semi-rural areas. *Transportation Science*, v. 5, n. 2, p. 193 – 203.
- WALKER, W. (1974). Using the set covering problem to assign fire companies to fire houses. *Operations Research*, v. 20, n. 3, p. 275-7.
- WEAVER, J. R.; CHURCH, R. L. (1983). Computational procedures for location problems on stochastic networks. *Transportation Science*, v. 17, p. 167-80.
- WEAVER, J. R. ; CHURCH, R. L. (1985). A median location model with nonclosest facility service. *Transportation Science*, v. 19, p. 58-71.
- WHITE, J. A.; CASE, K. E. (1974). On covering problems and the central facilities location problem. *Geographical Analysis*, v. 6, p. 281-93.
- WORLD HEALTH ORGANIZATION (2000). *The world health report 2000: health systems – improving performance*. Geneva, Switzerland. World Health Organization, 1211.
- WOLFF, R. W. (1989). *Stochastic modeling and theory of queues*. New Jersey, Prentice-Hall.

BIBLIOGRAFIA CONSULTADA

- ALI, A.; WHITE, J. (1978). Probabilistic formulation of the emergency service location problem. *Journal of Operational Research Society*, v. 17, p. 1167 – 1179.
- AVELLA, P.; BENATI, S.; MARTINEZ, L. C. (1998). Some personal views on the current state and the future of locational analysis. *European Journal of Operational Research*, v.104, 2, p. 269 – 287.
- BADRI, M. A.; MORTAGY, A. K.; ALSAYED, A. (1998). A multi-objective model for locating fire stations. *European Journal of Operational Research*, v. 111, n. 2, p. 243 – 260.
- BAKER, J. R.; CALYTON, E. R.; MOORE, L. J. (1989). Redesign of primary areas for county ambulance services. *European Journal of Operational Research*, v. 41, p. 23 – 32.
- BAKER, J. R.; FITTPATRICK, K. E. (1986). Determination of an optimal forecast model for ambulance demand using goal programming. *Journal of Operations Research Society*, v. 37, 11, p. 1047 – 1059.
- BALL, M. O.; LIN, F. L. (1993). A reliability model applied to emergency service vehicle location. *Operations Research*, v.41, 1, p. 18 – 36.
- BATTA, R.; LARSON, R. C.; ODONI, A. R. (1988). A single-server priority queueing-location model. *Networks*, v. 18, n. 2, p. 87-103.
- BERLIN, G.; LIEBMAN, J. (1974). Mathematical analysis of emergency ambulance location. *Journal of Socio-Economic Planning Science*, v. 8, p. 323 – 388.
- BERMAN, O.; LARSON, R. C.; FOUSKA, N. (1992). Optimal location of discretionary service facilities. *Transportation Science*, v. 26, n. 3, p. 201–211.
- BERMAN, O.; LARSON, R. C.; CHIU, S. S. (1985). Optimal server location on a network operating as an M/G/1 queue. *Operations Research*, v. 33, n. 4, p. 746 – 771.
- BRANDEAU, M.; CHIU, S. (1992). A center location problem with congestion. *Annals of Operations Research*, n. 40, p. 17 – 31.

- BRODSKY, H. (1992). Emergency medical service rescue time in fatal road accidents. *Transportation Research Record*, v. 1270, p. 89 – 96.
- CARRIZOSA, E.; CONDE, E.; MUNOZ-MARQUEZ, M. (1998). Admission policies in loss queueing models with heterogeneous arrivals. *Management Science*, v. 44, n. 3, p. 311 – 320.
- CARSON, Y. M.; BATTA, R. (1990). Locating an ambulance on the Amherst campus of the State University of New York at Buffalo. *Interfaces*, v. 20, n. 5, p. 43 – 49.
- CARTER, G. M.; CHAIKEN, J. M.; INGALL, E. (1972). Response areas for two emergency units. *Operations Research*, v. 20, p. 571 – 594.
- CARTER, G.; INGALL, E. (1970). A simulation model of fire department operations. *IEEE System Science and Cybernetics*, v. 6, p. 282 – 293.
- CHARNES, A.; STORBECK, J. (1980). A goal programming model for the siting of multilevel EMS systems. *Socio-Economic Planning Science*, v. 14, p. 155 – 161.
- CHELST, K. (1988). A public safety merger in Grosse Pointe Park, Michigan - a short and sweet study. *Interfaces*, v. 18, n. 4, p. 1 – 11.
- CHIU, S. S.; LARSON, R. C. (1985). Location and normal-server facility in a stochastic environment. *Computers and Operations Research*, v. 12, n. 6, p. 509 – 516.
- COOPER, L. (1963). Location-allocation problems. *Operations Research*, v. 11, p. 331 – 343.
- CRESSON WOOD, C. (1996). A computer emergency response team policy information. *Management and Computer Security*, v. 4, n. 2, p. 4.
- DABERKOW, S. G. (1977). Location and cost of ambulances serving a rural area. *Health Services Research*, v. 12, n. 3, p. 299 – 311.
- DASKIN, M. S. (1982). Application of an expected covering model to EMS system design. *Decision Science*, v. 13, n. 3, p. 416 – 439.
- DASKIN, M. S.; HAGHANI, A. (1984). Multiple vehicle routing and dispatching to an emergency scene. *Environment Planning*, n. A-16, p. 1349 – 1359.
- DEARING, P. M.; JARVIS, J. P. (1978). A location model with queueing constraints. *Journal of Computers and Operations Research*, v. 5, p. 273 – 277.

- DESROCHERS, M.; LENSTRA, J. K.; SAVELSBERGH, M. W. P. (1990). A classification scheme for vehicle routing and scheduling problem. *European Journal of Operational Research*, v. 46, p. 322 – 332.
- EATON, D. J.; HECTOR, M. L.; SANCHEZ, U.; LANTIGUA, R. R.; MORGAN, J. (1986). Determining ambulance deployment in Santo Domingo, Dominican Republic. *Journal of Operations Research Society*, v. 37, n. 2, p. 113 – 26.
- FITZSIMMONS, J. A.; SRIKAR, B. N. (1982). Emergency ambulance location using the contiguous zone search routine. *Journal of Operations Management*, v. 2, n. 4, p. 225 – 237.
- FORTER, S. T.; GORR, W. L. (1986). An adaptive filter for estimating spatially-varying parameters: application to modeling police hours spent in response to calls for service. *Management Science*, v. 32, n. 7, p. 878 – 889.
- FRANCIS, R.; MCGINNIS, L.; WHITE, J. (1983). Location Analysis. *European Journal of Operational Research*, v. 12, p. 220 – 252.
- GOLDBERG, J.; DIETRICH, R.; CHEN, J. M.; MITWASI, M. G.; VALENZUELA, T.; CRISS, E. (1990). Validating and applying a model for locating emergency medical vehicles in Tucson, AZ. *European journal of Operational Research*, v. 49, p. 308 – 324.
- GOLDBERG, J.; SZIDAROVSKY, F. (1991). A general model and convergence results for determining vehicle utilization in emergency systems. *Stochastic Models*, v. 7, p. 137 – 160.
- GOLDBERG, J.; SZIDAROVSKY, F. (1991). Methods for solving nonlinear equations used in evaluating emergency vehicle busy probabilities. *Operations Research*, v. 39, n. 6, p. 903 – 916.
- GOLDBERG, J.; PAZ, L. (1991). Locating emergency vehicle bases when service time depends on call location. *Transportation Science*, v. 24, p. 264 – 280.
- GROOM, K. N. (1977). Planning emergency ambulance services. *Operational Research Quarterly*, v. 28, n. II, p. 641 – 651.
- HAKIMI, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, v. 13, p. 450 – 459.

- HALL, W. K. (1972). The application of multifunction stochastic service systems in allocating ambulances to an urban area. *Operations Research*, v. 20, n. 3, p. 558 – 570.
- HAREL, A.; CHENG, H. (1997). Order statistics applications to queueing and scheduling problems. *Queueing Systems*, v. 27, n. 3, p. 325 – 350.
- INGALL, E. J.; WALKER, W. E. (1977). Deployment of ambulances in Washington, D.C. *Urban Analysis*, v. 4, p. 59 – 92.
- KOLESAR, P. (1975). A model for predicting average fire engine travel times. *Operations Research*, v. 23, n. 4, p. 603 – 613.
- KOLESAR, P.; WALKER, W.; HAUSNER, J. (1975). Determining the relation between fire engine travel times and travel distances in New York City. *Operations Research*, v. 23, n. 4, p. 614 – 627.
- KOLESAR, P.; RIDER, K.; CRABILL, T.; WALKER, W. (1975). A queueing linear programming approach to scheduling police patrol cars. *Operations Research*, v. 23, n. 6, p. 1045 – 1062.
- KOLESAR, P.; WALKER, W. (1974). An algorithm for the dynamic relocation of fire companies. *Operations Research*, v. 22, n. 2, p. 249 – 274.
- LAPORTE, G. (1992). The vehicle routing problem: an overview of exact and approximate algorithms. *European Journal of Operational Research*, v. 59, p. 345 – 358.
- LARSON, R. C. (1975). Overview of development and use of hypercube model for planning in urban emergency services. *Operations Research*, v. 23, Supl. 2, p. B-275 - B-275.
- LARSON, R. C. (1978). Police deployment – introduction. *Management Science*, v. 24, n. 12, p. 1278 – 1279.
- LARSON, R. C. (1981). The process of regionalization - an appropriate conceptual and methodological approach. *Socio Economic Planning Science*, v. 15, n. 5, p. 199 – 205.
- LARSON, R. C. (1987). Perspectives on queues - social-justice and the psychology of queueing. *Operations Research*, v. 35, n. 6, p. 895 – 905.
- LARSON, R. C.; FRANCK, E. (1978). Evaluating dispatching consequences of automatic vehicle location in emergency services. *Journal of Computers and Operations Research*, v. 5, p. 11 – 30.

- LARSON, R. C.; SADIQ, G. (1983). Facility locations with the Manhattan metric in the presence of barriers to travel. *Operations Research*, v. 31, n. 4, p. 652 – 669.
- LARSON, R. C.; STEVENSON, K. A. (1972). On insensitivities in urban redistricting and facility location. *Operations Research*, v. 20, p. 595 – 612.
- LARSON, R. C.; MCNEW, M. A. (1982). Police patrol-initiated activities within a systems queueing model. *Management Science*, v. 28, n. 7, p. 759 – 774.
- LOVE, R. F.; MORRIS, J. G. (1988). *Facilities location: models & methods*. New York, Elsevier Science Publishing.
- LU, X. J.; GERCHAK, Y. (1998). Minimizing the expected response time of an idle server on a line. *IEEE Transactions*, v. 30, n. 4, p. 401 – 408.
- MAÇADA, A. C. G.; BECKER, J. L. (1993). Modelos de transporte aplicado ao planejamento em saúde. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 25., Campinas, SP, 1993. *Anais*. Campinas, v. 1, p. 106-10.
- MARIANOV, V.; REVELLE, C. (1992). The capacitated standard response fire protection siting problem: deterministic and probabilistic models. *Annals of Operations Research*, v. 40, p. 303 – 322.
- MARIANOV, V.; SERRA, D. (1998). Probabilistic, maximal covering location-allocation models for congested systems. *Journal of Regional Science*, v. 38, n. 3, p. 401 – 424.
- NEEBE, A. W. (1988). A procedure for locating emergency-service facilities for all possible response distances. *Journal of Operations Research Society*, v. 39, n. 8, p. 743 – 748.
- NOVAES, A. G.; ALVARENGA, A. C. (1994). *Logística Aplicada: suprimento e distribuição física*. São Paulo, Pioneira.
- PIRKUL, H.; SCHILLING, D. A. (1988). The siting of emergency service facilities with workload capacities and backup service. *Management Science*, v. 34, n. 7, p. 896 – 908.
- REVELLE, C.; MARKS, D.; LIEBMAN, J. (1970). An analysis of private and public sector location models. *Management Science*, v. 16, n. 11, p. 692 – 707.

- SACKS, S. R.; GRIEF, S. (1994). Orlando Police Department uses OR/MS methodology new software to design patrol district. *OR/MS Today*, p. 30 – 32.
- SAVAS, E. S. (1978). On equity in providing public services. *Management Science*, v. 24, n. 8, p. 800 – 808.
- SAYDAM, C.; MCKNEW, M. (1985). A separable programming approach to expected coverage: an application to ambulance location. *Decision Science*, v. 16, p. 381 – 397.
- SCOTT, A. J. (1970). Location-allocation systems: a review. *Geographical Analysis*, v. 2, p. 96 – 119.
- SCOTT, D. W.; FACTOR, L. E.; GORRY, G. A. (1978). Predicting the response time of an urban ambulance system. *Health Services Research*, v. 13, n. 4, p. 404 – 417.
- SILER, K. F. (1979). Evaluation of emergency ambulance characteristics under several criteria. *Health Services Research*, v. 14, n. 2, p. 160 – 176.
- SOUZA, J. C.; NOVAES, A. G.; GONÇALVES, M. B. (1996). Distribuição espacial de equipamentos em sistemas para atendimento de emergência. In: CONGRESSO DE PESQUISA E ENSINO EM TRANSPORTES, 10., Brasília, DF, 1996. *Anais*. Brasília, v. 1, p. 339 – 349.
- SOUZA, J. C.; NOVAES, A. G. (1994). Dimensionamento, localização e escalonamento temporal de serviços de emergência. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 26., Florianópolis, SC, 1994. *Anais*. Florianópolis, v. 1, p. 332 – 37.
- SWERSEY, A. J.; GOLDRING, L.; GEYER, E. (1993). Improving fire department productivity: merging fire and emergency medical unities in New Haven. *Interfaces*, v. 23, n. 1, p. 109 – 129.
- TIGHE, I. (1997). 999 Rescue time. *Computer Bulletin*, v. 9, n. 5, p. 1122 – 24.
- UYENO, D. H.; SEEBERG, C. (1984). A practical methodology for ambulance location. *Simulation*, v. 43, n. 2, p. 79 – 87.
- WATSON, E. F.; CHAWDA, P. P.; MCCARTHY, B. (1998). A simulation metamodel for response-time planning. *Decision Science*, v. 29, n. 1, p. 217 – 241.

APÊNDICE I

Ficha de regulação médica – SAMU de Campinas

PREFEITURA MUNICIPAL DE CAMPINAS

SAMU

SECRETARIA MUNICIPAL DE SAÚDE

192

FICHA DE REGULAÇÃO

01 – IDENTIFICAÇÃO DO CHAMADO

Data: __/__/____ N° da solicitação: _____ Hora da chamada: __:__ h.

02 – IDENTIFICAÇÃO DA VÍTIMA

Nome: _____ Sexo: _____ Idade: _____

Local da ocorrência: _____

Bairro: _____ Fone: _____ - _____ Ponto de referência: _____

Solicitante: _____ Relação com a vítima: _____

Via pública ☐ Domicílio ☐ C.S ☐ _____ P.S ☐ _____ Hospital ☐ _____ Outros ☐ _____

Motivo da chamada: _____

Nome do atendente: _____ Nome do controlador: _____

03 – REGULAÇÃO MÉDICA

Relato do solicitante: _____

Hipótese diagnóstica: _____

Resposta: _____

Remoção simples ☐ Suporte básico ☐ Suporte avançado ☐ Psiquiátrica ☐ Neonatal ☐ Reorientação ☐

Destino da vítima: _____ Médico regulador: _____ CRM: _____

04 - FICHA DO MOTORISTA

Viatura: _____ km de saída: _____ km final: _____

Hora saída: __:__ h. Hora chegada: __:__ h. Hora saída: __:__ h/ Hora chegada: __:__ h.

Hora retorno à base: __:__ h. Intercorrência com o veículo: sim ☐ não ☐ defeito ☐ colisão ☐

Observações: _____

Intercorrência no transporte: sim ☐ não ☐ vítima já socorrida ☐ evadiu-se do local ☐Vítima consciente: sim ☐ não ☐ Respiração: normal ☐ anormal ☐ ausente ☐ Sangramento: sim ☐ não ☐

Atuação do motorista: _____

Observações sobre o paciente: _____

05 – CASOS DE ACIDENTE

Tipo de acidente: atropelamento ☐ carro ☐ moto ☐ outros ☐ _____Posição da vítima: motorista ☐ passageiro ☐ traseiro ☐ dianteiro ☐Uso de equipamento de segurança: cinto ☐ capacete ☐ nenhum ☐ Tempo: bom ☐ chuva ☐ neblina ☐Mão de direção: única ☐ dupla ☐ cruzamento ☐ defesa/barreira ☐ Sinalização: boa ☐ ruim ☐ ausente ☐Tipo de pavimento: asfalto ☐ cascalho ☐ paralelepípedo ☐ terra ☐ concreto ☐ areia ☐

06 - DADOS SOBRE RECEBIMENTO

Destino do paciente: _____ Hora recebimento: __:__ h.

Recebido por: _____ Função: _____ Assinatura _____

07 – CERTIFICADO DE RECUSA

Eu _____, portador do R.G. nº: _____ assumo inteira responsabilidade na recusa do atendimento prestado pelo SAMU – 192 a minha pessoa ou a que previamente acompanhei para este atendimento, mesmo que isto traga danos a minha saúde ou a da pessoa acima relacionada.

DATA: __/__/____ Assinatura: _____

APÊNDICE II

Exemplo de planilha de coleta de dados

Chamado hh:mm	Região origem	Tipo de urgência (*)				Veículo	Saída Base hh:mm	Chegada Local hh:mm	Saída Local hh:mm	Liberação h/pa/ps/... hh:mm	Retorno Base hh:mm
		1	2	3	4						
0 5	3		1			76	0 7	0 14	0 20	0 35	
0 30	4		1			74	0 30	0 58	1 0	1 12	1 40
0 32	4				1	75	0 32	0 40	0 50	1 10	0 35
0 50	3		1			79	0 50	1 10	1 15	1 25	1 50
0 56	1		1			79	0 57	1 10	1 20	1 35	
1 29	1		1			75	1 30	1 40	1 50	2 10	
1 30	2		1			77	1 38	1 47	1 50	2 0	
1 34	4	1				73	1 35	1 40	1 50	2 0	8
1 37	1	1				73	1 39				1 5
1 40	5		1			76	1 40	1 43	1 50	2 0	2 25
2 40	2		1			77	2 40				
3 10	3	1				72	3 12	3 26	3 36	3 41	4 7
3 11	5	1				78	3 14	3 20	3 20	3 30	3 30
3 15	4			1		53	3 20	3 30	3 40	4 0	4 0
4 45	1	1				72	4 46	5 10	5 14	5 15	5 50
5 40	4		1			77	5 42				
6 10	1		1			74	6 12	6 26	6 30	6 43	
7 4	5		1			73	7 12	7 23	7 37		
7 5	4	1				72	7 18	7 40	8 7	8 25	11 25
7 30	2		1			74	7 40	7 50	7 52	8 20	8 30
7 40	5		1			73	7 45	7 50	7 53	8 3	8 13
8 29	3	1				78	8 40	9 5	9 10	9 18	
8 35	4			1		52	8 35	9 0	9 25	9 56	10 15
9 0	4	1				72	9 0	9 8	9 19	9 27	9 35
9 0	4		1			76	9 10	9 19	9 26	9 34	10 13
9 15	3	1				79	9 20	9 40	9 45	9 59	11 30
9 33	2		1			73	9 40	9 50	10 0	10 10	10 20
9 35	4		1			78	9 35	9 47	9 50	10 7	10 35
9 45	3			1		53	9 50	9 58	10 45	11 0	11 25
9 50	2		1			74	10 0	10 7	10 13	10 44	10 47
9 50	1	1				75	9 55	10 0	10 5	10 15	
9 55	5	1				77	9 55	10 10	10 13	10 25	10 35
10 0	4	1				74	10 16	10 23	10 26	10 35	
10 5	5	1				78	11 15	11 28	11 35	12 6	
10 6	2					75	10 41	10 45	10 47	11 3	
10 10	3	1				79	10 15	10 38	10 40	10 56	11 30
10 15	5	1				78	10 20	10 34	10 50	11 50	
10 23	5	1				76	10 38	10 43	10 50	10 57	13 10

(*) 1: VRS

2: VSB

3: VSA

4: PSQ

APÊNDICE III

Exemplos de equações de balanço - sistema com 10 servidores

Neste apêndice apresentam-se as 100 primeiras equações geradas pelo modelo hipercubo para um sistema com 10 servidores.

Notação:

$$\ddot{e}[j] = \ddot{e}_j, j = 1, 2, \dots, N_A$$

$$\dot{i}[n] = \dot{i}_n, n = 1, 2, \dots, N$$

$$p[b_1 b_2 \dots b_N] = p_{b_1 b_2 \dots b_N} = p_B$$

Para todo estado $[b_1 b_2 \dots b_N]$, lê-se: $[B] \cdot p[b_1 b_2 \dots b_N]$, onde $[B]$ representa a soma de todas as possíveis taxas de transição para o estado $[b_1 b_2 \dots b_N]$.

Equação 1 - estado 0000000000

$$(\lambda[2] + \lambda[6]) + (\lambda[4] + \lambda[8] + \lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + (\lambda[1]) + (\lambda[3]) p[0000000000] = \mu[1] p[1000000000] + \mu[2] p[0100000000] + \mu[3] p[0010000000] + \mu[4] p[0001000000] + \mu[5] p[0000100000] + \mu[6] p[0000010000] + \mu[7] p[0000001000] + \mu[8] p[0000000100] + \mu[9] p[0000000010] + \mu[10] p[0000000001]$$

Equação 2 - estado 0000000001

$$(\lambda[2] + \lambda[6]) + (\lambda[4] + \lambda[8] + \lambda[10]) + (\lambda[3] + \lambda[5]) + (\lambda[7]) + (\lambda[9]) + (\lambda[1]) + \mu[10] p[0000000001] = \mu[1] p[1000000001] + \mu[2] p[0100000001] + \mu[3] p[0010000001] + \mu[4] p[0001000001] + \mu[5] p[0000100001] + \mu[6] p[0000010001] + \mu[7] p[0000001001] + \mu[8] p[0000000101] + \mu[9] p[0000000011] + (\lambda[3]) p[0000000000]$$

Equação 3 - estado 0000000010

$$(\lambda[2] + \lambda[6]) + (\lambda[4] + \lambda[8] + \lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[9] + (\lambda[1] + \lambda[3]) p[0000000010] = \mu[1] p[1000000010] + \mu[2] p[0100000010] + \mu[3] p[0010000010] + \mu[4] p[0001000010] + \mu[5] p[0000100010] + \mu[6] p[0000010010] + \mu[7] p[0000001010] + \mu[8] p[0000000110] + (\lambda[1]) p[0000000000] + \mu[10] p[0000000011]$$

Equação 4 - estado 0000000011

$$(\lambda[2] + \lambda[6]) + (\lambda[4] + \lambda[8] + \lambda[10]) + (\lambda[1] + \lambda[3] + \lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[9] + \mu[10] p[0000000011] = \mu[1] p[1000000011] + \mu[2] p[0100000011] + \mu[3] p[0010000011] + \mu[4] p[0001000011] + \mu[5] p[0000100011] + \mu[6] p[0000010011] + \mu[7] p[0000001011] + \mu[8] p[0000000111] + (\lambda[1]) p[0000000001] + (\lambda[1] + \lambda[3]) p[0000000010]$$

Equação 5 - estado 0000000100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[8] + (\lambda[1]) + (\lambda[3]) p[0000000100] = \mu[1] p[1000000100] + \mu[2] p[0100000100] + \mu[3] p[0010000100] + \mu[4] p[0001000100] + \mu[5] p[0000100100] + \mu[6] p[0000010100] + \mu[7] p[0000001100] + \mu[9] p[0000000110] + \mu[10] p[0000000101]$$

Equação 6 - estado 0000000101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[8] + (\lambda[1]) + \mu[10] p[0000000101] = \mu[1] p[1000000101] + \mu[2] p[0100000101] + \mu[3] p[0010000101] + \mu[4] p[0001000101] + \mu[5] p[0000100101] + \mu[6] p[0000010101] + \mu[7] p[0000001101] + \mu[9] p[0000000111] + (\lambda[3]) p[0000000100]$$

Equação 7 - estado 0000000110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) p[0000000110] = \mu[1] p[1000000110] + \mu[2] p[0100000110] + \mu[3] p[0010000110] + \mu[4] p[0001000110] + \mu[5] p[0000100110] + \mu[6] p[0000010110] + \mu[7] p[0000001110] + (\lambda[1]) p[0000000100] + \mu[10] p[0000000111]$$

Equação 8 - estado 0000000111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[8] + \mu[9] + \mu[10] p[0000000111] = \mu[1] p[1000000111] + \mu[2] p[0100000111] + \mu[3] p[0010000111] + \mu[4] p[0001000111] + \mu[5] p[0000100111] + \mu[6] p[0000010111] + \mu[7] p[0000001111] + (\lambda[1]) p[0000000101] + (\lambda[1]+\lambda[3]) p[0000000110]$$

Equação 9 - estado 0000001000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[7] + (\lambda[1]) + (\lambda[3]) p[0000001000] = \mu[1] p[1000001000] + \mu[2] p[0100001000] + \mu[3] p[0010001000] + \mu[4] p[0001001000] + \mu[5] p[0000101000] + \mu[6] p[0000011000] + \mu[8] p[0000001100] + \mu[9] p[0000001010] + \mu[10] p[0000001001]$$

Equação 10 - estado 0000001001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[7] + (\lambda[1]) + \mu[10] p[0000001001] = \mu[1] p[1000001001] + \mu[2] p[0100001001] + \mu[3] p[0010001001] + \mu[4] p[0001001001] + \mu[5] p[0000101001] + \mu[6] p[0000011001] + \mu[8] p[0000001101] + \mu[9] p[0000001011] + (\lambda[3]) p[0000001000]$$

Equação 11 - estado 0000001010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[7] + \mu[9] + (\lambda[1]+\lambda[3]) p[0000001010] = \mu[1] p[1000001010] + \mu[2] p[0100001010] + \mu[3] p[0010001010] + \mu[4] p[0001001010] + \mu[5] p[0000101010] + \mu[6] p[0000011010] + \mu[8] p[0000001110] + (\lambda[1]) p[0000001000] + \mu[10] p[0000001011]$$

Equação 12 - estado 0000001011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[7] + \mu[9] + \mu[10] p[0000001011] = \mu[1] p[1000001011] + \mu[2] p[0100001011] + \mu[3] p[0010001011] + \mu[4] p[0001001011] + \mu[5] p[0000101011] + \mu[6] p[0000011011] + \mu[8] p[0000001111] + (\lambda[1]) p[0000001001] + (\lambda[1]+\lambda[3]) p[0000001010]$$

Equação 13 - estado 0000001100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[7] + \mu[8] + (\lambda[1]) + (\lambda[3]) p[0000001100] = \mu[1] p[1000001100] + \mu[2] p[0100001100] + \mu[3] p[0010001100] + \mu[4] p[0001001100] + \mu[5] p[0000101100] + \mu[6] p[0000011100] + \mu[9] p[0000001110] + \mu[10] p[0000001101]$$

Equação 14 - estado 0000001101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[7] + \mu[8] + (\lambda[1]) + \mu[10] \\ p[0000001101] = \mu[1] p[1000001101] + \mu[2] p[0100001101] + \mu[3] p[0010001101] + \mu[4] \\ p[0001001101] + \mu[5] p[0000101101] + \mu[6] p[0000011101] + \mu[9] p[0000001111] + (\lambda[3]) \\ p[0000001100]$$

Equação 15 - estado 0000001110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[7] + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000001110] = \mu[1] p[1000001110] + \mu[2] p[0100001110] + \mu[3] p[0010001110] + \mu[4] \\ p[0001001110] + \mu[5] p[0000101110] + \mu[6] p[0000011110] + (\lambda[1]) p[0000001100] + \mu[10] \\ p[0000001111]$$

Equação 16 - estado 0000001111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[7] + \mu[8] + \mu[9] + \mu[10] \\ p[0000001111] = \mu[1] p[1000001111] + \mu[2] p[0100001111] + \mu[3] p[0010001111] + \mu[4] \\ p[0001001111] + \mu[5] p[0000101111] + \mu[6] p[0000011111] + (\lambda[1]) p[0000001101] + (\lambda[1]+\lambda[3]) \\ p[0000001110]$$

Equação 17 - estado 0000010000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + (\lambda[1]) + (\lambda[3]) p[0000010000] = \\ \mu[1] p[1000010000] + \mu[2] p[0100010000] + \mu[3] p[0010010000] + \mu[4] p[0001010000] + \mu[5] \\ p[0000101000] + \mu[7] p[0000011000] + \mu[8] p[0000010100] + \mu[9] p[0000010010] + \mu[10] \\ p[0000010001]$$

Equação 18 - estado 0000010001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + (\lambda[1]) + \mu[10] \\ p[0000010001] = \mu[1] p[1000010001] + \mu[2] p[0100010001] + \mu[3] p[0010010001] + \mu[4] \\ p[0001010001] + \mu[5] p[0000101001] + \mu[7] p[0000011001] + \mu[8] p[0000010101] + \mu[9] \\ p[0000010011] + (\lambda[3]) p[0000010000]$$

Equação 19 - estado 0000010010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000010010] = \mu[1] p[1000010010] + \mu[2] p[0100010010] + \mu[3] p[0010010010] + \mu[4] \\ p[0001010010] + \mu[5] p[0000101010] + \mu[7] p[0000011010] + \mu[8] p[0000010110] + (\lambda[1]) \\ p[0000010000] + \mu[10] p[0000010011]$$

Equação 20 - estado 0000010011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[9] + \mu[10] \\ p[0000010011] = \mu[1] p[1000010011] + \mu[2] p[0100010011] + \mu[3] p[0010010011] + \mu[4] \\ p[0001010011] + \mu[5] p[0000101011] + \mu[7] p[0000011011] + \mu[8] p[0000010111] + (\lambda[1]) \\ p[0000010001] + (\lambda[1]+\lambda[3]) p[0000010010]$$

Equação 21 - estado 0000010100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[8] + (\lambda[1]) + (\lambda[3]) \\ p[0000010100] = \mu[1] p[1000010100] + \mu[2] p[0100010100] + \mu[3] p[0010010100] + \mu[4] \\ p[0001010100] + \mu[5] p[0000101100] + \mu[7] p[0000011100] + \mu[9] p[0000010110] + \mu[10] \\ p[0000010101]$$

Equação 22 - estado 0000010101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[8] + (\lambda[1]) + \mu[10] \\ p[0000010101] = \mu[1] p[1000010101] + \mu[2] p[0100010101] + \mu[3] p[0010010101] + \mu[4] \\ p[0001010101] + \mu[5] p[0000101101] + \mu[7] p[0000011101] + \mu[9] p[0000010111] + (\lambda[3]) \\ p[0000010100]$$

Equação 23 - estado 0000010110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000010110] = \mu[1] p[1000010110] + \mu[2] p[0100010110] + \mu[3] p[0010010110] + \mu[4] \\ p[0001010110] + \mu[5] p[0000110110] + \mu[7] p[0000011110] + (\lambda[1]) p[0000010100] + \mu[10] \\ p[0000010111]$$

Equação 24 - estado 0000010111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[8] + \mu[9] + \mu[10] \\ p[0000010111] = \mu[1] p[1000010111] + \mu[2] p[0100010111] + \mu[3] p[0010010111] + \mu[4] \\ p[0001010111] + \mu[5] p[0000110111] + \mu[7] p[0000011111] + (\lambda[1]) p[0000010101] + (\lambda[1]+\lambda[3]) \\ p[0000010110]$$

Equação 25 - estado 0000011000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[7] + (\lambda[1]) + (\lambda[3]) \\ p[0000011000] = \mu[1] p[1000011000] + \mu[2] p[0100011000] + \mu[3] p[0010011000] + \mu[4] \\ p[0001011000] + \mu[5] p[0000111000] + \mu[8] p[0000011100] + \mu[9] p[0000011010] + \mu[10] \\ p[0000011001]$$

Equação 26 - estado 0000011001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[7] + (\lambda[1]) + \mu[10] \\ p[0000011001] = \mu[1] p[1000011001] + \mu[2] p[0100011001] + \mu[3] p[0010011001] + \mu[4] \\ p[0001011001] + \mu[5] p[0000111001] + \mu[8] p[0000011101] + \mu[9] p[0000011011] + (\lambda[3]) \\ p[0000011000]$$

Equação 27 - estado 0000011010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[7] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000011010] = \mu[1] p[1000011010] + \mu[2] p[0100011010] + \mu[3] p[0010011010] + \mu[4] \\ p[0001011010] + \mu[5] p[0000111010] + \mu[8] p[0000011110] + (\lambda[1]) p[0000011000] + \mu[10] \\ p[0000011011]$$

Equação 28 - estado 0000011011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[7] + \mu[9] + \mu[10] \\ p[0000011011] = \mu[1] p[1000011011] + \mu[2] p[0100011011] + \mu[3] p[0010011011] + \mu[4] \\ p[0001011011] + \mu[5] p[0000111011] + \mu[8] p[0000011111] + (\lambda[1]) p[0000011001] + (\lambda[1]+\lambda[3]) \\ p[0000011010]$$

Equação 29 - estado 0000011100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[7] + \mu[8] + (\lambda[1]) + (\lambda[3]) \\ p[0000011100] = \mu[1] p[1000011100] + \mu[2] p[0100011100] + \mu[3] p[0010011100] + \mu[4] \\ p[0001011100] + \mu[5] p[0000111100] + \mu[9] p[0000011110] + \mu[10] p[0000011101]$$

Equação 30 - estado 0000011101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[7] + \mu[8] + (\lambda[1]) + \mu[10] \\ p[0000011101] = \mu[1] p[1000011101] + \mu[2] p[0100011101] + \mu[3] p[0010011101] + \mu[4] \\ p[0001011101] + \mu[5] p[0000111101] + \mu[9] p[0000011111] + (\lambda[3]) p[0000011100]$$

Equação 31 - estado 0000011110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[7] + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000011110] = \mu[1] p[1000011110] + \mu[2] p[0100011110] + \mu[3] p[0010011110] + \mu[4] \\ p[0001011110] + \mu[5] p[0000111110] + (\lambda[1]) p[0000011100] + \mu[10] p[0000011111]$$

Equação 32 - estado 0000011111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + (\lambda[9]) + \mu[6] + \mu[7] + \mu[8] + \mu[9] + \\ \mu[10] p[0000011111] = \mu[1] p[1000011111] + \mu[2] p[0100011111] + \mu[3] p[0010011111] + \mu[4] \\ p[0001011111] + \mu[5] p[0000111111] + (\lambda[1]) p[0000011101] + (\lambda[1]+\lambda[3]) p[0000011110]$$

Equação 33 - estado 0000100000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + (\lambda[1]) + (\lambda[3]) p[0000100000] = \mu[1] p[1000100000] + \mu[2] p[0100100000] + \mu[3] p[0010100000] + \mu[4] p[0001100000] + (\lambda[9]) p[0000000000] + \mu[6] p[0000110000] + \mu[7] p[0000101000] + \mu[8] p[0000100100] + \mu[9] p[0000100010] + \mu[10] p[0000100001]$$

Equação 34 - estado 0000100001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + (\lambda[1]) + \mu[10] p[0000100001] = \mu[1] p[1000100001] + \mu[2] p[0100100001] + \mu[3] p[0010100001] + \mu[4] p[0001100001] + (\lambda[9]) p[0000000001] + \mu[6] p[0000110001] + \mu[7] p[0000101001] + \mu[8] p[0000100101] + \mu[9] p[0000100011] + (\lambda[3]) p[0000100000]$$

Equação 35 - estado 0000100010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[9] + (\lambda[1]+\lambda[3]) p[0000100010] = \mu[1] p[1000100010] + \mu[2] p[0100100010] + \mu[3] p[0010100010] + \mu[4] p[0001100010] + (\lambda[9]) p[0000000010] + \mu[6] p[0000110010] + \mu[7] p[0000101010] + \mu[8] p[0000100110] + (\lambda[1]) p[0000100000] + \mu[10] p[0000100011]$$

Equação 36 - estado 0000100011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[9] + \mu[10] p[0000100011] = \mu[1] p[1000100011] + \mu[2] p[0100100011] + \mu[3] p[0010100011] + \mu[4] p[0001100011] + (\lambda[9]) p[0000000011] + \mu[6] p[0000110011] + \mu[7] p[0000101011] + \mu[8] p[0000100111] + (\lambda[1]) p[0000100001] + (\lambda[1]+\lambda[3]) p[0000100010]$$

Equação 37 - estado 0000100100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[8] + (\lambda[1]) + (\lambda[3]) p[0000100100] = \mu[1] p[1000100100] + \mu[2] p[0100100100] + \mu[3] p[0010100100] + \mu[4] p[0001100100] + (\lambda[9]) p[0000000100] + \mu[6] p[0000110100] + \mu[7] p[0000101100] + \mu[9] p[0000100110] + \mu[10] p[0000100101]$$

Equação 38 - estado 0000100101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[8] + (\lambda[1]) + \mu[10] p[0000100101] = \mu[1] p[1000100101] + \mu[2] p[0100100101] + \mu[3] p[0010100101] + \mu[4] p[0001100101] + (\lambda[9]) p[0000000101] + \mu[6] p[0000110101] + \mu[7] p[0000101101] + \mu[9] p[0000100111] + (\lambda[3]) p[0000100100]$$

Equação 39 - estado 0000100110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) p[0000100110] = \mu[1] p[1000100110] + \mu[2] p[0100100110] + \mu[3] p[0010100110] + \mu[4] p[0001100110] + (\lambda[9]) p[0000000110] + \mu[6] p[0000110110] + \mu[7] p[0000101110] + (\lambda[1]) p[0000100100] + \mu[10] p[0000100111]$$

Equação 40 - estado 0000100111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[8] + \mu[9] + \mu[10] p[0000100111] = \mu[1] p[1000100111] + \mu[2] p[0100100111] + \mu[3] p[0010100111] + \mu[4] p[0001100111] + (\lambda[9]) p[0000000111] + \mu[6] p[0000110111] + \mu[7] p[0000101111] + (\lambda[1]) p[0000100101] + (\lambda[1]+\lambda[3]) p[0000100110]$$

Equação 41 - estado 0000101000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[7] + (\lambda[1]) + (\lambda[3]) p[0000101000] = \mu[1] p[1000101000] + \mu[2] p[0100101000] + \mu[3] p[0010101000] + \mu[4] p[0001101000] + (\lambda[9]) p[0000000100] + \mu[6] p[0000111000] + \mu[8] p[0000101100] + \mu[9] p[0000101010] + \mu[10] p[0000101001]$$

Equação 42 - estado 0000101001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[7] + (\lambda[1]) + \mu[10] \\ p[0000101001] = \mu[1] p[1000101001] + \mu[2] p[0100101001] + \mu[3] p[0010101001] + \mu[4] \\ p[0001101001] + (\lambda[9]) p[0000001001] + \mu[6] p[0000111001] + \mu[8] p[0000101101] + \mu[9] \\ p[0000101011] + (\lambda[3]) p[0000101000]$$

Equação 43 - estado 0000101010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[7] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000101010] = \mu[1] p[1000101010] + \mu[2] p[0100101010] + \mu[3] p[0010101010] + \mu[4] \\ p[0001101010] + (\lambda[9]) p[0000001010] + \mu[6] p[0000111010] + \mu[8] p[0000101110] + (\lambda[1]) \\ p[0000101000] + \mu[10] p[0000101011]$$

Equação 44 - estado 0000101011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[7] + \mu[9] + \mu[10] \\ p[0000101011] = \mu[1] p[1000101011] + \mu[2] p[0100101011] + \mu[3] p[0010101011] + \mu[4] \\ p[0001101011] + (\lambda[9]) p[0000001011] + \mu[6] p[0000111011] + \mu[8] p[0000101111] + (\lambda[1]) \\ p[0000101001] + (\lambda[1]+\lambda[3]) p[0000101010]$$

Equação 45 - estado 0000101100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[7] + \mu[8] + (\lambda[1]) + (\lambda[3]) \\ p[0000101100] = \mu[1] p[1000101100] + \mu[2] p[0100101100] + \mu[3] p[0010101100] + \mu[4] \\ p[0001101100] + (\lambda[9]) p[0000001100] + \mu[6] p[0000111100] + \mu[9] p[0000101110] + \mu[10] \\ p[0000101101]$$

Equação 46 - estado 0000101101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[7] + \mu[8] + (\lambda[1]) + \mu[10] \\ p[0000101101] = \mu[1] p[1000101101] + \mu[2] p[0100101101] + \mu[3] p[0010101101] + \mu[4] \\ p[0001101101] + (\lambda[9]) p[0000001101] + \mu[6] p[0000111101] + \mu[9] p[0000101111] + (\lambda[3]) \\ p[0000101100]$$

Equação 47 - estado 0000101110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[7] + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000101110] = \mu[1] p[1000101110] + \mu[2] p[0100101110] + \mu[3] p[0010101110] + \mu[4] \\ p[0001101110] + (\lambda[9]) p[0000001110] + \mu[6] p[0000111110] + (\lambda[1]) p[0000101100] + \mu[10] \\ p[0000101111]$$

Equação 48 - estado 0000101111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + (\lambda[9]) + \mu[7] + \mu[8] + \mu[9] + \\ \mu[10] p[0000101111] = \mu[1] p[1000101111] + \mu[2] p[0100101111] + \mu[3] p[0010101111] + \mu[4] \\ p[0001101111] + (\lambda[9]) p[0000001111] + \mu[6] p[0000111111] + (\lambda[1]) p[0000101101] + (\lambda[1]+\lambda[3]) \\ p[0000101110]$$

Equação 49 - estado 0000110000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + (\lambda[9]) + (\lambda[1]) + (\lambda[3]) \\ p[0000110000] = \mu[1] p[1000110000] + \mu[2] p[0100110000] + \mu[3] p[0010110000] + \mu[4] \\ p[0001110000] + (\lambda[9]) p[0000010000] + (\lambda[9]) p[0000100000] + \mu[7] p[0000111000] + \mu[8] \\ p[0000110100] + \mu[9] p[0000110010] + \mu[10] p[0000110001]$$

Equação 50 - estado 0000110001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + (\lambda[9]) + (\lambda[1]) + \mu[10] \\ p[0000110001] = \mu[1] p[1000110001] + \mu[2] p[0100110001] + \mu[3] p[0010110001] + \mu[4] \\ p[0001110001] + (\lambda[9]) p[0000010001] + (\lambda[9]) p[0000100001] + \mu[7] p[0000111001] + \mu[8] \\ p[0000110101] + \mu[9] p[0000110011] + (\lambda[3]) p[0000110000]$$

Equação 51 - estado 0000110010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + (\lambda[9]) + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000110010] = \mu[1] p[1000110010] + \mu[2] p[0100110010] + \mu[3] p[0010110010] + \mu[4] \\ p[0001110010] + (\lambda[9]) p[0000010010] + (\lambda[9]) p[0000100010] + \mu[7] p[0000111010] + \mu[8] \\ p[0000110110] + (\lambda[1]) p[0000110000] + \mu[10] p[0000110011]$$

Equação 52 - estado 0000110011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + (\lambda[9]) + \mu[9] + \mu[10] \\ p[0000110011] = \mu[1] p[1000110011] + \mu[2] p[0100110011] + \mu[3] p[0010110011] + \mu[4] \\ p[0001110011] + (\lambda[9]) p[0000010011] + (\lambda[9]) p[0000100011] + \mu[7] p[0000111011] + \mu[8] \\ p[0000110111] + (\lambda[1]) p[0000110001] + (\lambda[1]+\lambda[3]) p[0000110010]$$

Equação 53 - estado 0000110100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[8] + (\lambda[1]) + (\lambda[3]+\lambda[9]) \\ p[0000110100] = \mu[1] p[1000110100] + \mu[2] p[0100110100] + \mu[3] p[0010110100] + \mu[4] \\ p[0001110100] + (\lambda[9]) p[0000010100] + (\lambda[9]) p[0000100100] + \mu[7] p[0000111100] + (\lambda[9]) \\ p[0000110000] + \mu[9] p[0000110110] + \mu[10] p[0000110101]$$

Equação 54 - estado 0000110101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[8] + (\lambda[1]+\lambda[9]) + \mu[10] \\ p[0000110101] = \mu[1] p[1000110101] + \mu[2] p[0100110101] + \mu[3] p[0010110101] + \mu[4] \\ p[0001110101] + (\lambda[9]) p[0000010101] + (\lambda[9]) p[0000100101] + \mu[7] p[0000111101] + (\lambda[9]) \\ p[0000110001] + \mu[9] p[0000110111] + (\lambda[3]+\lambda[9]) p[0000110100]$$

Equação 55 - estado 0000110110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]+\lambda[9]) \\ p[0000110110] = \mu[1] p[1000110110] + \mu[2] p[0100110110] + \mu[3] p[0010110110] + \mu[4] \\ p[0001110110] + (\lambda[9]) p[0000010110] + (\lambda[9]) p[0000100110] + \mu[7] p[0000111110] + (\lambda[9]) \\ p[0000110010] + (\lambda[1]) p[0000110100] + \mu[10] p[0000110111]$$

Equação 56 - estado 0000110111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]+\lambda[9]) + \mu[5] + \mu[6] + \mu[8] + \mu[9] + \\ \mu[10] p[0000110111] = \mu[1] p[1000110111] + \mu[2] p[0100110111] + \mu[3] p[0010110111] + \mu[4] \\ p[0001110111] + (\lambda[9]) p[0000010111] + (\lambda[9]) p[0000100111] + \mu[7] p[0000111111] + (\lambda[9]) \\ p[0000110011] + (\lambda[1]+\lambda[9]) p[0000110101] + (\lambda[1]+\lambda[3]+\lambda[9]) p[0000110110]$$

Equação 57 - estado 0000111000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[7] + (\lambda[9]) + (\lambda[1]) + (\lambda[3]) \\ p[0000111000] = \mu[1] p[1000111000] + \mu[2] p[0100111000] + \mu[3] p[0010111000] + \mu[4] \\ p[0001111000] + (\lambda[9]) p[0000011000] + (\lambda[9]) p[0000101000] + \mu[8] p[0000111100] + \mu[9] \\ p[0000111010] + \mu[10] p[0000111001]$$

Equação 58 - estado 0000111001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[7] + (\lambda[9]) + (\lambda[1]) + \mu[10] \\ p[0000111001] = \mu[1] p[1000111001] + \mu[2] p[0100111001] + \mu[3] p[0010111001] + \mu[4] \\ p[0001111001] + (\lambda[9]) p[0000011001] + (\lambda[9]) p[0000101001] + \mu[8] p[0000111101] + \mu[9] \\ p[0000111011] + (\lambda[3]) p[0000111000]$$

Equação 59 - estado 0000111010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[7] + (\lambda[9]) + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0000111010] = \mu[1] p[1000111010] + \mu[2] p[0100111010] + \mu[3] p[0010111010] + \mu[4] \\ p[0001111010] + (\lambda[9]) p[0000011010] + (\lambda[9]) p[0000101010] + \mu[8] p[0000111110] + (\lambda[1]) \\ p[0000111000] + \mu[10] p[0000111011]$$

Equação 60 - estado 0000111011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[7] + (\lambda[9]) + \mu[9] + \mu[10] p[0000111011] = \mu[1] p[1000111011] + \mu[2] p[0100111011] + \mu[3] p[0010111011] + \mu[4] p[0001111011] + (\lambda[9]) p[0000011011] + (\lambda[9]) p[0000101011] + \mu[8] p[0000111111] + (\lambda[1]) p[0000111001] + (\lambda[1]+\lambda[3]) p[0000111010]$$

Equação 61 - estado 0000111100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[7] + \mu[8] + \lambda[1] + \lambda[3] + \lambda[9] p[0000111100] = \mu[1] p[1000111100] + \mu[2] p[0100111100] + \mu[3] p[0010111100] + \mu[4] p[0001111100] + (\lambda[9]) p[0000011100] + (\lambda[9]) p[0000101100] + (\lambda[9]) p[0000111000] + \mu[9] p[0000111110] + \mu[10] p[0000111101]$$

Equação 62 - estado 0000111101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[7] + \mu[8] + (\lambda[1]+\lambda[9]) + \mu[10] p[0000111101] = \mu[1] p[1000111101] + \mu[2] p[0100111101] + \mu[3] p[0010111101] + \mu[4] p[0001111101] + (\lambda[9]) p[0000011101] + (\lambda[9]) p[0000101101] + (\lambda[9]) p[0000111001] + \mu[9] p[0000111111] + (\lambda[3]+\lambda[9]) p[0000111100]$$

Equação 63 - estado 0000111110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + (\lambda[7]) + \mu[5] + \mu[6] + \mu[7] + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]+\lambda[9]) p[0000111110] = \mu[1] p[1000111110] + \mu[2] p[0100111110] + \mu[3] p[0010111110] + \mu[4] p[0001111110] + (\lambda[9]) p[0000011110] + (\lambda[9]) p[0000101110] + (\lambda[9]) p[0000111010] + (\lambda[1]) p[0000111100] + \mu[10] p[0000111111]$$

Equação 64 - estado 0000111111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + (\lambda[7]+\lambda[9]) + \mu[5] + \mu[6] + \mu[7] + \mu[8] + \mu[9] + \mu[10] p[0000111111] = \mu[1] p[1000111111] + \mu[2] p[0100111111] + \mu[3] p[0010111111] + \mu[4] p[0001111111] + (\lambda[9]) p[0000011111] + (\lambda[9]) p[0000101111] + (\lambda[9]) p[0000111011] + (\lambda[1]+\lambda[9]) p[0000111101] + (\lambda[1]+\lambda[3]+\lambda[9]) p[0000111110]$$

Equação 65 - estado 0001000000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + (\lambda[1]) + (\lambda[3]) p[0001000000] = \mu[1] p[1001000000] + \mu[2] p[0101000000] + \mu[3] p[0011000000] + (\lambda[7]) p[0000000000] + \mu[5] p[0001100000] + \mu[6] p[0001010000] + \mu[7] p[0001001000] + \mu[8] p[0001000100] + \mu[9] p[0001000010] + \mu[10] p[0001000001]$$

Equação 66 - estado 0001000001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + (\lambda[1]) + \mu[10] p[0001000001] = \mu[1] p[1001000001] + \mu[2] p[0101000001] + \mu[3] p[0011000001] + (\lambda[7]) p[0000000001] + \mu[5] p[0001100001] + \mu[6] p[0001010001] + \mu[7] p[0001001001] + \mu[8] p[0001000101] + \mu[9] p[0001000011] + (\lambda[3]) p[0001000000]$$

Equação 67 - estado 0001000010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[9] + (\lambda[1]+\lambda[3]) p[0001000010] = \mu[1] p[1001000010] + \mu[2] p[0101000010] + \mu[3] p[0011000010] + (\lambda[7]) p[0000000010] + \mu[5] p[0001100010] + \mu[6] p[0001010010] + \mu[7] p[0001001010] + \mu[8] p[0001000110] + (\lambda[1]) p[0001000000] + \mu[10] p[0001000011]$$

Equação 68 - estado 0001000011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[9] + \mu[10] p[0001000011] = \mu[1] p[1001000011] + \mu[2] p[0101000011] + \mu[3] p[0011000011] + (\lambda[7]) p[0000000011] + \mu[5] p[0001100011] + \mu[6] p[0001010011] + \mu[7] p[0001001011] + \mu[8] p[0001000111] + (\lambda[1]) p[0001000001] + (\lambda[1]+\lambda[3]) p[0001000010]$$

Equação 69 - estado 0001000100

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[8] + (\lambda[1]) + (\lambda[3]) \\
& p[0001000100] = \mu[1] p[1001000100] + \mu[2] p[0101000100] + \mu[3] p[0011000100] + (\lambda[7]) \\
& p[0000000100] + \mu[5] p[0001100100] + \mu[6] p[0001010100] + \mu[7] p[0001001100] + \mu[9] \\
& p[0001000110] + \mu[10] p[0001000101]
\end{aligned}$$

Equação 70 - estado 0001000101

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[8] + (\lambda[1]) + \mu[10] \\
& p[0001000101] = \mu[1] p[1001000101] + \mu[2] p[0101000101] + \mu[3] p[0011000101] + (\lambda[7]) \\
& p[0000000101] + \mu[5] p[0001100101] + \mu[6] p[0001010101] + \mu[7] p[0001001101] + \mu[9] \\
& p[0001000111] + (\lambda[3]) p[0001000100]
\end{aligned}$$

Equação 71 - estado 0001000110

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) \\
& p[0001000110] = \mu[1] p[1001000110] + \mu[2] p[0101000110] + \mu[3] p[0011000110] + (\lambda[7]) \\
& p[0000000110] + \mu[5] p[0001100110] + \mu[6] p[0001010110] + \mu[7] p[0001001110] + (\lambda[1]) \\
& p[0001000100] + \mu[10] p[0001000111]
\end{aligned}$$

Equação 72 - estado 0001000111

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[8] + \mu[9] + \mu[10] \\
& p[0001000111] = \mu[1] p[1001000111] + \mu[2] p[0101000111] + \mu[3] p[0011000111] + (\lambda[7]) \\
& p[0000000111] + \mu[5] p[0001100111] + \mu[6] p[0001010111] + \mu[7] p[0001001111] + (\lambda[1]) \\
& p[0001000101] + (\lambda[1]+\lambda[3]) p[0001000110]
\end{aligned}$$

Equação 73 - estado 0001001000

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[7] + (\lambda[1]) + (\lambda[3]) \\
& p[0001001000] = \mu[1] p[1001001000] + \mu[2] p[0101001000] + \mu[3] p[0011001000] + (\lambda[7]) \\
& p[0000001000] + \mu[5] p[0001101000] + \mu[6] p[0001011000] + \mu[8] p[0001001100] + \mu[9] \\
& p[0001001010] + \mu[10] p[0001001001]
\end{aligned}$$

Equação 74 - estado 0001001001

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[7] + (\lambda[1]) + \mu[10] \\
& p[0001001001] = \mu[1] p[1001001001] + \mu[2] p[0101001001] + \mu[3] p[0011001001] + (\lambda[7]) \\
& p[0000001001] + \mu[5] p[0001101001] + \mu[6] p[0001011001] + \mu[8] p[0001001101] + \mu[9] \\
& p[0001001011] + (\lambda[3]) p[0001001000]
\end{aligned}$$

Equação 75 - estado 0001001010

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[7] + \mu[9] + (\lambda[1]+\lambda[3]) \\
& p[0001001010] = \mu[1] p[1001001010] + \mu[2] p[0101001010] + \mu[3] p[0011001010] + (\lambda[7]) \\
& p[0000001010] + \mu[5] p[0001101010] + \mu[6] p[0001011010] + \mu[8] p[0001001110] + (\lambda[1]) \\
& p[0001001000] + \mu[10] p[0001001011]
\end{aligned}$$

Equação 76 - estado 0001001011

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[7] + \mu[9] + \mu[10] \\
& p[0001001011] = \mu[1] p[1001001011] + \mu[2] p[0101001011] + \mu[3] p[0011001011] + (\lambda[7]) \\
& p[0000001011] + \mu[5] p[0001101011] + \mu[6] p[0001011011] + \mu[8] p[0001001111] + (\lambda[1]) \\
& p[0001001001] + (\lambda[1]+\lambda[3]) p[0001001010]
\end{aligned}$$

Equação 77 - estado 0001001100

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[7] + \mu[8] + (\lambda[1]) + (\lambda[3]) \\
& p[0001001100] = \mu[1] p[1001001100] + \mu[2] p[0101001100] + \mu[3] p[0011001100] + (\lambda[7]) \\
& p[0000001100] + \mu[5] p[0001101100] + \mu[6] p[0001011100] + \mu[9] p[0001001110] + \mu[10] \\
& p[0001001101]
\end{aligned}$$

Equação 78 - estado 0001001101

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[7] + \mu[8] + (\lambda[1]) + \mu[10] \\
&p[0001001101] = \mu[1] p[1001001101] + \mu[2] p[0101001101] + \mu[3] p[0011001101] + (\lambda[7]) \\
&p[0000001101] + \mu[5] p[0001101101] + \mu[6] p[0001011101] + \mu[9] p[0001001111] + (\lambda[3]) \\
&p[0001001100]
\end{aligned}$$

Equação 79 - estado 0001001110

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[7] + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) \\
&p[0001001110] = \mu[1] p[1001001110] + \mu[2] p[0101001110] + \mu[3] p[0011001110] + (\lambda[7]) \\
&p[0000001110] + \mu[5] p[0001101110] + \mu[6] p[0001011110] + (\lambda[1]) p[0001001100] + \mu[10] \\
&p[0001001111]
\end{aligned}$$

Equação 80 - estado 0001001111

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[7] + \mu[8] + \mu[9] + \\
&\mu[10] p[0001001111] = \mu[1] p[1001001111] + \mu[2] p[0101001111] + \mu[3] p[0011001111] + (\lambda[7]) \\
&p[0000001111] + \mu[5] p[0001101111] + \mu[6] p[0001011111] + (\lambda[1]) p[0001001101] + (\lambda[1]+\lambda[3]) \\
&p[0001001110]
\end{aligned}$$

Equação 81 - estado 0001010000

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + (\lambda[1]) + (\lambda[3]) \\
&p[0001010000] = \mu[1] p[1001010000] + \mu[2] p[0101010000] + \mu[3] p[0011010000] + (\lambda[7]) \\
&p[0000010000] + \mu[5] p[0001110000] + \mu[7] p[0001011000] + \mu[8] p[0001010100] + \mu[9] \\
&p[0001010010] + \mu[10] p[0001010001]
\end{aligned}$$

Equação 82 - estado 0001010001

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + (\lambda[1]) + \mu[10] \\
&p[0001010001] = \mu[1] p[1001010001] + \mu[2] p[0101010001] + \mu[3] p[0011010001] + (\lambda[7]) \\
&p[0000010001] + \mu[5] p[0001110001] + \mu[7] p[0001011001] + \mu[8] p[0001010101] + \mu[9] \\
&p[0001010011] + (\lambda[3]) p[0001010000]
\end{aligned}$$

Equação 83 - estado 0001010010

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[9] + (\lambda[1]+\lambda[3]) \\
&p[0001010010] = \mu[1] p[1001010010] + \mu[2] p[0101010010] + \mu[3] p[0011010010] + (\lambda[7]) \\
&p[0000010010] + \mu[5] p[0001110010] + \mu[7] p[0001011010] + \mu[8] p[0001010110] + (\lambda[1]) \\
&p[0001010000] + \mu[10] p[0001010011]
\end{aligned}$$

Equação 84 - estado 0001010011

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[9] + \mu[10] \\
&p[0001010011] = \mu[1] p[1001010011] + \mu[2] p[0101010011] + \mu[3] p[0011010011] + (\lambda[7]) \\
&p[0000010011] + \mu[5] p[0001110011] + \mu[7] p[0001011011] + \mu[8] p[0001010111] + (\lambda[1]) \\
&p[0001010001] + (\lambda[1]+\lambda[3]) p[0001010010]
\end{aligned}$$

Equação 85 - estado 0001010100

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[8] + (\lambda[1]) + (\lambda[3]) \\
&p[0001010100] = \mu[1] p[1001010100] + \mu[2] p[0101010100] + \mu[3] p[0011010100] + (\lambda[7]) \\
&p[0000010100] + \mu[5] p[0001110100] + \mu[7] p[0001011100] + \mu[9] p[0001010110] + \mu[10] \\
&p[0001010101]
\end{aligned}$$

Equação 86 - estado 000 1 0 1 0 1 0 1

$$\begin{aligned}
&(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[8] + (\lambda[1]) + \mu[10] \\
&p[0001010101] = \mu[1] p[1001010101] + \mu[2] p[0101010101] + \mu[3] p[0011010101] + (\lambda[7]) \\
&p[0000010101] + \mu[5] p[0001110101] + \mu[7] p[0001011101] + \mu[9] p[0001010111] + (\lambda[3]) \\
&p[0001010100]
\end{aligned}$$

Equação 87 - estado 0001010110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[8] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0001010110] = \mu[1] p[1001010110] + \mu[2] p[0101010110] + \mu[3] p[0011010110] + (\lambda[7]) \\ p[0000010110] + \mu[5] p[0001110110] + \mu[7] p[0001011110] + (\lambda[1]) p[0001010100] + \mu[10] \\ p[0001010111]$$

Equação 88 - estado 0001010111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[8] + \mu[9] + \\ \mu[10] p[0001010111] = \mu[1] p[1001010111] + \mu[2] p[0101010111] + \mu[3] p[0011010111] + (\lambda[7]) \\ p[0000010111] + \mu[5] p[0001110111] + \mu[7] p[0001011111] + (\lambda[1]) p[0001010101] + (\lambda[1]+\lambda[3]) \\ p[0001010110]$$

Equação 89 - estado 0001011000

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[7] + (\lambda[1]) + (\lambda[3]) \\ p[0001011000] = \mu[1] p[1001011000] + \mu[2] p[0101011000] + \mu[3] p[0011011000] + (\lambda[7]) \\ p[0000011000] + \mu[5] p[0001111000] + \mu[8] p[0001011100] + \mu[9] p[0001011010] + \mu[10] \\ p[0001011001]$$

Equação 90 - estado 0001011001

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[7] + (\lambda[1]) + \mu[10] \\ p[0001011001] = \mu[1] p[1001011001] + \mu[2] p[0101011001] + \mu[3] p[0011011001] + (\lambda[7]) \\ p[0000011001] + \mu[5] p[0001111001] + \mu[8] p[0001011101] + \mu[9] p[0001011011] + (\lambda[3]) \\ p[0001011000]$$

Equação 91 - estado 0001011010

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[7] + \mu[9] + (\lambda[1]+\lambda[3]) \\ p[0001011010] = \mu[1] p[1001011010] + \mu[2] p[0101011010] + \mu[3] p[0011011010] + (\lambda[7]) \\ p[0000011010] + \mu[5] p[0001111010] + \mu[8] p[0001011110] + (\lambda[1]) p[0001011000] + \mu[10] \\ p[0001011011]$$

Equação 92 - estado 0001011011

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[7] + \mu[9] + \\ \mu[10] p[0001011011] = \mu[1] p[1001011011] + \mu[2] p[0101011011] + \mu[3] p[0011011011] + (\lambda[7]) \\ p[0000011011] + \mu[5] p[0001111011] + \mu[8] p[0001011111] + (\lambda[1]) p[0001011001] + (\lambda[1]+\lambda[3]) \\ p[0001011010]$$

Equação 93 - estado 0001011100

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[7] + \mu[8] + (\lambda[1]) + (\lambda[3]) \\ p[0001011100] = \mu[1] p[1001011100] + \mu[2] p[0101011100] + \mu[3] p[0011011100] + (\lambda[7]) \\ p[0000011100] + \mu[5] p[0001111100] + \mu[9] p[0001011110] + \mu[10] p[0001011101]$$

Equação 94 - estado 0001011101

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[7] + \mu[8] + (\lambda[1]) + \\ \mu[10] p[0001011101] = \mu[1] p[1001011101] + \mu[2] p[0101011101] + \mu[3] p[0011011101] + (\lambda[7]) \\ p[0000011101] + \mu[5] p[0001111101] + \mu[9] p[0001011111] + (\lambda[3]) p[0001011100]$$

Equação 95 - estado 0001011110

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[7] + \mu[8] + \mu[9] + \\ (\lambda[1]+\lambda[3]) p[0001011110] = \mu[1] p[1001011110] + \mu[2] p[0101011110] + \mu[3] p[0011011110] + \\ (\lambda[7]) p[0000011110] + \mu[5] p[0001111110] + (\lambda[1]) p[0001011100] + \mu[10] p[0001011111]$$

Equação 96 - estado 0001011111

$$(\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + (\lambda[7]+\lambda[9]) + \mu[6] + \mu[7] + \mu[8] + \mu[9] \\ + \mu[10] p[0001011111] = \mu[1] p[1001011111] + \mu[2] p[0101011111] + \mu[3] p[0011011111] + (\lambda[7]) \\ p[0000011111] + \mu[5] p[0001111111] + (\lambda[1]) p[0001011101] + (\lambda[1]+\lambda[3]) p[0001011110]$$

Equação 97 - estado 0001100000

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + \mu[5] + (\lambda[9]) + (\lambda[7]) + (\lambda[1]) + (\lambda[3]) \\
& p[0001100000] = \mu[1] p[1001100000] + \mu[2] p[0101100000] + \mu[3] p[0011100000] + (\lambda[7]) \\
& p[0000100000] + (\lambda[7]+\lambda[9]) p[0001000000] + \mu[6] p[0001110000] + \mu[7] p[0001101000] + \mu[8] \\
& p[0001100100] + \mu[9] p[0001100010] + \mu[10] p[0001100001]
\end{aligned}$$

Equação 98 - estado 0001100001

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[3]+\lambda[5]) + \mu[4] + \mu[5] + (\lambda[9]) + (\lambda[7]) + (\lambda[1]) + \mu[10] \\
& p[0001100001] = \mu[1] p[1001100001] + \mu[2] p[0101100001] + \mu[3] p[0011100001] + (\lambda[7]) \\
& p[0000100001] + (\lambda[7]+\lambda[9]) p[0001000001] + \mu[6] p[0001110001] + \mu[7] p[0001101001] + \mu[8] \\
& p[0001100101] + \mu[9] p[0001100011] + (\lambda[3]) p[0001100000]
\end{aligned}$$

Equação 99 - estado 0001100010

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[5]) + \mu[4] + \mu[5] + (\lambda[9]) + (\lambda[7]) + \mu[9] + (\lambda[1]+\lambda[3]) \\
& p[0001100010] = \mu[1] p[1001100010] + \mu[2] p[0101100010] + \mu[3] p[0011100010] + (\lambda[7]) \\
& p[0000100010] + (\lambda[7]+\lambda[9]) p[0001000010] + \mu[6] p[0001110010] + \mu[7] p[0001101010] + \mu[8] \\
& p[0001100110] + (\lambda[1]) p[0001100000] + \mu[10] p[0001100011]
\end{aligned}$$

Equação 100 - estado 0001100011

$$\begin{aligned}
& (\lambda[2]+\lambda[6]) + (\lambda[4]+\lambda[8]+\lambda[10]) + (\lambda[1]+\lambda[3]+\lambda[5]) + \mu[4] + \mu[5] + (\lambda[9]) + (\lambda[7]) + \mu[9] + \mu[10] \\
& p[0001100011] = \mu[1] p[1001100011] + \mu[2] p[0101100011] + \mu[3] p[0011100011] + (\lambda[7]) \\
& p[0000100011] + (\lambda[7]+\lambda[9]) p[0001000011] + \mu[6] p[0001110011] + \mu[7] p[0001101011] + \mu[8] \\
& p[0001100111] + (\lambda[1]) p[0001100001] + (\lambda[1]+\lambda[3]) p[0001100010]
\end{aligned}$$