



APLICAÇÃO DO MODELO HIPERCUBO PARA ANÁLISE DE UM SISTEMA MÉDICO-EMERGENCIAL EM RODOVIA

Fernando César Mendonça
Reinaldo Morabito

Departamento de Engenharia de Produção
Universidade Federal de São Carlos
Caixa Postal 676 – 13565-905 – São Carlos – SP
E-mail: morabito@power.ufscar.br

Resumo

Em um sistema médico-emergencial, o nível de serviço é fundamental, pois a qualidade de atendimento poderá ser a diferença entre a vida e a morte do usuário. Além da qualidade do atendimento, um dos componentes mais importantes do nível de serviço é o tempo médio de resposta a uma chamada emergencial. Entretanto, adquirir equipamentos e/ou treinamentos para a equipe, para reduzir este tempo de resposta, incorre em custos e investimentos adicionais. Nesta situação, aparece um importante trade-off entre o nível de serviço e o investimento no sistema. Para auxiliar a análise deste trade-off, há um modelo conhecido da literatura chamado hipercubo, que é baseado em teoria de filas espacialmente distribuídas. Neste artigo, adaptamos e aplicamos o modelo hipercubo para analisar os Anjos do Asfalto, um sistema médico-emergencial localizado ao longo da Rodovia Presidente Dutra entre as cidades do Rio de Janeiro e São Paulo. Os resultados mostraram que o modelo hipercubo pode ser uma ferramenta efetiva de apoio às decisões estratégicas e operacionais em sistemas emergenciais em rodovias.

Palavras-chave: *atendimento médico-emergencial, teoria de filas, modelo hipercubo.*

1. Introdução

Nosso cotidiano está cercado de situações em que serviços básicos são requisitados, tais

como atendimentos médicos, policiamento, coleta de lixo, reparos a veículos, transporte coletivo, etc. Com o aumento da população, cresce a potencial demanda por tais serviços, e surge a

necessidade de melhorar o gerenciamento dos sistemas que os oferecem, para que seja garantido um nível de atendimento aceitável aos usuários. Em geral, o nível de serviço está relacionado com o nível de congestão do sistema, pois, é comum termos situações em que, num certo intervalo de tempo, há mais demanda por serviço do que a capacidade de atendimento do sistema. Um baixo nível de serviço pode gerar um prejuízo ao sistema não propriamente financeiro, mas um prejuízo intangível, que é sua confiabilidade. Os serviços de atendimentos à população podem ser classificados em três categorias:

- (i) Serviços de rotina (ou periódicos): coleta domiciliar de lixo, entrega de jornais, distribuição de correspondências.
- (ii) Serviços semi-emergenciais: reparos em rede de energia elétrica, água, saneamentos, telefonias.
- (iii) Serviços de emergência: ambulâncias, polícia, bombeiros.

Nas duas primeiras classes, podem haver pequenos atrasos ou filas de espera; já na terceira, o custo intangível é muito alto, fazendo com que sistemas dessa categoria devam ser planejados com altos índices de ociosidade, para que a probabilidade de congestão ou espera em fila seja pequena. Algumas vezes, em um acidente traumático, poucos minutos de atraso no atendimento às vítimas podem representar a diferença entre a vida e a morte.

Quando pensamos em gerenciamento de recursos de um sistema de atendimento médico-emergencial, não podemos esquecer que melhor qualidade em serviço, em geral, implica em maiores custos e investimentos. Independente da forma de financiamento da empresa que gerencia o sistema de atendimento (público, privado ou cooperativo), é desejável que os recursos, não apenas financeiros, sejam alocados da melhor forma possível. Cabe observar que este *trade-off* entre nível de serviço e custo (ou investimento) é típico em diversos outros sistemas de serviço e manufatura (FONTANELLA & MORABITO, 1997; BITRAN & MORABITO, 1999; MORABITO & LIMA, 1999).

Neste artigo, adaptamos e aplicamos uma ferramenta conhecida da literatura, o *modelo hiper-cubo*, para auxiliar nas decisões envolvidas em um sistema de atendimento médico-emergencial na Rodovia Presidente Dutra (Via Dutra), entre as cidades do Rio de Janeiro e São Paulo. Em particular, tratamos apenas do trecho coberto pelos *Anjos do Asfalto*, uma fundação não-governamental e sem fins lucrativos, que presta socorro a acidentados entre as regiões de Lavrinhas (SP) e Pavuna (RJ) (figura 1). O desempenho do modelo é analisado a partir de dados reais coletados junto aos Anjos do Asfalto.

O modelo hiper-cubo, proposto por LARSON (1974) e estudado por diversos autores (SWERSEY, 1994), é uma importante ferramenta para o planejamento de sistemas de serviços, especialmente os sistemas urbanos em que servidores se deslocam para fornecer algum tipo de serviço para clientes. O modelo aborda complexidades geográficas e temporais da região, baseado em resultados de teoria de filas espacialmente distribuídas e aproximações a partir de análise Markoviana. Basicamente, a idéia é expandir a descrição do espaço de estados de um sistema de fila com múltiplos servidores, para poder representar cada servidor individualmente e incorporar políticas de despacho mais complexas. O modelo consiste em resolver um sistema linear que fornece as probabilidades de equilíbrio dos possíveis estados do sistema. Estas probabilidades permitem estimar diversas medidas de desempenho interessantes para o gerenciamento do sistema, tais como *workloads* (cargas de trabalho) dos servidores, tempo médio de resposta do sistema ou de cada servidor, frequência de atendimento de cada servidor em cada região, entre outras.

Alguns exemplos de aplicação do modelo hiper-cubo na literatura são: a localização de ambulâncias em Boston nos EUA (BRANDEAU & LARSON, 1986), o patrulhamento policial em Orlando nos EUA (SACKS & GRIEF, 1994), o atendimento a interrupções de redes de energia elétrica em Santa Catarina (ALBINO, 1994), a

localização de ambulâncias em um trecho da BR-111 (GONÇALVES *et al.*, 1994, 1995), o programa de visitas de serviço social (LARSON & ODONI, 1981), entre outros. Extensões do modelo hipercubo foram consideradas em HALPERN (1977) e BURWELL *et al.* (1993). Outros modelos para localização de ambulâncias em cidades como Austin e Tucson nos EUA, e Bangkok na Tailândia, foram apresentados em EATON *et al.* (1985), GOLDBERG & PAZ (1991) e FUJIWARA (1987). Um exame do modelo hipercubo e outros modelos para análise de sistemas de atendimento emergencial pode ser encontrado em SWERSEY (1994).

O presente artigo está organizado da seguinte maneira: na seção 2, descrevemos resumidamente o sistema Anjos do Asfalto, com ênfase nos aspectos do seu funcionamento importantes para a aplicação do modelo hipercubo. Na seção 3, revisamos brevemente o modelo hipercubo e discutimos algumas pequenas adaptações para sua aplicação no sistema estudado. Na seção 4, apresentamos e analisamos as medidas de desempenho obtidas por meio do modelo hipercubo, ao utilizarmos dados reais coletados junto à empresa. Na seção 5, aplicamos o modelo para explorar cenários alternativos, em particular, modificando a configuração do sistema para melhorar o balanceamento do *workload* entre as ambulâncias. Também analisamos os efeitos de variações nos parâmetros do sistema, por exemplo, a demanda dos usuários, para investigar a robustez dos resultados obtidos. Finalmente, na seção 6, apresentamos as conclusões do estudo e algumas perspectivas para pesquisa futura.

2. O Sistema Anjos do Asfalto

Conforme mencionado, o sistema Anjos do Asfalto é uma fundação não-governamental sem fins lucrativos, que presta socorro a acidentados e feridos ou pessoas que se sentem mal na Via Dutra. Tal sistema possui seis bases fixas ao longo da rodovia, e cada base é composta por ambulâncias de atendimentos de urgência e uma equipe composta por resgatistas,

que fazem os resgates necessários, e médicos, que fazem os atendimentos de socorro propriamente ditos, além dos motoristas. Esta equipe só se movimenta em conjunto. As localizações das bases são: km 165-RJ (Pavuna), km 206-RJ (Japeri), km 227-RJ (Caiçara), km 258-RJ (Volta Redonda), km 311-RJ (Penedo) e km 18-SP (Lavrinhas), conforme ilustrado na figura 1.

Há um posto central na cidade do Rio de Janeiro que monitora a rodovia por rádio. Qualquer movimento da ambulância, mesmo abastecimento, é avisado à central. Em princípio, a alocação de ambulâncias é coordenada: quando há um acidente, alguma testemunha avisa diretamente a central ou via Polícia Rodoviária, e a central encaminha a ambulância disponível mais próxima ao local do acidente. Entretanto, pode acontecer de a testemunha ou a polícia avisar diretamente a base. Neste caso, a base avisa a central via rádio e comunica o local do acidente. Após a privatização da Via Dutra, a concessionária possui veículos especiais que circulam pela rodovia, avisando a central se houver algum acidente. Nos dados coletados, grande parte dos chamados foram feitos dessa forma.

A política de despacho é sempre alocar a ambulância mais próxima ao local da ocorrência. Caso esta esteja ocupada, a central procura alocar a segunda ambulância mais próxima do chamado (nesse caso, diz-se que ocorreu um atendimento de *backup*). Se a segunda ambulância também estiver ocupada, o chamado é transferido para a Polícia Rodoviária ou o Corpo de Bombeiros da cidade mais próxima ao local, ou seja, nunca é alocada a terceira ambulância mais próxima do chamado. Desta forma, o sistema nunca terá filas de chamados, pois, se um chamado encontrar as duas ambulâncias mais próximas ocupadas, ele será transferido para outro sistema de atendimento. Também é possível que a ambulância fique ocupada sem sair para atender um chamado, mas atendendo alguém que peça auxílio na própria base (p.e., epiléticos, cardíacos com mal-estar, pessoas alcoolizadas). Neste caso, a central também é

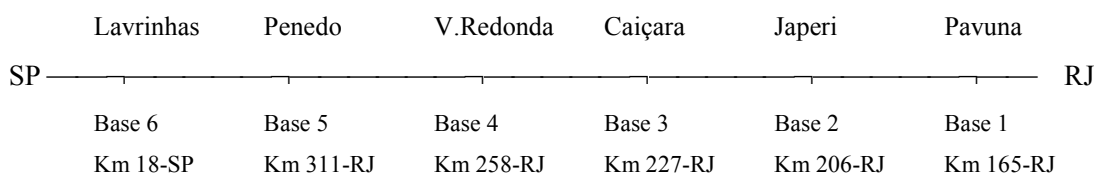


Figura 1 – Esboço da rodovia com as suas bases de atendimento.

informada. Quando um chamado não é atendido pelos Anjos do Asfalto, é considerado como perda do sistema.

Em geral os médicos dos Anjos do Asfalto dão os primeiros socorros, e transportam a vítima até o hospital mais próximo, só depois retornando à base original. No caso da vítima não ter condições clínicas para os métodos terrestres convencionais de transporte, os socorristas prestam os primeiros atendimentos e, via rádio, solicitam o helicóptero do Corpo de Bombeiros, retornando depois para a base. Os dados estão disponíveis em livros de registro, nos quais o operador da central anota todos os eventos ocorridos durante o dia. Não há nada registrado em computador. No livro é anotado o instante em que a central foi chamada, o local da ocorrência, a ambulância que foi despachada, e o instante em que a ambulância retornou à base, além de detalhes técnicos do chamado, tais como espécie do acidente, número de vítimas, encaminhamentos.

3. O Modelo Hiper-cubo no Sistema Anjos do Asfalto

O modelo hiper-cubo proposto em LARSON (1974) baseia-se na divisão da região em um conjunto de *átomos geográficos*. Cada átomo é considerado como uma fonte independente solicitadora de serviços ao longo do tempo. Desta maneira, o modelo considera os chamados distribuídos tanto espacialmente quanto temporalmente na região. O atendimento aos chamados de cada átomo é realizado por *servidores* (no caso, ambulâncias), espacialmente distribuídos na região, que podem estar fixos em alguns pontos (bases), ou movimentando-se em sub-regiões (neste caso, seus movimentos deverão

ser conhecidos ao menos probabilisticamente). Para um servidor, sua *área de cobertura primária* é o conjunto dos átomos para os quais este servidor tem prioridade de atendimento (i.e., é o primeiro a ser chamado). No caso deste servidor estar ocupado, outros servidores são chamados para atender à solicitação.

O nome hiper-cubo deriva da descrição da disponibilidade dos servidores por meio do espaço de estados. Cada servidor pode estar livre (0) ou ocupado (1) num certo instante. Um estado particular do sistema é dado pela lista dos servidores que estão livres e ocupados. Por exemplo, o estado 110 corresponde a um sistema com três servidores, com o servidor 1 livre e os servidores 2 e 3 ocupados (note que 110 descreve o estado dos servidores da direita para a esquerda). Desta maneira, o espaço de estados do sistema é dado pelos vértices de um cubo. No caso de termos mais de três servidores, temos um hiper-cubo. O modelo trata tanto sistemas em que os chamados não atendidos aguardam em fila, quanto sistemas que não admitem a formação de filas, como é o caso dos Anjos do Asfalto. A descrição detalhada deste modelo está além dos objetivos deste artigo; para mais detalhes, o leitor pode consultar LARSON & ODONI (1981) e MENDONÇA (1999).

Para a aplicação do modelo, há nove hipóteses básicas:

1) Átomos geográficos: a região analisada é dividida em N_A átomos geográficos. No caso dos Anjos do Asfalto, estes átomos correspondem a trechos da rodovia, conforme é visto adiante.

2) Processos de chegada conforme processos de *Poisson* independentes: admite-se que os chamados de cada átomo j são gerados conforme um processo de *Poisson*, independente dos

outros átomos, com taxa média λ_j ($j = 1, 2, \dots, N_A$). A validade desta hipótese para os Anjos do Asfalto é verificada na seção 4.

3) Tempos de deslocamento do servidor: os tempos médios de viagem τ_{ij} entre o átomo i e o átomo j ($i, j = 1, 2, \dots, N_A$) deverão ser conhecidos ou estimados pelos conceitos de probabilidade geométrica. No presente trabalho, estes tempos foram estimados com base nas estatísticas de velocidades médias de viagem entre átomos, fornecidas pela empresa (MENDONÇA, 1999).

4) Servidores: existem N servidores espacialmente distribuídos ao longo de região, que podem se deslocar e atender qualquer um dos átomos. Essa hipótese é relaxada adiante, devido à política de despacho particular dos Anjos do Asfalto descrita na seção 2.

5) Localização dos servidores: cada servidor, quando disponível, pode ficar fixo em um átomo, ou se mover (p.e., em patrulhamento) dentro de uma área (neste caso, sua localização deverá ser conhecida, ao menos, probabilisticamente). Neste estudo, uma ambulância permanece em cada base da figura 1, à espera de um chamado.

6) Despacho de um servidor: a cada serviço solicitado, apenas um servidor é despachado para atendê-lo. Se não houver servidores disponíveis, poderá haver formação de filas (no caso de sistemas que permitem filas), ou perda do chamado (no caso de sistemas que não permitem filas), que então é transferido para outro sistema de atendimento. Conforme já discutido, este último é o caso dos Anjos do Asfalto.

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 do átomo, caso contrário, o próximo servidor disponível da lista (i.e., o *backup*) é despachado. A lista de preferências é fixada *a priori* e permanece inalterada durante a operação do sistema. Note que esta política de despacho é compatível com a dos Anjos do Asfalto descrita na seção 2, exceto que a lista contém apenas dois servidores para cada átomo.

8) Tempo de serviço: o tempo de serviço de um chamado inclui o tempo de *setup* (tempo de preparo do servidor), tempo de viagem do servidor até o local onde houve o chamado, o tempo de realização do serviço propriamente dito (em cena), e o tempo de retorno à base. Os servidores têm taxas médias de serviço μ_n ($n = 1, \dots, N$) que podem ser diferentes, como é o caso dos Anjos do Asfalto, conforme discutido na seção 4. No caso do sistema permitir formação de filas, o modelo funciona melhor à medida que os tempos médios de serviço se aproximam dos respectivos desvios-padrões, isto é, o processo de serviço tende a ser exponencialmente distribuído. Entretanto, conforme LARSON & ODONI (1981), desvios razoáveis desta hipótese não alteram sensivelmente a precisão do modelo. Se o sistema não permitir filas, como é o caso dos Anjos do Asfalto, esta hipótese é ainda menos necessária, dado que sistemas de filas $M/M/N/N$ e $M/G/N/N$ têm a mesma distribuição de equilíbrio.

9) Dependência do tempo de serviço em relação ao tempo de viagem: variações no tempo de serviço devidas às variações no tempo de viagem são assumidas como sendo de segunda ordem, quando comparadas com as variações dos tempos em cena e/ou *setup* (isso não significa que o tempo de deslocamento deva ser ignorado ao se computar o tempo de serviço). Esta hipótese é verificada para os Anjos do Asfalto na seção 4.

Conforme mencionado na seção 1, o modelo hipercubo é baseado em resultados de teoria de filas espacialmente distribuídas e aproximações a partir de análise Markoviana. Para um sistema que não permite filas, a idéia básica é estender o espaço de estados do modelo $M/M/N/N$ ou $M/G/N/N$, para descrever o estado de cada um dos N servidores individualmente, e analisar o sistema sob hipótese de equilíbrio (*steady state*). Como é visto adiante (veja equações de equilíbrio abaixo), o modelo envolve resolver um sistema de 2^N equações lineares, que fornece as probabilidades de equilíbrio dos possíveis estados do sistema.

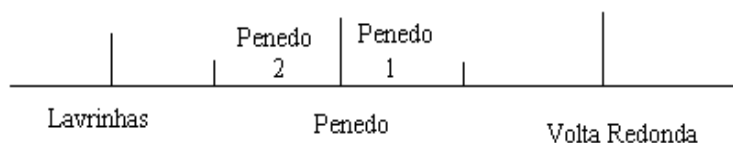


Figura 2 – Base exemplo e suas vizinhanças.

Tabela 1 – Preferências de despacho de servidores para cada átomo.

Átomo	1 ^a . opção	2 ^a . opção
1 – Pavuna	1	2
2 – Japeri 1	2	1
3 – Japeri 2	2	3
4 – Caiçara 1	3	2
5 – Caiçara 2	3	4
6 – Volta Redonda 1	4	3
7 – Volta Redonda 2	4	5
8 – Penedo 1	5	4
9 – Penedo 2	5	6
10 – Lavrinhas	6	5

O sistema Anjos do Asfalto, por possuir 6 bases, corresponde a um sistema com $N=6$ servidores. Portanto, cada estado do sistema é representado por uma lista sêxtupla de 0s (servidor disponível) e 1s (servidor ocupado). Como devemos ter uma equação de equilíbrio para cada estado, temos $2^6 = 64$ estados possíveis e, portanto, o modelo resulta em 64 equações lineares. Se o número de servidores for muito grande (digamos, $N > 15$), o modelo pode ser aproximado por um sistema de N equações não-lineares (LARSON & ODONI, 1981), ou o problema pode ser tratado dividindo-se a rodovia em subtrechos, tendo em vista a política de despacho particular. Isto não será explorado neste trabalho, dado que a solução de sistemas com apenas 64 equações lineares não envolve grandes dificuldades computacionais.

A numeração das bases 1,2,...,6 foi definida começando pelas bases mais próximas do Rio de Janeiro, conforme mostra a figura 1. As regiões próximas a cada base são divididas em dois

átomos diferentes, visto que as opções de despachos não são as mesmas para ambos os lados da região. Tomemos, por exemplo, a região de Penedo (figura 2): sua parte leste (Penedo 1) tem como primeira opção de atendimento a base 5 (Penedo), e como *backup* a base 4 (Volta Redonda). Já a parte oeste (Penedo 2) também tem como primeira opção a base 5, mas como *backup* a base 6 (Lavrinhas). A figura 2 representa tal situação, e a tabela 1 apresenta as preferências de cada um dos $N_A=10$ átomos, que é compatível com a hipótese 7 do modelo hipercubo.

Com exceção das bases 1 e 6, todas as demais podem vir a atender quatro átomos (dois como primeira opção e dois como *backup*). Se duas bases adjacentes estão ocupadas (p.e., as bases 2 e 3) e ocorrer um chamado em um átomo onde estas bases são as únicas opções (p.e., os átomos 3 ou 4), haverá atendimento por parte dos bombeiros ou Polícia Rodoviária, sendo considerado, para efeito de modelagem, um chamado perdido

e transferido para outro sistema. Isto torna o modelo hipercubo dos Anjos do Asfalto diferente do originalmente proposto em LARSON (1974), uma vez que a política de despacho é mais particular do que a da hipótese 4. Assim, a hipótese 4, que admite que qualquer servidor atenda qualquer átomo, terá de ser relaxada para que o modelo possa ser aplicado.

Equações de equilíbrio

As equações de equilíbrio (balanço) são definidas supondo-se que o sistema atinja estado estável (*steady state*). Para cada um dos 64 estados do sistema, o fluxo com que se entra num estado (i.e., o produto entre a probabilidade do sistema estar em outro estado e a taxa de transição daquele para o estado atual) deve ser igual ao fluxo com que o sistema sai do estado. Para mais detalhes sobre isso, o leitor pode consultar LARSON & ODONI (1981) e MENDONÇA (1999). A figura 3 ilustra os fluxos de entrada e saída de um estado com apenas um servidor ocupado, no caso o estado 001000, com apenas o servidor 4 ocupado. Note na figura que o fluxo de entrada no estado 001000 é:

$$\mu_1 p_{001001} + \mu_2 p_{001010} + \mu_3 p_{001100} + \mu_5 p_{001100} + \mu_6 p_{101000} + (\lambda_6 + \lambda_7) p_{000000}$$

onde p_B denota a probabilidade de equilíbrio do estado B . O fluxo de saída do estado 001000 é:

$$(\mu_4 + \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10}) p_{001000}$$

Igualando estes fluxos de entrada e saída, obtemos a equação de equilíbrio para o estado 001000. Note que, diferente das equações de equilíbrio de um sistema $M/M/N/N$, aqui as taxas médias de serviço μ_n variam para cada servidor $n=1, \dots, N$. Outros estados com apenas um dos servidores ocupados possuem equações de equilíbrio similares. Em estados com dois ou mais servidores ocupados, se estes servidores forem adjacentes, aparece uma probabilidade de perda. Para ilustrar isso, considere, por exemplo, a equação de equilíbrio do estado 010001

(expressão (1)), em que os dois servidores ocupados 1 e 5 não são adjacentes. Os fluxos de entrada e saída deste estado estão representados na figura 4:

$$\begin{aligned} &\mu_6 p_{110001} + \mu_4 p_{011001} + \mu_3 p_{010101} + \mu_2 p_{010011} + \\ &+ \lambda_1 p_{010000} + (\lambda_8 + \lambda_9) p_{000001} = \\ &= ((\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10}) + \\ &+ \mu_1 + \mu_5) p_{010001} \end{aligned} \tag{1}$$

Note que, se o sistema estiver no estado 010001, qualquer chegada no sistema é atendida, pois, o termo $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10}$ na taxa de saída deste estado (veja expressão (1)) é a soma de *todas* as taxas de chegada dos átomos. Por outro lado, a equação de equilíbrio do estado 000110 (expressão (2)), em que os dois servidores ocupados 2 e 3 são adjacentes, envolve perda de chamados. Note agora que o termo $\lambda_1 + \lambda_2 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10}$ na taxa de saída do estado 000110 (expressão (2)) não contém as taxas de chegada λ_3 e λ_4 dos átomos 3 e 4, uma vez que, se o sistema estiver neste estado, estas chegadas não são atendidas, pois, tanto a primeira opção quanto o *backup* (i.e., os servidores 2 e 3) estão ocupados. A figura 5 ilustra os fluxos de entrada e saída do estado 000110.

$$\begin{aligned} &\mu_1 p_{000111} + \mu_4 p_{001110} + \mu_5 p_{010110} + \mu_6 p_{100110} + \\ &+ (\lambda_4 + \lambda_5 + \lambda_3) p_{000010} + (\lambda_2 + \lambda_3 + \lambda_4) p_{000100} = \\ &= (\mu_2 + \mu_3 + (\lambda_1 + \lambda_2 + \lambda_5 + \lambda_6 + \lambda_7 + \\ &+ \lambda_8 + \lambda_9 + \lambda_{10})) p_{000110} \end{aligned} \tag{2}$$

Procedendo-se de maneira similar para os outros estados, obtém-se 64 equações de equilíbrio, linearmente dependentes entre si. Por serem probabilidades, temos que:

$$p_{000000} + p_{000001} + p_{000010} + p_{000100} + \dots + p_{101111} + p_{011111} + p_{111111} = 1$$

e substituindo-se esta equação por qualquer uma das equações de equilíbrio, tornamos o sistema determinado, com 64 equações linearmente independentes.

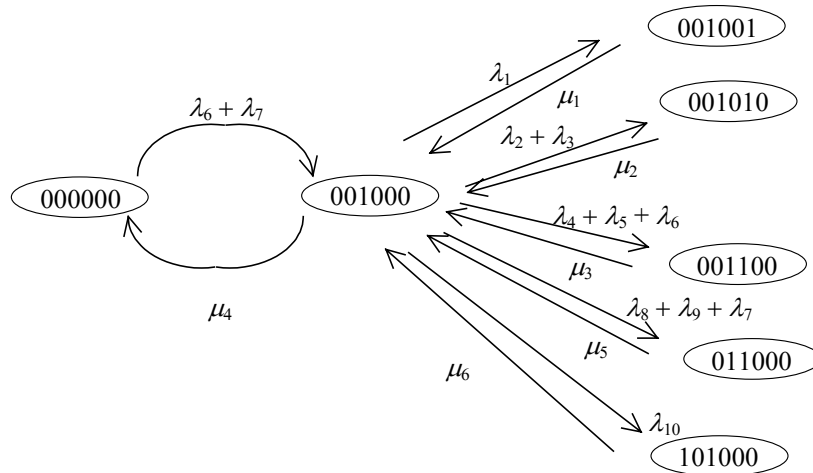


Figura 3 – Fluxos de entrada e saída do estado 001000.

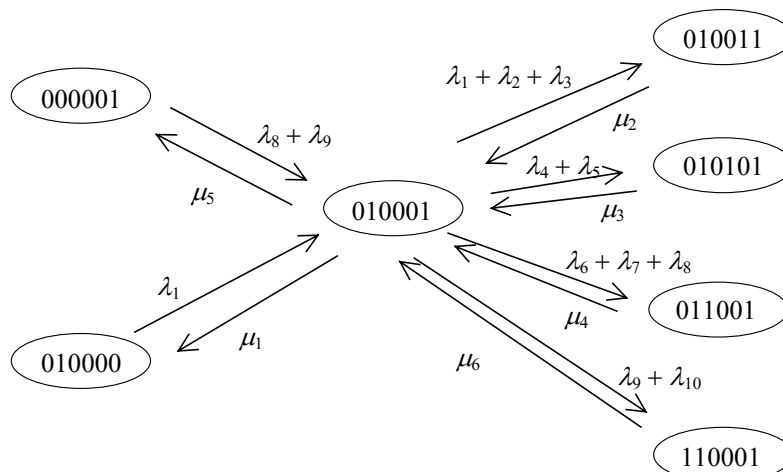


Figura 4 – Fluxos de entrada e saída do estado 010001.

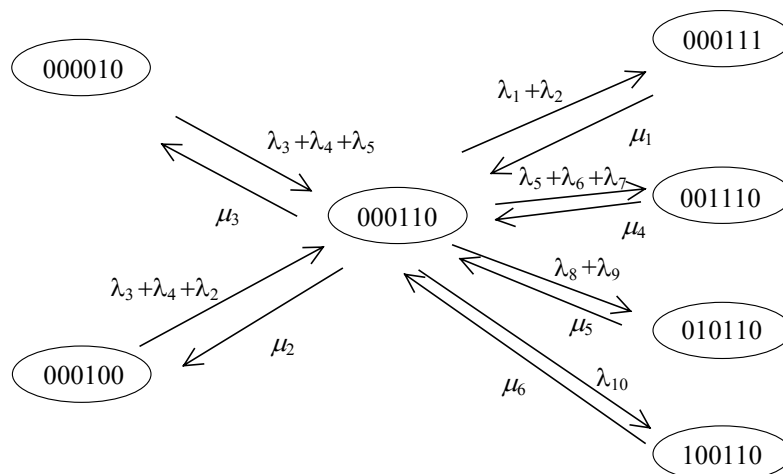


Figura 5 – Fluxos de entrada e saída do estado 000110.

Medidas de desempenho

Com a distribuição de equilíbrio de estados computada, podemos estimar diversas medidas de desempenho, conforme LARSON (1974) e LARSON & ODONI (1981). Entretanto, devido às características do sistema Anjos do Asfalto, algumas medidas de desempenho precisam ser modificadas, conforme discutido a seguir.

O *workload* ρ_n de cada servidor n (fração do tempo em que o servidor está ocupado) é definido, conforme LARSON & ODONI (1981), como sendo a soma das probabilidades em que o servidor em questão está ocupado. Por exemplo, o *workload* do servidor 6 é dado por:

$$\begin{aligned} \rho_6 = & p_{100000} + p_{100001} + p_{100010} + p_{100100} + \\ & p_{101000} + p_{110000} + p_{100011} + p_{100101} + \\ & p_{101001} + p_{110001} + p_{100110} + p_{101010} + \\ & p_{110010} + p_{101100} + p_{110100} + p_{111000} + \\ & p_{100111} + p_{101011} + p_{101101} + p_{101110} + \\ & p_{110011} + p_{110101} + p_{110110} + p_{111001} + \\ & p_{111010} + p_{111100} + p_{011111} + p_{110111} + \\ & p_{111011} + p_{111101} + p_{111110} + p_{111111} \end{aligned} \quad (3)$$

Já a frequência de despacho do servidor n ao átomo j , f_{nj} , precisa ser modificada em relação à expressão em LARSON & ODONI (1981). No caso dos Anjos do Asfalto, conforme discussão nas equações de equilíbrio, o estado 111111 não é o único em que pode haver perda de chamados, mas qualquer estado que tenha dois ou mais servidores ocupados adjacentes (p.e., o estado 000110 da figura 5). Desta forma, o denominador $1-p_{111111}$ da expressão original deverá ser substituído por $1-p_p$, onde p_p é definida como sendo a probabilidade de perda do sistema (veja formulação no anexo 1), e esta é obtida como a soma das probabilidades dos estados que levam à perda de usuários no sistema. Cada termo é composto pela probabilidade do sistema se encontrar em um estado de perda, multiplicada pela probabilidade do chamado vir de um átomo que não pode ser atendido. Desta forma, a expressão da frequência de despacho f_{nj} passa a ser:

$$f_{nj} = \frac{\lambda_j \sum_{B \in E_{nj}} p_B}{1 - p_p} \quad (4)$$

onde B é o estado do sistema e E_{nj} é o conjunto dos estados nos quais um chamado ao átomo j é atendido pelo servidor n . A fração de despachos do servidor n como *backup*, f_{ln} , é calculada pela expressão abaixo:

$$f_{ln} = \frac{\sum_{\substack{j \in \text{area_de} \\ \text{cobertura_n}}} f_{nj}}{\sum_{j=1}^{N_A} f_{nj}} \quad (5)$$

e a fração de despachos como *backup* no átomo i , f'_{li} , é dada por:

$$f'_{li} = \frac{\sum_{n \neq i} \sum_{\substack{j \in \text{area_de} \\ \text{cobertura_i}}} f_{nj}}{\sum_{n=1}^N \sum_{\substack{j \in \text{area_de} \\ \text{cobertura_i}}} f_{nj}} \quad (6)$$

O tempo médio de viagem do sistema T é obtido pela expressão:

$$T = \sum_{n=1}^N \sum_{j=1}^{N_A} f_{nj} t_{nj} \quad (7)$$

onde t_{nj} é o tempo médio de viagem que o servidor n , quando disponível, gasta para ir ao átomo j (note que t_{nj} é facilmente computado a partir dos tempos médios de viagem entre átomos, τ_{ij}). O tempo médio de viagem de cada servidor n , TU_n , como o sistema não permite filas, é calculado como sendo:

$$TU_n = \frac{\sum_{j=1}^{N_A} f_{nj} t_{nj}}{\sum_{j=1}^{N_A} f_{nj}} \quad (8)$$

e o tempo médio de viagem para cada átomo j , T_j , é dado por:

$$T_j = \frac{\sum_{n=1}^N f_{nj} t_{nj}}{\sum_{n=1}^N f_{nj}} \quad (9)$$

Outras medidas de desempenho podem ser definidas. Para mais detalhes destas e outras medidas de desempenho, o leitor pode consultar LARSON & ODoni (1981) e MENDONÇA (1999).

4. Coleta de Dados e Resultados

Realizamos a coleta de dados em 1997 no escritório central do Anjos do Asfalto no Rio de Janeiro. Conforme sugestão de um dos operadores da empresa, coletamos dados relativos a fins de semana, devido a estes períodos envolverem maior demanda por serviço. Procuramos excluir da coleta fins de semana atípicos, por exemplo, fins de semana prolongados ou contendo feriados, devido à hipótese de dados independentes e identicamente distribuídos. Os dados levantados continham os instantes em que cada servidor foi despachado (com isso pudemos caracterizar o processo de chegada), os instantes em que cada servidor retornou à base (processo de atendimento), além do local do chamado e o servidor que foi alocado. Os resultados obtidos para o processo de chegada são apresentados na tabela 2, e os resultados do processo de atendimento, na tabela 3. Nos dias coletados, não houve demanda no átomo de Volta Redonda 1, assim, consideramos sua taxa média de chegada como aproximadamente nula (símbolo # na tabela 2).

Para cada processo de chegada e atendimento, realizamos testes de hipóteses para verificar a aderência à distribuição exponencial (processo de *Poisson*). Esta verificação foi feita utilizando três testes: teste do qui-quadrado, *Kolmogorov & Smirnov* e *Lillefors* (SHAPIRO & GROSS, 1981; CONOVER, 1971). Os testes foram feitos com níveis de significância de 5%. No processo de chegada, em todos os átomos, não pudemos rejeitar a hipótese de processo de *Poisson*, o que valida a hipótese 2 do modelo hiper-cubo (seção 3). De fato, note na tabela 2 que os coeficientes de variação (razão entre desvio padrão e média) são relativamente próximos de 1. Já para o processo de atendimento, para alguns servidores

foi possível rejeitar a hipótese dos tempos de serviço serem exponencialmente distribuídos. Note que isto não invalida a hipótese 8 do modelo hiper-cubo (seção 3). Consideramos cada servidor com processo de atendimento distinto dos demais, isto é, com taxas médias de serviço $\mu_n, n=1, \dots, N$, diferentes.

Medidas de desempenho

A partir dos dados coletados e das equações de equilíbrio, obtivemos a probabilidade de equilíbrio de cada um dos 64 estados, apresentada no anexo 2. Com essas probabilidades, calculamos a probabilidade de perda p_p conforme expressão do anexo 1, que resultou 0,0502, ou seja, o sistema atende cerca de 95% dos chamados. Note no anexo 2 que a probabilidade do sistema estar completamente vazio, $p_{000000}=0,3092$, é a probabilidade do sistema se encontrar no estado 000000, isto é, todos os servidores estarem disponíveis. O *workload* ρ_n foi calculado para cada servidor n conforme expressão (3) (tabela 4 e figura 6). Podemos perceber que o sistema não está balanceado entre as bases. Enquanto que a base de Volta Redonda fica cerca 30% do tempo ocupada, a base de Lavrinhas fica menos de 15%.

Também calculamos a fração dos atendimentos de um servidor n como *backup*, f_{ln} , (expressão (5)) (tabela 4), e a fração dos atendimentos num átomo j como *backup*, f'_{lj} (expressão (6)) (tabela 5). Isto significa, por exemplo, que 19% dos atendimentos da base de Japeri são *backup* (neste caso, as regiões de Pavuna e a parte leste da região de Caiçara), enquanto que os 81% restantes atendem a própria região.

O tempo médio de viagem do sistema T foi calculado conforme expressão (7), e resultou em 8,6 min. O tempo médio de viagem de um servidor n , TU_n , é o tempo médio que este servidor leva para chegar até o local do atendimento, calculado conforme expressão (8) (tabela 4). Pode-se observar das tabelas 3 e 4 que os tempos médios de viagem são pequenos em relação aos tempos médios de serviço (em geral, em torno de 10%); as variações nos tempos de

Tabela 2 – Processo de chegada de chamadas em cada átomo.

Átomo j	Intervalo Médio (min)	Desvio Padrão (min)	Coefficiente de Variação	Taxa de chegada λ_j (min^{-1})
1 – Pavuna	360,9	290,3	0,8	0,0028
2 – Japeri 1	1192,3	1234,1	1,0	0,0008
3 – Japeri 2	507,3	426,9	0,8	0,0020
4 – Caiçara 1	898,2	690,4	0,8	0,0011
5 – Caiçara 2	561,4	354,6	0,6	0,0018
6 – Volta Redonda 1	#	#	#	#
7 – Volta Redonda 2	266,5	228,8	0,9	0,0038
8 – Penedo 1	544,1	461,2	0,9	0,0018
9 – Penedo - 2	438,8	375,2	0,9	0,0023
10 – Lavrinhas	555,6	683,8	1,2	0,0018

Tabela 3 – Processo de atendimento de chamadas para cada servidor.

Servidor n	Tempo Médio (min)	Desvio Padrão (min)	Coefficiente de Variação	Taxa μ_n (min^{-1})
1 – Pavuna	53,3	32,8	0,6	0,019
2 – Japeri	71,7	48,1	0,7	0,014
3 – Caiçara	60,3	27,7	0,5	0,017
4 – Volta Redonda	98,9	119,0	1,2	0,010
5 – Penedo	41,6	28,5	0,7	0,024
6 – Lavrinhas	79,8	34,9	0,4	0,013

Tabela 4 – *Workload* (expressão (3)), fração de atendimento como *backup* (expressão (5)), tempo médio de viagem do modelo (expressão (8)), tempo médio de viagem da amostra, desvio absoluto e desvio relativo de cada servidor.

Servidor n	<i>Workloads</i> ρ_n	Fração f_n	Tempo médio de viagem do modelo TU_n (min)	Tempo médio de viagem da amostra (min)	Desvio absoluto (min)	Desvio relativo (%)
1 – Pavuna	0,1351	0,0517	4,7	4,9	-0,2	-4,3
2 – Japeri	0,1923	0,1892	7,5	7,2	0,3	4,0
3 – Caiçara	0,1641	0,1089	5,9	5,7	0,2	3,4
4 – V Redonda	0,3037	0,1303	11,5	13,0	-1,5	-13,0
5 – Penedo	0,1834	0,2393	10,5	5,2	5,3	50,4
6 – Lavrinhas	0,1488	0,1771	9,4	8,9	0,5	5,3

viagem contribuem pouco para as variações nos tempos de serviço (veja desvios padrões dos tempos de serviço na tabela 3). Isto está de acordo com a hipótese 9 do modelo hipercubo (seção 3). O tempo médio de viagem no átomo j ,

T_j , é o tempo médio que os servidores levam para chegar neste átomo, calculado conforme expressão (9) (tabela 5). Outras medidas de desempenho deste estudo podem ser encontradas em MENDONÇA (1999).

Tabela 5 – Fração de atendimento como *backup* (expressão (6)) e tempo médio de viagem (expressão (9)) de cada átomo.

Átomo j	Fração f_{ij}^*	Tempo médio de viagem T_j
1 – Pavuna	0,1036	6,2
2 – Japeri 1	0,1617	9,2
3 – Japeri 2	0,1574	5,1
4 – Caiçara 1	0,2135	5,7
5 – Caiçara 2	0,1163	7,0
6 – Volta Redonda 1	0,2599	8,9
7 – Volta Redonda 2	0,2447	14,7
8 – Penedo 1	0,1190	8,7
9 – Penedo 2	0,1505	6,0
10 – Lavrinhas	0,1187	9,4

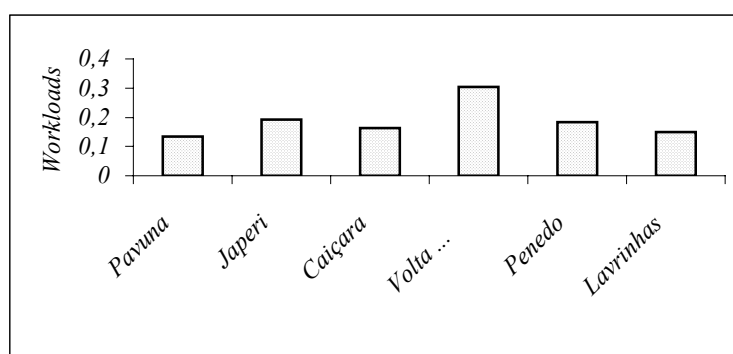


Figura 6 – Gráfico comparativo dos *workloads*.

Para verificar se o modelo está calibrado e tentar validar os resultados produzidos, calculamos os tempos médios de viagem de cada servidor na amostra coletada (tabela 4). As duas últimas colunas da tabela comparam estes tempos com os tempos produzidos pelo modelo. Note que o modelo está razoavelmente calibrado, pois, cinco dos seis servidores têm tempos amostrais suficientemente próximos (desvios de -4,3%, 4,0%, 3,4%, -13,0% e 5,3%) dos tempos produzidos pelo modelo. A exceção é Penedo, cujo desvio (cerca de 50%) pode ser explicado pelo fato da amostra coletada quase não conter casos de *backup*, o que reduziu o tempo amostral, enquanto que o modelo previa 24% de tais ocorrências (conforme tabela 4). Também

acreditamos que, nos fins de semana amostrados, pode ter acontecido do operador dos Anjos do Asfalto ter evitado despachar a ambulância de Penedo para atender chamados de *backup* no átomo 10 (Lavrinhas), devido ao tempo de viagem da base de Penedo até o átomo 10 ser muito longo (28 minutos, conforme MENDONÇA, 1999). Desta maneira, o operador teria preferido, dependendo da emergência do chamado, não respeitar a política de despacho, e alocar este chamado de *backup* para outro sistema de atendimento mais próximo a Lavrinhas, ao invés de despachar a ambulância de Penedo. Infelizmente o sistema não dispõe de dados registrados para podermos validar esta nossa conjectura.

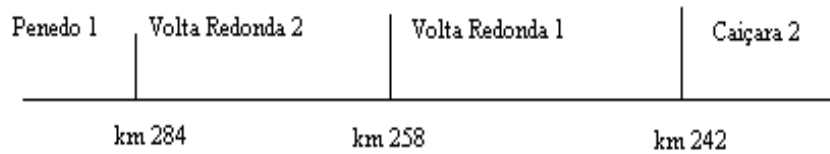


Figura 7 – Configuração original dos Anjos do Asfalto.

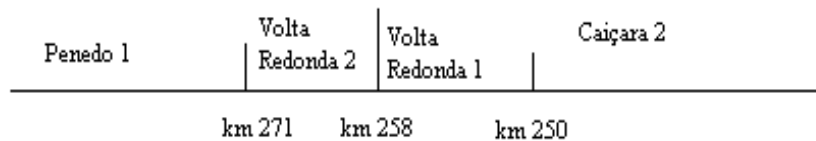


Figura 8 – Configuração modificada para o sistema.

5. Cenários Alternativos

Nesta seção aplicamos o modelo para explorar cenários alternativos, em particular, modificando a configuração do sistema para melhorar o balanceamento do *workload* entre as ambulâncias. Também analisamos os efeitos de variações nos parâmetros do sistema, por exemplo, a demanda dos usuários, para investigar a robustez dos resultados obtidos.

Balaceamento

Como podemos perceber na seção 4, o sistema não está balanceado, pois o servidor de Volta Redonda trabalha bem mais que os demais. Resolvemos então estudar um cenário alternativo, que consistia na redução no tamanho dos átomos em que o servidor de Volta Redonda é prioritário, e na incorporação dessas metades nos átomos adjacentes. Esta modificação pode ser vista nas figuras 7 e 8.

A tabela 6 compara algumas medidas do sistema original e o sistema modificado. O tempo médio de viagem T no sistema original era de 8,6 min., já no caso do sistema modificado, esse tempo caiu para 7,7 min. (expressão (7)). A probabilidade de perda p_p no sistema original era de 0,0502, já no caso do sistema modificado, essa probabilidade caiu para 0,0439

(conforme anexo 1). Pode-se notar na tabela 6 e na figura 9 que a modificação no sistema original balanceia melhor o sistema, reduzindo a diferença entre os *workloads* dos servidores, ou seja, com menor desvio padrão. Note na tabela 6 e nas figuras 9 e 10 que essa alteração no tamanho do átomo de Volta Redonda provocou alterações globais, e não apenas nas vizinhanças de Volta Redonda.

Análise de sensibilidade

Outra análise do sistema que foi feita com o auxílio do modelo foi verificar como a atual configuração dos Anjos do Asfalto responderia a um aumento da demanda. Para isto, acrescentou-se às taxas médias de chegada λ_j , $j=1, \dots, N_A$, incrementos de 10%, 20%, 30%, 40%, 50%, 100%, 200% e 300%. Os resultados obtidos podem ser vistos na tabela 7 e nas figuras 11-13. Outros parâmetros do modelo hipercubo, além da taxa média de chegadas, também poderiam ter sido alterados. O método de análise seria similar.

Como podemos observar na tabela 7 e figura 13, o tempo médio de viagem do sistema é pouco sensível com relação ao aumento de demanda. Isto ocorre porque nesse sistema não há filas. Dessa forma, o chamado que encontra o sistema congestionado não espera em fila, o que não aumenta muito o tempo médio de viagem do

Tabela 6 – Comparação entre os *workloads* (expressão (3)), fração de atendimento como *backup* (expressão (5)) e tempos médios de viagem (expressão (8)) dos sistemas original e modificado.

Servidor n	<i>Workloads</i> ρ_n		Fração de atendimento como <i>backup</i> f_n		Tempos médios de viagem TU_n	
	Sistema original	Sistema modificado	Sistema original	Sistema modificado	Sistema original	Sistema modificado
1 – Pavuna	0,1351	0,1346	0,0517	0,0497	4,7	4,7
2 – Japeri	0,1923	0,1848	0,1892	0,1650	7,5	7,3
3 – Caiçara	0,1641	0,1199	0,1089	0,1064	5,9	7,5
4 – V Redonda	0,3037	0,2222	0,1303	0,3287	11,5	10,6
5 – Penedo	0,1834	0,2155	0,2393	0,0952	10,5	7,5
6 – Lavrinhas	0,1488	0,1527	0,1771	0,2009	9,4	9,8

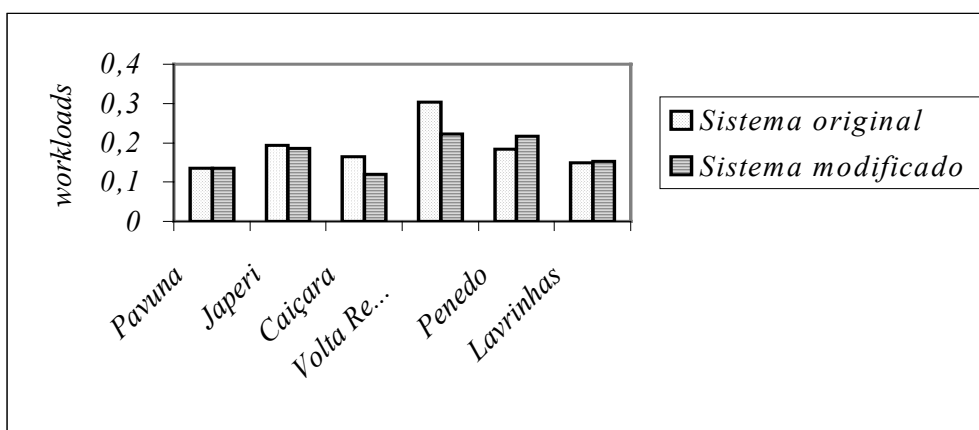


Figura 9 – Comparativo entre os *workloads* dos sistemas original e modificado.

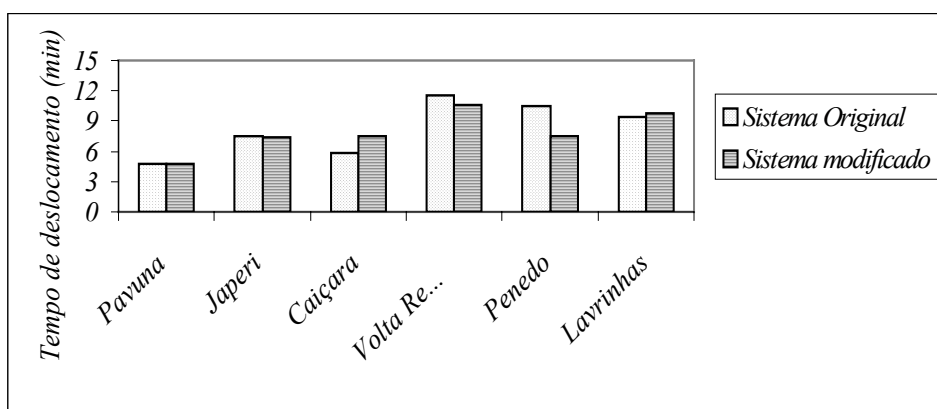


Figura 10 – Comparação entre os tempos médios de viagem de cada servidor para os sistemas original e modificado.

Tabela 7 – Impacto do aumento da demanda sobre a probabilidade de sistema vazio (p_{000000}), probabilidade de perda (p_p), fração dos atendimentos como *backup*, tempo médio de viagem do sistema (expressão (7)).

Aumento da demanda (%)	Probabilidade do sistema vazio p_{000000}	Probabilidade de perda p_p	Fração dos atendimentos como <i>backup</i> f_i	Tempo médio de viagem do sistema T (min)
0	0,3090	0,0502	0,1596	8,6
10	0,2762	0,0590	0,1704	8,8
20	0,2463	0,0680	0,1814	9,9
30	0,2206	0,0768	0,1910	9,1
40	0,1967	0,0866	0,2005	9,2
50	0,1768	0,0959	0,2093	9,4
100	0,1057	0,1441	0,2461	10,0
200	0,0409	0,2382	0,2975	10,8
300	0,0011	0,5438	0,3724	11,2

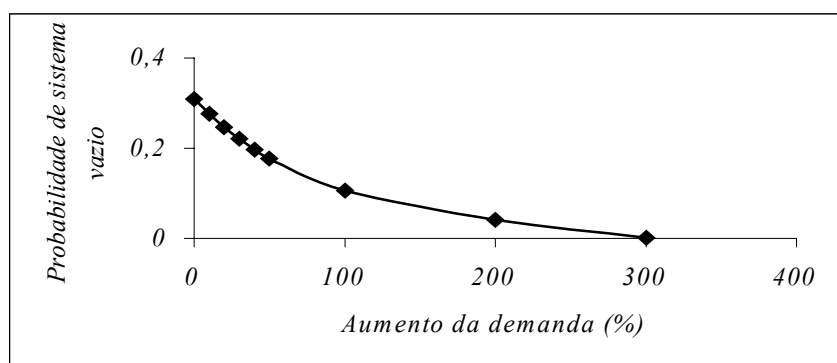


Figura 11 – Impacto do aumento da demanda sobre a probabilidade de sistema ocioso.

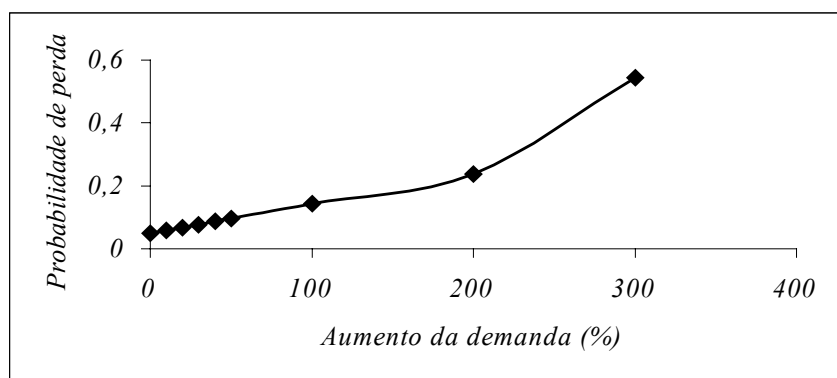


Figura 12 – Impacto do aumento da demanda sobre a probabilidade de perda.

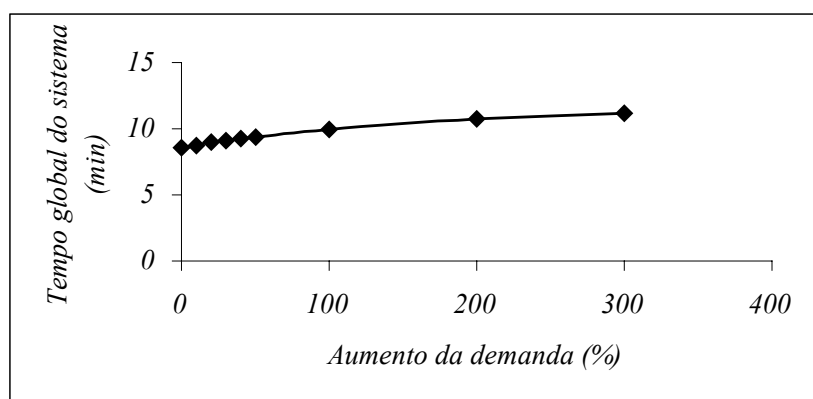


Figura 13 – Gráfico do impacto do aumento da demanda no tempo médio de viagem do sistema.

sistema para os chamados que conseguem ser atendidos. Note que, ao dobrarmos a demanda, o tempo médio de viagem cresce de 8,6 min. para apenas 10 min., no entanto, a perda de chamadas sobe de 5% para mais de 14%, e a proporção de tempo que o sistema fica completamente vazio cai de 31% para 11%.

6. Conclusões e Perspectivas

Neste trabalho mostramos como o modelo hipercubo pode ser adaptado e aplicado para analisar um sistema de atendimento médico-emergencial em rodovia, como os Anjos do Asfalto. Os resultados indicaram que os Anjos do Asfalto têm prestado um bom nível de atendimento aos seus usuários, por exemplo, o tempo médio de resposta é de apenas 10,6 minutos, isto é, 8,6 min. de tempo de viagem, mais cerca de 2 min. de tempo de *setup*. Convém salientar que esse tempo de resposta é próximo de outros estudos reportados na literatura, por exemplo, VOLZ (1971) e FUJIWARA (1987) reportam 8,8 min. Para a configuração atual do sistema, esse resultado parece ser robusto, isto é, esse tempo de resposta é pouco sensível a aumentos de demanda, mesmo quando dobramos a demanda, principalmente porque o sistema não permite a formação de filas de chamados.

O sistema trabalha com uma alta ociosidade (em quase 31% do tempo, o sistema permanece

completamente vazio; veja p_{000000} no anexo 2), e há uma baixa probabilidade de perda de chamados (cerca de 5%). No entanto, os Anjos do Asfalto poderiam balancear melhor a carga de trabalho entre suas ambulâncias, por exemplo, por meio de uma pequena mudança na definição dos seus átomos conforme ilustrado na seção 5, sem investimentos ou custos adicionais e sem maiores prejuízos para as demais medidas de desempenho. Uma perspectiva interessante para pesquisa futura seria integrar o modelo hipercubo num modelo de otimização, para determinar como o tamanho dos átomos deve ser alterado, de maneira a maximizar o balanceamento do sistema. Outra pesquisa futura interessante seria realizar estudos de caso em cidades brasileiras, similarmente ao que foi feito neste trabalho para o caso de uma rodovia, para analisar a viabilidade de aplicação do modelo hipercubo para apoiar decisões estratégicas e operacionais nestes sistemas urbanos.

Agradecimentos

Os autores são gratos aos três revisores anônimos e à doutoranda Renata Algisi Takeda pelos úteis comentários e sugestões. Também agradecemos ao sistema Anjos do Asfalto pelo fornecimento das informações.

Anexo 1 – Probabilidade de perda no sistema.

$$\begin{aligned}
 p_p = & ((\lambda_1 + \lambda_2)/\lambda) p_{000011} + ((\lambda_3 + \lambda_4)/\lambda) p_{000110} + ((\lambda_5 + \lambda_6)/\lambda) p_{001100} + ((\lambda_7 + \lambda_8)/\lambda) p_{011000} + ((\lambda_9 + \lambda_{10})/\lambda) \\
 & p_{110000} + ((\lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})/\lambda) p_{111000} + ((\lambda_9 + \lambda_{10})/\lambda) p_{110001} + ((\lambda_9 + \lambda_{10})/\lambda) p_{110010} + ((\lambda_9 + \lambda_{10})/\lambda) \\
 & p_{110100} + ((\lambda_5 + \lambda_6)/\lambda) p_{101100} + ((\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8)/\lambda) p_{011100} + ((\lambda_7 + \lambda_8)/\lambda) p_{011010} + ((\lambda_3 + \lambda_4)/\lambda) \\
 & p_{100110} + ((\lambda_3 + \lambda_4)/\lambda) p_{010110} + ((\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6)/\lambda) p_{001110} + ((\lambda_7 + \lambda_8)/\lambda) p_{011001} + ((\lambda_5 + \lambda_6)/\lambda) \\
 & p_{001101} + ((\lambda_1 + \lambda_2)/\lambda) p_{100011} + ((\lambda_1 + \lambda_2)/\lambda) p_{010011} + ((\lambda_1 + \lambda_2)/\lambda) p_{001011} + ((\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)/\lambda) \\
 & p_{000111} + ((\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})/\lambda) p_{111100} + ((\lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})/\lambda) p_{111010} + ((\lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})/\lambda) \\
 & p_{111001} + ((\lambda_3 + \lambda_4 + \lambda_9 + \lambda_{10})/\lambda) p_{110110} + ((\lambda_9 + \lambda_{10})/\lambda) p_{110101} + ((\lambda_1 + \lambda_2 + \lambda_9 + \lambda_{10})/\lambda) p_{110011} + \\
 & ((\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8)/\lambda) p_{011110} + ((\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6)/\lambda) p_{101110} + ((\lambda_5 + \lambda_6)/\lambda) p_{101101} + \\
 & ((\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8)/\lambda) p_{011101} + ((\lambda_1 + \lambda_2)/\lambda) p_{101011} + ((\lambda_1 + \lambda_2 + \lambda_7 + \lambda_8)/\lambda) p_{011011} + \\
 & ((\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)/\lambda) p_{100111} + ((\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4)/\lambda) p_{010111} + ((\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6)/\lambda) p_{001111} + \\
 & ((\lambda(\lambda_1 + \lambda_2))/\lambda) p_{111110} + ((\lambda(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4))/\lambda) p_{111101} + ((\lambda(\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6))/\lambda) p_{111011} + \\
 & ((\lambda(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8))/\lambda) p_{110111} + ((\lambda(\lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10}))/\lambda) p_{101111} + ((\lambda(\lambda_9 + \lambda_{10}))/\lambda) p_{011111} + p_{111111}.
 \end{aligned}$$

onde $\lambda = \sum_{j=1}^{N_A} \lambda_j$.

Anexo 2 - Probabilidade de equilíbrio dos 64 estados possíveis para o sistema.

$p_{000000}=0,3090$	$p_{010000}=0,0529$	$p_{100000}=0,0483$	$p_{110000}=0,0139$
$p_{000001}=0,0435$	$p_{010001}=0,0075$	$p_{100001}=0,0068$	$p_{110001}=0,0020$
$p_{000010}=0,0633$	$p_{010010}=0,0109$	$p_{100010}=0,0099$	$p_{110010}=0,0028$
$p_{000011}=0,0142$	$p_{010011}=0,0024$	$p_{100011}=0,0022$	$p_{110011}=0,0006$
$p_{000100}=0,0524$	$p_{010100}=0,0090$	$p_{100100}=0,0082$	$p_{110100}=0,0024$
$p_{000101}=0,0075$	$p_{010101}=0,0013$	$p_{100101}=0,0012$	$p_{110101}=0,0003$
$p_{000110}=0,0166$	$p_{010110}=0,0029$	$p_{100110}=0,0026$	$p_{110110}=0,0007$
$p_{000111}=0,0036$	$p_{010111}=0,0006$	$p_{100111}=0,0006$	$p_{110111}=0,0002$
$p_{001000}=0,1148$	$p_{011000}=0,0345$	$p_{101000}=0,0186$	$p_{111000}=0,0086$
$p_{001001}=0,0162$	$p_{011001}=0,0049$	$p_{101001}=0,0026$	$p_{111001}=0,0012$
$p_{001010}=0,0237$	$p_{011010}=0,0071$	$p_{101010}=0,0038$	$p_{111010}=0,0018$
$p_{001011}=0,0053$	$p_{011011}=0,0016$	$p_{101011}=0,0009$	$p_{111011}=0,0004$
$p_{001100}=0,0232$	$p_{011100}=0,0068$	$p_{101100}=0,0037$	$p_{111100}=0,0017$
$p_{001101}=0,0033$	$p_{011101}=0,0010$	$p_{101101}=0,0005$	$p_{111101}=0,0002$
$p_{001110}=0,0073$	$p_{011110}=0,0021$	$p_{101110}=0,0012$	$p_{111110}=0,0005$
$p_{001111}=0,0016$	$p_{011111}=0,0005$	$p_{101111}=0,0003$	$p_{111111}=0,0001$

Referências Bibliográficas

ALBINO, J.C.C.: *Quantificação e localização de unidades móveis de atendimento de emergência e interrupções em redes de distribuição de energia elétrica: aplicação do Modelo Hipercubo*, Dissertação de Mestrado, Departamento de Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, SC, 1994.

BITRAN, G. & MORABITO, R.: “An overview of tradeoff curves in manufacturing system design”, *Production and operations management*, v.8, n.1, p.56-75, 1999.

- BRANDEAU, M.L. & LARSON, R.C.:** “Extending and applying the hiper-cube model to deploy ambulances in Boston”. In *Delivery of Urban services* (ed. SWERSEY, A. & IGNALL, E.) TIMS Studies in the Management Sciences 22, Elsevier, Amsterdam, 1986.
- BURWELL, T.H.; JARVIS, J.P. & McKNEW, M.A.:** “Modeling co-located servers and dispatch ties in the hiper-cube model”, *Computers and Operations Research*, v.20, n.2, p.113-119, 1993.
- CONOVER, J.:** *Practical nonparametric statistics*, John Wiley & Sons, New York, 1971.
- EATON, D.J. et al.:** “Determining emergency medical service vehicle deployment in Austin, Texas”, *Interfaces*, v.15, n.1., p.96-108, 1985.
- FONTANELLA, G.C. & MORABITO, R.:** “Modelagem por meio de teoria de filas do *tradeoff* entre investir em canais de atendimento e satisfazer o nível de serviço em provedores *internet*”, *Gestão & Produção*, v.4, n.3, p.278-295, 1997.
- FUJIWARA, O.:** “Ambulance deployment analyses: a case study of Bangkok”, *European Journal of Operations Research*, v.31, p.9-18, 1987.
- GOLDBERG, J. & PAZ, L.:** “Location of emergency vehicle bases when service times depends on call location”, *Transportation Science*, v.25, n.4, p.264-280, 1991.
- GONÇALVES, M.B.; NOVAES, A.G.N. & ALBINO, J.C.C.:** “Modelos de localização de serviços emergenciais em rodovias”. *Anais do XXVI Simpósio Brasileiro de Pesquisa Operacional*, v.1, Florianópolis, SC, p.591-596, 1994.
- GONÇALVES, M.B., NOVAES, A.G.N. & SCHMITZ, R.:** “Um modelo de otimização para localizar unidades de serviços emergenciais em rodovias”. *Anais do IX ANPET – Congresso de Pesquisa e Ensino em Transportes*, São Carlos, SP, p.962-972, 1995.
- HALPERN, J.:** “Accuracy of estimates for the performance criteria in certain emergency service queueing systems”, *Transportation Science*, v.11, n.3. p.223-242, 1977
- LARSON, R.C.:** “A Hiper-cube queueing Model for facility location and redistricting in urban emergency services”, *Computers and operations research*, v.1, p.67-95. 1974.
- LARSON, R.C. & ODoni, A.R.:** *Urban operations research*, Prentice Hall, New Jersey, 1981.
- MENDONÇA, F.C.:** “Aplicação do Modelo Hiper-cubo, baseado em Teoria de Filas, para análise de um sistema médico-emergencial em rodovia”, *Dissertação de Mestrado*, Departamento de Engenharia de Produção, Universidade Federal de São Carlos, São Carlos, SP, 1999.
- MORABITO, R. & LIMA, F.C.:** “Um modelo para analisar o problema de filas em caixas de supermercados: Um estudo de caso”, aceito para publicação na *Pesquisa Operacional*, 1999.
- SACKS, S. R. & GRIEF, S.:** “Orlando Police Department uses OR/MS methodology, new software to design patrol districts”, *OR/MS Today*, p.30-32, Fev. 1994.
- SHAPIRO, S.S. & GROSS, A.J.:** *Statistical modeling techniques*. Marcel Dekker, New York, 1981.
- SWERSEY, A.J.:** “The deployment of police, fire and emergency medical units”. In *Handbooks in OR & MS*, v.6 (ed. POLLOCK, S.M. et al.), Elsevier Science B.V., Amsterdam, 151-200, 1994.
- VOLZ, R.A.:** “Optimum ambulance location in semi-rurals areas”, *Transportation Science*, v.5, n.2, p.193-203, 1971.

***AN APPLICATION OF THE HYPERCUBE MODEL TO ANALYZE AN
EMERGENCY MEDICAL SYSTEM ON A HIGHWAY***

Abstract

In an emergency medical system, the quality of the service is fundamental since this might be the difference between life and death of the user. Beyond the attendance quality, one of the most important components of the service level is the mean response time to an emergency call. However, acquiring equipment and training for the team to reduce this time imply additional costs and investments. In this case, an important trade-off appears between the service level and the investment in the system. To analyze this trade-off, there is a well-known model named hypercube, which is based in spatially distributed queueing theory. In this paper, we adapt and apply the hypercube model to analyze the “Anjos do Asfalto”, an emergency medical system on the highway Presidente Dutra between the cities of Rio de Janeiro and São Paulo, in Brazil. The results showed that the hypercube model can be an effective tool to support strategic and operational decisions of emergency systems on highways.

Key words: emergency medical system, queueing theory, hypercube model.