

Modelo de apoio à decisão para um problema de posicionamento de bases, alocação e realocação de ambulâncias em centros urbanos: estudo de caso no município de São Paulo

Luiz Augusto Andrade¹, Claudio Barbieri da Cunha²

Resumo: Este artigo apresenta uma proposta de modelo matemático para o problema de localização de bases de atendimento emergencial, alocação de ambulâncias a essas bases em múltiplos períodos de tempo num horizonte de planejamento definido e realocação das viaturas entre períodos subsequentes. Esse problema é relevante para planejamento de sistemas de atendimento emergencial em grandes centros urbanos, nos quais existem variações das condições de tráfego e da concentração de pessoas em diferentes zonas ao longo do dia, fazendo com que os sistemas emergenciais nesses centros precisem ser dinâmicos o suficiente para acompanhar essas variações. Adicionalmente, em grandes metrópoles o número de ambulâncias é elevado (superior a uma centena), assim como o número de distritos em que a cidade é dividida. Como objetivo do planejamento tem-se a maximização de probabilidade de atendimento de um determinado chamado dentro de um tempo máximo de cobertura pré-definido. Neste artigo também é apresentada uma aplicação prática do modelo no sistema de ambulâncias do município de São Paulo.

Palavras-chave: Serviço de ambulâncias. Problema de localização. Modelo de otimização.

Abstract: In this article a mathematical formulation for the problem of base location, ambulance allocation and relocation in multiple periods of time in a planning horizon is proposed. This problem is relevant for the planning of emergency services, especially in large urban centers where traffic conditions and population's concentration change during the day. These characteristics lead to the need of such services being dynamic enough to adjust to the change of city conditions in terms of traffic speeds and demand; in addition, the number of ambulances is usually elevated, as well as the number of districts in which the city is divided. Thus, the proposed model aims to maximize the probability of one determined call being served within a given covering time. We also describe a real world application for São Paulo's São Paulo's emergency system.

Keywords: Emergency medical services. Location problem. Optimization model.

1 INTRODUÇÃO

O serviço de atendimento de urgência, ou emergência, compreende os primeiros socorros e a remoção de pacientes sujeitos a acidentes, traumas e outras ocorrências médicas que podem representar risco a vidas humanas. Busca-se oferecer um serviço que maximize a probabilidade de sobrevivência dos socorridos, desde o acontecimento da situação de risco até a entrada do paciente em uma unidade de saúde especializada. Todo o trabalho é realizado por veículos de transporte e suporte à vida.

As chances de sobrevivência de um indi-

víduo que necessita de atendimento emergencial, devido a acidente ou outra ocorrência, aumentam com a diminuição do tempo de resposta, que é o tempo gasto entre o acontecimento do acidente e o momento da chegada de uma viatura de socorro. Uma parte importante deste tempo é o tempo de deslocamento da viatura de uma base até o local da ocorrência.

Conforme apontam Takeda et al. (2007), este tempo é afetado por diversos fatores, dentre os quais as condições de tráfego local, dia da semana e período do dia, tipo e número de veículos disponíveis e localização destes veículos. Ainda segundo os autores, a regulamentação americana para os serviços médicos de urgência estabelece que 95% das solicitações em área urbana devem ser atendidas em, no máximo, 10 minutos, sendo este período estendido para 30 minutos para a área rural (Ball & Lin, 1993). Nos casos dos serviços de ambulâncias das cidades de Londres e Montreal, 95% das solicitações devem ser servidas em 14 e 10 minutos, respectivamente, e 50% e 70% das solicitações devem ser servidas em 8

¹ Escola Politécnica da Universidade de São Paulo - Programa de Mestrado em Engenharia de Sistemas Logísticos (luizacga@gmail.com).

² Escola Politécnica da Universidade de São Paulo - Programa de Mestrado em Engenharia de Sistemas Logísticos (cbcunha@usp.br).

Manuscrito recebido em 24/09/2013 e aprovado para publicação em 29/04/2014. Este artigo é parte de TRANSPORTES v. 22, n. 2, 2014. ISSN: 2237-1346 (online).

DOI: <http://dx.doi.org/10.14295/transportes.v22i2.730>.

e 7 minutos, respectivamente (Geandreau et al., 2001; Galvão et al., 2003). No Brasil, entretanto, ainda não há legislação específica que estabeleça valores máximos para os tempos de resposta.

Um requisito importante desses sistemas é o planejamento da rede de atendimento, definida pelas localizações das bases de veículos e pelas alocações de viaturas de atendimento a essas bases, o que por sua vez impacta o tempo de deslocamento entre as viaturas e os locais dos acidentes. A operação desses sistemas é ainda mais crítica em grandes centros urbanos, nos quais as condições de trânsito e os padrões de variação da demanda por atendimento emergencial resultam numa maior dificuldade de realizar os atendimentos dentro de tempos de resposta pequenos.

Este trabalho trata do problema de planejamento das localizações de bases e alocação de viaturas ao longo de um período de planejamento, considerando características dinâmicas de variação espaço-temporal das demandas e dos tempos de deslocamento em centros urbanos, e também levando em consideração aspectos estocásticos do atendimento emergencial.

Mais especificamente, propõe-se uma ferramenta de planejamento, baseada em um modelo matemático, para os gestores de serviços de atendimento móvel pré-hospitalar de urgência, que auxilia na localização de bases e na alocação de veículos ao longo de um horizonte de tempo. Busca-se determinar a malha de atendimento que maximize a probabilidade dos chamados serem atendido dentro de um tempo de deslocamento pré-estabelecido, considerando aspectos dinâmicos e estocásticos do problema de atendimento emergencial, bem como as dimensões típicas do problema, dadas pelo número total de ambulâncias e de distritos em que a área de interesse é dividida. A ferramenta proposta permite considerar explicitamente a possibilidade de realocações de veículos ao longo do período de planejamento, de forma que a frota de ambulâncias possa acompanhar as variações espaço-temporais dos padrões de demanda e dos tempos de deslocamento entre as diversas partes da região de interesse a ser atendida.

A próxima seção corresponde a uma revisão bibliográfica de modelos matemáticos para problemas de localização de ambulâncias e instalações de atendimento emergencial. A seção 3

contempla uma descrição detalhada do problema tratado neste artigo e a seção 4 a formalização do modelo matemático correspondente que é proposto. Já a seção 5 descreve uma aplicação do mesmo a um estudo de caso no município de São Paulo. Na sexta e última seção são feitas algumas considerações sobre os resultados obtidos no estudo de caso e sobre a validade do modelo proposto, além de serem discutidas possíveis frentes de pesquisa futura.

2 REVISÃO BIBLIOGRÁFICA

Os problemas de localização de veículos para atendimento de emergências ocorrem em muitos casos práticos, por exemplo: localização de veículos do corpo de bombeiros, veículos de apoio mecânico, veículos de suporte médico e embarcações para atendimento de acidentes marítimos (Medina, 1996).

Os problemas de localização de ambulâncias estão, em geral, definidos em grafos não direcionados com pontos de demanda e pontos candidatos a receberem bases ou viaturas (Daskin, 1995). Nos casos reais, a demanda por serviços de atendimento de emergência é distribuída geograficamente numa região; contudo, na resolução de problemas desse tipo, o que geralmente se faz é determinar o nível de agregação de demandas que se deseja (por distritos ou por bairros, por exemplo) e acumular a demanda de cada subdivisão num único ponto, sendo esse ponto tratado matematicamente no grafo do problema.

Na definição de problemas de localização de bases de veículos de emergência, usualmente assume-se que determinado ponto de demanda é coberto se ele pode ser atendido num intervalo de tempo máximo pré-estabelecido. Segundo Rajagopalan et al. (2008), essa noção de cobertura é amplamente aceita e inclusive utilizada como meio de definição de níveis de serviço. A demanda dos pontos é definida genericamente como um número de ocorrências por unidade de tempo, originadas dentro do distrito representado pelo ponto. Alguns autores definem a demanda como um número de ocorrências médio tomado num horizonte de tempo suficientemente grande; outros ainda definem a demanda como uma frequência de ocorrências computada num período de análise.

Toregas et al. (1971) e Church e ReVelle

(1974) trataram do caso mais simples do problema de localização de bases de ambulâncias, em que se considera um conjunto de pontos de demanda i e um conjunto de pontos candidatos j dispostos num grafo. Cada arco entre quaisquer pontos i e j representa o tempo de deslocamento entre esses pontos. Em ambos os trabalhos define-se um tempo máximo de atendimento S , acima do qual uma viatura localizada em j não consegue cobrir um ponto i adequadamente. Toregas et al. (1971) formalizam o problema como sendo o de encontrar o menor número possível de viaturas necessário para que todos os pontos de demanda sejam cobertos. O modelo resultante foi denominado *Location Set Covering Model* (LSCM), que é o modelo clássico do conjunto de cobertura aplicado ao caso do posicionamento de ambulâncias.

Do ponto de vista dos planejadores de sistemas de ambulâncias, a quantidade de recursos é limitada e, portanto, um parâmetro de entrada do problema. Uma alternativa para a formulação LSCM foi proposta por Church e ReVelle (1974), chamada de *Maximal Covering Location Problem* (MCLP). Sendo fixo e conhecido o número de instalações que se deseja posicionar, o MCLP busca maximizar a demanda coberta por essa quantidade pré-definida.

Grande parte do desenvolvimento posterior dos modelos para o problema de localização de ambulâncias foi baseado nesses dois artigos. A abordagem de Toregas et al. (1971) resulta, em geral, em um número muito grande de viaturas, o que do ponto de vista prático é inviável, dadas as restrições orçamentárias dos sistemas de ambulâncias, ao passo que a abordagem proposta por Church e ReVelle (1974) é mais condizente com as restrições enfrentadas pelos planejadores dos sistemas de ambulância, e como consequência, os modelos posteriormente desenvolvidos aderem mais a essa segunda vertente de modelagem.

Como apontado por Brotcorne et al. (2003), esses modelos matemáticos mais antigos, propostos para o problema de localização de ambulâncias consideram definições muito restritas e genéricas para o problema. Existem extensões desses dois modelos que consideram aspectos mais realistas do problema, como por exemplo, o fato da localização de bases ser independente da localização de viaturas, ou o fato de existirem diferentes tipos de veículos com tempos máximos de cobertura distintos; alguns

modelos ainda introduziram o conceito de cobertura múltipla, tal que um ponto de demanda é atendido se ele é coberto por mais de uma viatura.

Na formulação proposta por Schilling et al. (1979), denominada *Tandem Equipment Allocation Model* (TEAM), que pode ser considerada como extensão do modelo MCLP, é possível considerar dois tipos ou categorias de veículos de emergência: básicos e avançados. No entanto, esse modelo não distingue entre a localização das bases e a localização dos veículos em si, e considera que um veículo avançado só pode ser posicionado num ponto candidato caso nesse ponto também seja posicionado um veículo básico. Nesse caso a localização de bases e veículos é feita de maneira conjunta, de modo que bases devem ser posicionadas em pontos que devem conter veículos.

Outra extensão do MCLP também desenvolvida por Schilling et al. (1979) é o modelo *Facility-Location Equipment-Emplacement Technique* (FLEET), criado para a localização de bases e veículos do corpo de bombeiros. Apesar de ser um modelo desenvolvido para a solução de problemas de bases e veículos de combate a incêndios, seus conceitos se aplicam ao problema de localização de ambulâncias. Um ponto do modelo FLEET que difere do modelo TEAM, é que no primeiro não existe hierarquia entre os veículos, porém é considerada explicitamente na modelagem uma hierarquia entre as bases e os veículos. Uma revisão detalhada de outros modelos determinísticos e probabilísticos para o problema de localização de ambulâncias pode ser encontrada em Schilling et al. (1993).

Levando em consideração a estocasticidade do processo de geração de demanda e do processo de atendimento dos acidentados, modelos probabilísticos foram também propostos com o intuito de aproximar os modelos matemáticos à realidade do problema, especialmente no tocante à aleatoriedade dos chamados e à indisponibilidade de veículos. Mais especificamente, os modelos determinísticos mencionados anteriormente não consideram uma característica importante do problema de localização de ambulâncias: a possibilidade de um ponto não ser atendido, pois o veículo que garantia a sua cobertura está alocado a um chamado. Um modelo probabilístico que considera essa situação foi proposto por Daskin (1983), denominado *Maximum Expected Covering Location Problem*

(MEXCLP). Nessa modelagem, considera-se que uma ambulância genérica possui uma probabilidade q , denominada fração de ocupação (*busy fraction*), de estar indisponível para atendimento. Os autores assumem que cada ambulância opera independentemente das demais e assumem que a fração de ocupação é igual para todas as ambulâncias do sistema e independente do estado do sistema, ou seja, independe de quantas ambulâncias estão ocupadas num determinado momento.

O modelo MEXCLP fornece meios para localizar apenas um tipo de veículo e não considera a localização de bases de veículos separadamente. Bianchi e Church (1988) desenvolveram um modelo híbrido entre os modelos FLEET e MEXCLP, denominado *Multiple cover, One unit, FLEET problem* (MOFLEET). Esse modelo, além de se tratar de uma formulação probabilística para o problema, considera explicitamente a separação entre a localização de bases e ambulâncias. Contudo, a formulação do MOFLEET não permite a localização de múltiplos tipos de veículos, algo que foi desenvolvido por Jayaraman e Srivastava (1995). Para localizar múltiplas instalações e veículos os autores desenvolveram um modelo probabilístico chamado *Multiple Equipment Multiple Cover Facility Location Allocation Problem* (MEMCOLA), o qual possibilita a localização de bases e dois tipos de veículos, cada qual com uma fração de ocupação específica.

Outros modelos probabilísticos foram propostos por ReVelle e Hogan (1989). Os autores formularam dois modelos chamados *Maximum Availability Location Problem I e II* (MALP I e MALP II). Assim como o MEXCLP, o modelo MALP I considera que a fração de ocupação q é a mesma para todos os pontos candidatos j e consequentemente igual e independente para todos os veículos. Sendo assim, pode-se calcular o número mínimo de ambulâncias necessárias para cobrir um ponto de demanda i com uma probabilidade α . A formulação do MALP I considera esse valor explicitamente em sua formulação, e busca maximizar a demanda coberta com uma probabilidade α .

Na formulação do MALP II, a premissa de frações de ocupação idênticas para todos os pontos candidatos não é tomada. Em vez disso, os autores associam uma fração de ocupação q_i para cada ponto i , que corresponde à probabilidade de uma ambulância localizada na vizinhança do ponto i estar ocupada, sendo que vizinhança de i é o subconjunto de pontos localizados a menos de um raio máximo de cobertura em relação ao ponto. Essas frações de ocupação locais proporcionam estimativas mais realistas da probabilidade de uma ambulância randomicamente selecionada estar ocupada. Dessa maneira, calcula-se para cada ponto i um número mínimo de ambulâncias necessárias b_i para que o ponto seja coberto com probabilidade α .

Um avanço maior nos conceitos presentes nos modelos MALP I e MALP II foi proposto por Marianov e ReVelle (1996), que consideram uma vizinhança de um ponto de demanda i como um sistema isolado com demandas e servidores funcionando num sistema de filas do tipo *M/G/s-loss*. São utilizados resultados da Teoria das Filas para fornecer melhores estimativas das frações de ocupação q_i . Esse modelo foi denominado *Queuing Maximal Availability Location Problem* (Q-MALP).

Uma abordagem probabilística do modelo FLEET também foi formulada por ReVelle e Marianov (1991). O *Probabilistic FLEET model* (P-FLEET) procura localizar bases, caminhões e bombas independentemente, de maneira a buscar uma maximização da cobertura da demanda com probabilidade α . Para cada ponto de demanda i , calculam-se frações de ocupação locais para os diferentes veículos e , com base nesses valores, calcula-se o número de servidores necessários para cobrir o ponto i com probabilidade α , para cada tipo de veículo. O P-FLEET é um modelo bastante completo e possui uma característica interessante para representar o problema prático: um ponto é considerado atendido somente se o mesmo for coberto com probabilidade α por mais de um tipo de veículo. Ele possui a desvantagem de permitir a alocação de apenas um veículo de cada tipo por base. Os autores apresentam também uma formulação alternativa para o P-FLEET, denominada *Probabilistic Facility-Location Equipment-Emplacement Technique with Multiple Co-location* (P-FLEET-MC), a qual permite relaxar essa restrição, possibilitando a localização de múltiplos veículos por base.

Outros modelos probabilísticos foram propostos baseados no Modelo Hipercubo (LARSON, 1974) que permite um tratamento detalhado das características estocásticas do problema. Conforme Galvão et al. (2005), o Modelo Hipercubo não corresponde a um modelo de oti-

mização, mas sim um modelo que possibilita a análise de diversos parâmetros de desempenho de um sistema, dado um cenário de localização e regras de despacho de veículos.

Dentre as aplicações práticas do Modelo Hiper cubo, pode-se citar Takeda et al. (2007). Os autores utilizaram o modelo para avaliação de cenários de posicionamento de viaturas no município de Campinas. Os cenários consideravam diversas alternativas de localização de ambulâncias em bases espalhadas na cidade, desde um cenário inteiramente centralizado (todas as viaturas posicionadas no distrito central do município) até mesmo a descentralização completa do sistema, considerando 10 áreas geradoras de demanda (denominados átomos geográficos) e 10 ambulâncias. Já Iannoni e Morabito (2007) desenvolveram uma aplicação do modelo para avaliação do posicionamento de ambulâncias ao longo de uma rodovia. Em ambos os trabalhos citados, o modelo permite uma análise minuciosa de indicadores de desempenho do sistema, o que subsidia a tomada de decisão pelos planejadores públicos. Uma extensão desse modelo foi proposta por Iannoni *et. al.* (2007) que consideraram a localização das bases de ambulância em rodovias em conjunto com a divisão das rodovias em segmentos que definem áreas de cobertura ou atendimento. Os experimentos realizados consideraram instâncias com até seis ambulâncias e 8 átomos geográficos.

Uma referência que cobre em detalhes o Modelo Hiper cubo, bem como extensões e desdobramentos possíveis é encontrada em Chiyoshi et al. (2011). Os autores, por meio de modelos reduzidos (tradução livre de *Toy Models*), apresentam formas de modelagem de diversas situações práticas encontradas nos sistemas de emergência.

Dentre as vantagens do uso do Modelo Hiper cubo para análise de sistemas de ambulâncias estão: (i) não é necessário considerar a homogeneidade entre os servidores (veículos), (ii) pode-se considerar que os veículos operam de maneira cooperativa e (iii) pode-se modelar regras de preferências de despacho específicas de cada sistema.

Para a finalidade de localização de bases e alocação de viaturas, o uso do Modelo Hiper cubo possui a desvantagem de tornar analiticamente complexa a análise de uma quantidade grande de cenários em busca do ótimo, algo necessário diante do problema de localização, es-

pecialmente quando o número total de bases e ambulâncias é elevado, como no caso do Sistema de Atendimento Médico-Emergencial (SAMU) de São Paulo, que possui aproximadamente 140 viaturas. Dessa forma, dentre os avanços feitos em modelos de planejamento de sistemas de atendimento emergencial, surgiram aplicações que agregam resultados do Modelo Hiper cubo com modelos de localização. Dentre esses, destaca-se o trabalho de Batta et al. (1989), em que o modelo proposto por Larson (1974) é utilizado em conjunto com o modelo MEXCLP de Daskin (1983). Galvão e Morabito (2008) apresentam uma revisão de modelos híbridos, que combinam o Modelo Hiper cubo e modelos de otimização.

Nos últimos anos, avanços na capacidade de processamento de computadores e o desenvolvimento de algoritmos de solução eficientes permitiram o desenvolvimento de modelos que consideram características dinâmicas do problema, como as variações da demanda e dos tempos de deslocamento entre pontos durante um ciclo de operação do sistema. Esses modelos, seguindo a nomenclatura dada por Brotcorne et al. (2003), são os modelos dinâmicos do problema, os quais resultam em planos de localização e alocação ao longo de horizontes de planejamento.

Um trabalho que considera essas características foi desenvolvido por Gendreau et al. (2001), e trata especificamente do problema de alocação e realocação de veículos de emergência, num contexto mais operacional que estratégico, em tempo real. A formulação proposta pelos autores, denominada *Redeployment Problem t* (RPt) pode ser considerada como uma extensão do modelo DSM (BROTCORNE et al., 2003). Assim como o DSM e outros modelos de cobertura múltipla, o RPt considera dois parâmetros de cobertura, r_1 e r_2 ($r_1 < r_2$). O modelo é determinístico, definido em um horizonte de planejamento para múltiplos períodos t e busca maximizar a demanda coberta pelo menos duas vezes dentro do raio de cobertura r_1 . Algumas características relevantes do modelo são: veículos realocados num determinado período t não poderão ser novamente realocados no período $t+1$; ciclos de realocação com apenas dois pontos são evitados; realocações com grandes distâncias são evitadas; apenas um número limitado de ambulâncias pode ser posicionado em cada base; apenas um número limitado de ambulâncias pode ser re-

alocado em cada período t ; a atribuição de uma ambulância a um chamado distante no final do expediente deve ser evitada.

Questões de realocação também foram analisadas por Schmid e Doerner (2010). O modelo formulado pelos autores foi denominado *Multi-period Double Standard Model* (mDSM). Trata-se de uma formulação determinística multi-período que, além das premissas de Gendreau et al. (2001), considera diferentes valores de tempos de deslocamento para cada período t . Isso retrata condições de tráfego de regiões densamente povoadas como grandes centros urbanos. Assim, os arcos do grafo, no qual o problema é definido, passam a possuir parâmetros dinâmicos s_{ij}^t de tempo de viagem entre os pontos i e j .

3 CARACTERIZAÇÃO DO PROBLEMA

Os sistemas de ambulâncias são caracterizados pelo despacho de veículos de emergência, que atendem situações de risco à saúde e vidas humanas. Busca-se maximizar a probabilidade de sobrevivência de um indivíduo acidentado por meio da minimização do tempo de chegada ao local do acidente, pelo rápido diagnóstico das equipes de resgate que operam as viaturas, pela aplicação dos procedimentos médicos corretamente e pela minimização do tempo de transporte do local do acidente até o centro de saúde mais próximo. Dentre esses objetivos citados, a minimização do tempo de chegada se relaciona com o planejamento da localização de bases de atendimento e com a alocação de viaturas a essas bases. Segundo Singer e Donoso (2008), esses sistemas podem ser vistos como sistemas de filas, nos quais os chamados representam a demanda ou o processo de chegada, e os servidores são representados pelos veículos e suas equipes.

A posição das ambulâncias de um sistema de atendimento de emergência impacta especificamente o tempo de resposta do sistema, sendo um fator que condiciona o desempenho do

mesmo. Sendo assim, busca-se uma formulação matemática que represente o problema de encontrar, em vários períodos, a localização de bases, alocação de veículos a essas bases, e as consequentes realocações de veículos entre períodos tal que nível de serviço do sistema seja maximizado, respeitando restrições mínimas de viabilidade e disponibilidade de recursos (bases e

viaturas). O nível de serviço é definido como a fração da demanda que se espera atender em tempos inferiores à tempos de cobertura pré-definidos para cada tipo de veículo do sistema. Outra definição para o nível de serviço é a de cobertura esperada, ou probabilidade de cobertura: dado um tempo de cobertura para cada tipo de veículo, qual a fração da demanda que possivelmente será atendida em tempos inferiores.

Para a caracterização do problema, deve-se considerar também que: (i) existe uma quantidade finita de bases e de ambulâncias de dois tipos; (ii) cada veículo possui um parâmetro de cobertura associado que define, em termos temporais, sua capacidade de cobertura; (iii) são conhecidos os padrões de variação das demandas pelos serviços de atendimento de cada tipo de viatura em uma região; (iv) são conhecidos também os padrões de variação dos tempos de deslocamento nessa região; e (v) uma vez que entre períodos subsequentes podem haver realocações de ambulâncias, deseja-se também minimizar o tempo de percurso dessas realocações de acordo com um fator de proporcionalidade.

Dessa forma, busca-se encontrar um plano de operação capaz de maximizar, em múltiplos períodos de um horizonte de planejamento, a cobertura esperada do sistema, e ao mesmo tempo capaz de minimizar as realocações de viaturas necessárias entre períodos subsequentes de acordo com um fator de proporcionalidade. Tal plano deve respeitar as seguintes restrições: (i) em todos os períodos, todos os pontos de demanda devem ser cobertos por uma viatura de cada tipo; (ii) em todos os períodos, a quantidade de bases e ambulâncias é constante; (iii) em todos os períodos, a quantidade de veículos posicionados em uma base não deve ultrapassar a capacidade de acomodação de viaturas dessa base; (iv) o plano de operação deve ser conexo, ou seja, no final do horizonte de planejamento, as realocações do último período devem resultar no posicionamento de viaturas do primeiro período, sendo cíclico o plano completo.

4 MODELO MATEMÁTICO PROPOSTO

O problema tratado é definido num grafo G não direcionado, com um conjunto de pontos de demanda $i \in V$ e um conjunto de pontos candidatos $j \in W$ a receberem bases e veículos; assume-se que $W \subset V$, o que é verdadeiro na maioria dos casos práticos. Esses pontos constituem

uma simplificação da realidade uma vez que representam uma determinada localização geográfica concentrada em um único ponto. A determinação do nível de agregação da demanda que resulta nos pontos i depende da precisão desejada na localização de bases. Esse nível de agregação dos pontos de demanda é considerado o mesmo para os pontos candidatos. O que define se um ponto é candidato é a sua capacidade de receber uma base de veículos, por exemplo, pontos que representam distritos com instalações do corpo de bombeiros, hospitais próximos, ou zonas muito isoladas e distantes de um município. São considerados também períodos de tempo $t \in \tau = \{0, 1, 2, \dots, t, \dots, T\}$, sendo a soma dos períodos t equivalente ao horizonte de planejamento para o qual serão definidas as localizações de bases e alocações de ambulâncias.

Para cada período t define-se o tempo de deslocamento s_{ij}^t entre dois pontos $i \in \{V \cup W\}$ e $j \in \{V \cup W\}$. A definição dos tempos s_{ij}^t , que expressam a variação dinâmica das condições de tráfego da região modelada, pode ser tanto determinística quanto probabilística, dependendo da disponibilidade de informações sobre os tempos de transporte entre i e j . Para o problema sendo modelado, considera-se que esses tempos de deslocamento sejam determinísticos e variem apenas de acordo com o período t , uma vez que dados dos veículos em tempo real (os chamados *Floating Car Data*, ou FCD), que permitiriam estimar funções de probabilidades, ainda não estão disponíveis nas cidades brasileiras. No entanto, a modelagem é genérica a ponto de permitir uma modelagem probabilística desses tempos, e mais adiante nesta seção são apresentadas as modificações necessárias ao modelo para este caso. Com isso define-se o grafo não direcionado G .

$$G(N, A^t); N = V \cup W \quad (1)$$

$$A^t = \{s_{ij}^t : i \in V \cup W; j \in V \cup W\}$$

A formulação é definida para dois tipos de veículos k , básicos (*Basic Life Support* - BLS) e avançados (*Advanced Life Support* - ALS). O índice k igual a um é utilizado para representar veículos do tipo BLS, e o índice k igual a dois é utilizado para representar veículos do tipo ALS. Como condição mínima de desempenho do sistema deseja-se que, em todos os períodos, todos

os pontos de demanda possuam pelo menos uma unidade BLS localizada num ponto candidato j a menos de um tempo de deslocamento r_1 ; e deseja-se também que todos os pontos de demanda possuam pelo menos uma unidade ALS localizada num ponto candidato j a menos de um tempo de deslocamento r_2 . Em geral, um sistema de ambulâncias possui mais veículos do tipo BLS do que ALS, o que resulta na maioria dos casos práticos em $r_1 \leq r_2$.

Cada ponto de demanda i possui uma demanda d_i^{kt} , avaliada em termos de frequência de chamados por unidade de tempo, em cada período t para cada tipo de veículo k . Definem-se também os conjuntos W_i^{kt} , V_j^{kt} e N_i^{kt} conforme as expressões (2), (3) e (4).

$$W_i^{kt} = \{j \in W : s_{ij}^t \leq r_k\}; k = \{1, 2\} \quad (2)$$

$$V_j^{kt} = \{i \in V : s_{ij}^t \leq r_k\}; k = \{1, 2\} \quad (3)$$

$$N_i^{kt} = \{z \in V : s_{iz}^t \leq r_k\}; k = \{1, 2\} \quad (4)$$

Nota-se que a definição dos conjuntos W_i^{kt} , V_j^{kt} e N_i^{kt} é intrinsecamente ligada à definição dos parâmetros s_{ij}^t , ou seja, uma definição determinística dos tempos de deslocamento resulta em conjuntos determinísticos. Cada vizinhança W_i^{kt} representa o conjunto de pontos candidatos j que estão localizados a menos de um dado tempo de deslocamento r_k em relação a um ponto i no período t . Cada vizinhança V_j^{kt} representa o conjunto de pontos de demanda i que estão localizados a menos de um tempo de deslocamento r_k em relação a um ponto candidato j no período t . Cada vizinhança N_i^{kt} representa o conjunto de pontos de demanda i que estão localizados a menos de um tempo de deslocamento r_k de um ponto de demanda i .

O modelo matemático tem o intuito de localizar no grafo G uma quantidade p_z de bases e alocar, nos diversos períodos de tempo t , quantidades p_B de ambulâncias básicas e p_A de ambulâncias avançadas. Considera-se também que em cada ponto candidato j , em qualquer instante de tempo, não podem ser alocados mais do que C_j veículos.

As bases devem ser localizadas nos pontos candidatos e, em cada período, as viaturas devem ser alocadas as bases. Para isso definem-se

as variáveis de decisão z_j , y_j^{kt} e x_i^{wkt} de acordo com as expressões (5), (6) e (7).

Simultaneamente à questão do posicionamento de bases e ambulâncias, existe o problema de realocação. Sendo diferente a alocação de viaturas entre períodos subsequentes, deve-se movimentar as viaturas entre esses períodos, partindo da alocação de um período para o próximo de maneira a minimizar o tempo total de percurso de todas as ambulâncias. Considerando essa situação definem-se as variáveis de decisão $r_{jj'}^{kt}$, conforme mostrado na expressão (8).

De maneira análoga ao modelo Q-MALP desenvolvido por Marianov e ReVelle (1996), consideram-se dois tipos de vizinhanças do ponto i , definidas para cada parâmetro de cobertura, r_1 e r_2 , ou seja, para cada ponto i , em cada período de tempo t e para cada tipo de veículo k , define-se uma vizinhança. Admite-se que essas vizinhanças funcionam como sistemas de filas *M/G/s-loss*, um sistema de filas com s servidores tal que: a chegada de clientes ocorre de acordo com um processo de Poisson com média $1/\lambda$, o serviço de atendimento ocorre com um tempo definido segundo uma distribuição de probabilidade genérica com média $1/\mu$, e quando um cli-

ente entra no sistema e não existem servidores disponíveis ele não é atendido e sai do sistema, não havendo a formação de filas. Detalhes sobre a formulação de filas *M/G/s-loss* podem ser encontrados, por exemplo, em Hillier e Lieberman (2004). Para cada uma dessas vizinhanças, é calculada uma fração de ocupação q , que equivale à probabilidade de uma ambulância randomicamente selecionada estar ocupada. Como o modelo trata de dois tipos de veículos, para cada ponto de demanda i em cada período t são consideradas duas frações de ocupação: uma referente à cobertura por veículos BLS (vizinhança relativa ao parâmetro r_1) e outra referente à cobertura por veículos ALS (vizinhança relativa ao parâmetro r_2). Essas frações de ocupação $q_{k,i}^t$ podem ser calculadas segundo a expressão (9).

Sendo que \bar{t} é o tempo médio de atendimento em horas, d_z^{kt} é a demanda, expressa em chamados por dia, do ponto genérico z por veículos do tipo k durante o período t , e y_j^{kt} é a quantidade de veículos do tipo k localizados no ponto j no período t . O divisor 24 serve apenas para compatibilizar a unidade de tempo da demanda e do tempo de atendimento.

$$z_j = \begin{cases} 1, & \text{se é aberta uma base no ponto candidato } j \in W \\ 0, & \text{caso contrário} \end{cases} \quad (5)$$

$$y_j^{kt} = \text{número de veículos do tipo } k \text{ posicionados no ponto } j \in W, \text{ no período } t \in \tau \quad (6)$$

$$x_i^{wkt} = \begin{cases} 1, & \text{se o ponto de demanda } i \text{ é coberto por } w \text{ veículos do tipo } k \text{ no período } t \in \tau \\ 0, & \text{caso contrário} \end{cases} \quad (7)$$

$$r_{jj'}^{kt} = \text{número de veículos } k \text{ realocados de } j \in W \text{ para } j' \in W \text{ entre os períodos } t \text{ e } t+1 \quad (8)$$

$$q_{k,i}^t \cong \frac{\bar{t} \sum_{z \in N_i^{kt}} d_z^{k,t}}{24 \sum_{j \in W_i^{kt}} y_j^{kt}} \quad (9)$$

Considerando que a soma das demandas, expressas em frequências de chamadas por dia, é equivalente a uma taxa de geração de chamados e que o inverso do tempo médio de atendimento, definido em horas, é equivalente a uma taxa de atendimento de servidores em sistemas de filas, o quociente entre eles é análogo a uma

taxa de congestionamento do sistema ρ_i^{kt} . Além disso, reescrevendo o somatório de y_j^{kt} em todos os pontos candidatos $j \in W_i^{kt}$ como uma variável b_i^{kt} que representa a quantidade total de ambulâncias do tipo k localizadas na vizinhança W_i^{kt} ,

a expressão (9) pode ser reescrita conforme a expressão (10).

$$q_{r_k,i}^t \cong \frac{\lambda_i^{k,t}}{\mu_i^{k,t} \sum_{j \in W_i^{kt}} y_j^{1,t}} = \frac{\rho_i^{k,t}}{b_i^{k,t}} \quad (10)$$

As taxas de congestionamento, ρ_i^{kt} , são utilizadas para calcular a probabilidade de um servidor selecionado randomicamente estar ocupado considerando o modelo de filas *M/G/s-loss* para a vizinhança W_i^{kt} .

Marianov e ReVelle (1996) formulam o modelo Q-MALP utilizando as equações de estado estacionário para o referido sistema de filas e assumindo que o processo de chegada é um processo de Poisson. Nesta pesquisa toma-se como premissa as vizinhanças de fato se comportarem como sistemas de filas em estado estacionário e o processo de chegadas ser um processo de Poisson. Considerando uma taxa genérica de congestionamento ρ de um sistema de filas *M/G/s-loss* em estado estacionário, a probabilidade $p(w)$ de w servidores estarem ocupados é dada pela expressão (11); detalhes sobre o seu desenvolvimento podem ser encontrados, por exemplo, em Hillier e Lieberman (2004).

$$p(w) = \frac{\frac{1}{w!} \rho^w}{1 + \rho + \frac{1}{2!} \rho^2 + \dots + \frac{1}{w!} \rho^w} \quad (11)$$

Outros estudos de sistemas de ambulâncias adotam ou validam as premissas adotadas quanto ao sistema de filas. Singer e Donoso (2008) analisam um serviço de ambulância no Chile no qual os autores assumem a estacionariedade do sistema de filas, baseado no fato de o serviço ser oferecido ininterruptamente, e validam a premissa de processo poissoniano para descrever a demanda por atendimento por meio

de uma análise de dados coletados. Mendonça e Morabito (2001) analisam um sistema de atendimento emergencial de uma rodovia Brasileira para a qual os autores assumem a premissa de estacionariedade do sistema além de testar e validar estatisticamente a hipótese de processo poissoniano de demanda. Channouf et al. (2007), realizam um estudo do sistema de ambulâncias de Amsterdã no qual concluem, através de uma análise estatística de dados de demanda, que a premissa de processo poissoniano é válida para o serviço em estudo.

Além disso, como argumentam Matteson et al. (2010), a referida hipótese de processo poissoniano é utilizada na maioria dos estudos sobre serviços de ambulância, e é suportada pelo teorema de Palm-Khintchine, o qual afirma que um grande número de processos, não necessariamente poissonianos, se sobrepostos, originam um processo aproximadamente poissoniano. Dessa forma, pode-se argumentar que essas duas premissas não limitam a aplicação do modelo a casos reais de sistemas de ambulância.

Com a expressão (11) é possível calcular, num sistema de filas, a probabilidade de atendimento $E(w)$, que é simplesmente a probabilidade complementar de $p(w)$, representando a probabilidade de haver ao menos um servidor disponível no momento de ocorrência de um chamado.

$$E(w) = (1 - p(w)) \quad (12)$$

Assim, a cobertura incremental obtida por haver w ao invés de $(w-1)$ veículos atendendo chamados dentro do sistema pode ser obtida de acordo com a expressão (13), que desenvolvida algebricamente, considerando especificamente as vizinhanças W_i^{kt} , resulta nas coberturas incrementais de cada vizinhança, C_i^{wkt} , dadas pela expressão (14).

$$C^w = E(w) - E(w-1) = [1 - p(w)] - [1 - p(w-1)] = p(w-1) - p(w) \quad (13)$$

$$C_i^{wkt} = \left[\frac{\frac{1}{(w-1)!} (\rho_i^{kt})^{w-1}}{1 + \rho_i^{kt} + \frac{1}{2!} (\rho_i^{kt})^2 + \dots + \frac{1}{(w-1)!} (\rho_i^{kt})^{w-1}} \right] - \left[\frac{\frac{1}{w!} (\rho_i^{kt})^w}{1 + \rho_i^{kt} + \frac{1}{2!} (\rho_i^{kt})^2 + \dots + \frac{1}{w!} (\rho_i^{kt})^w} \right] \quad (14)$$

$$1 - p(M_i^{kt}) \geq \alpha \Leftrightarrow \frac{\frac{1}{M_i^{kt}!} (\rho_i^{kt})^{M_i^{kt}}}{1 + \rho_i^{kt} + \frac{1}{2!} (\rho_i^{kt})^2 + \dots + \frac{1}{M_i^{kt}!} (\rho_i^{kt})^{M_i^{kt}}} \leq 1 - \alpha \quad (15)$$

A expressão (14) é obtida através da substituição de (12) em (13) e desenvolvendo as probabilidades $p(w)$ e $p(w-1)$ em função da taxa de congestionamento ρ_i^{kt} de cada vizinhança.

Além disso, seguindo os conceitos dos modelos MALP I e II propostos por ReVelle e Hogan (1989), pode-se calcular com o uso da expressão (15) a quantidade M_i^{kt} que é a quantidade mínima de veículos do tipo k de modo que a probabilidade de todos os veículos desse tipo estarem ocupados na vizinhança W_i^{kt} do ponto i no período t seja inferior a $(1-\alpha)$.

Assim, considerando todos os pontos de demanda do conjunto V , todos os períodos do conjunto τ e os dois tipos de veículo, $k=1$ e $k=2$, é possível calcular a cobertura esperada em um sistema de atendimento emergencial pela expressão (16).

$$\sum_{t \in \tau} \sum_{k=1}^2 \sum_{i \in V} \sum_{w=0}^{M_i^{kt}} d_i^{kt} C_i^{wkt} x_i^{wkt} \quad (16)$$

Vale ressaltar que a cobertura esperada do sistema, dada pela expressão (16), é limitada superiormente pelo produto entre a demanda total do sistema e a probabilidade α , uma vez que a

quantidade w de ambulâncias do tipo k que cobrem um ponto i em um período de tempo t é sempre menor ou igual a M_i^{kt} .

O modelo proposto busca maximizar a cobertura esperada do sistema, ao mesmo tempo em que busca minimizar o tempo total de realocação de viaturas entre períodos subsequentes. Esse tempo total de realocação, que depende dos tempos de deslocamento s_{ij}^t e das variáveis de decisão $r_{jj'}^{kt}$, pode ser calculado segundo a expressão (17).

$$\sum_{t \in \tau} \sum_{j \in W} \sum_{j' \in W} \sum_{k=1}^2 s_{jj'}^t r_{jj'}^{kt} \quad (17)$$

O modelo matemático para o problema de localização de bases, alocação de ambulâncias em múltiplos períodos e realocação entre períodos subsequentes, proposto neste artigo, pode ser definido conforme as expressões (18) a (32). A sua resolução permite determinar um plano de operação num horizonte pré-definido de tempo, ou seja, resulta na localização de bases que deve ser estabelecida, na alocação de viaturas que varia nos múltiplos períodos de tempo e nas realocações necessárias entre períodos subsequentes.

$$[\max] \sum_{t \in \tau} \left[\sum_{k=1}^2 \sum_{i \in V} \sum_{w=0}^{M_i^{kt}} d_i^{kt} C_i^{wkt} x_i^{wkt} - \beta \sum_{j \in W} \sum_{j' \in W} \sum_{k=1}^2 s_{jj'}^t r_{jj'}^{kt} \right] \quad (18)$$

Sujeito a:

$$\sum_{j \in W_i^{k,t}} y_j^{kt} \geq 1, \forall i \in V, \forall t \in \tau, k = \{1,2\} \quad (19)$$

$$\sum_{j \in W_i^{k,t}} y_j^{kt} \geq \sum_{w=0}^{M_i^{kt}} x_i^{wkt}, \forall i \in V, \forall t \in \tau, k = \{1,2\} \quad (20)$$

$$x_i^{wkt} \leq x_i^{(w-1)kt}, \forall i \in V, \forall t \in \tau, k = \{1,2\}, w = \{1,2, \dots, M_i^{kt}\} \quad (21)$$

$$p_k z_j \geq y_j^{kt}, \forall j \in W, \forall t \in \tau, k = \{1,2\} \quad (22)$$

$$y_j^{kt} + \sum_{i \in W} r_{ij}^{kt} - \sum_{i \in W} r_{ji}^{kt} = y_j^{k(t+1)}, \forall j \in W, \forall t \in \tau - \{T\}, k = \{1,2\} \quad (23)$$

$$y_j^{kT} + \sum_{i \in W} r_{ij}^{kT} - \sum_{i \in W} r_{ji}^{kT} = y_j^{k1}, \forall j \in W, k = \{1,2\} \quad (24)$$

$$\sum_{j \in W} z_j = p_z \quad (25)$$

$$\sum_{j \in W} y_j^{kt} = p_B, \forall t \in \tau \quad (26)$$

$$\sum_{j \in W} y_j^{2t} = p_A, \forall t \in \tau \quad (27)$$

$$y_j^{1t} + y_j^{2t} \leq C_j, \forall j \in W, \forall t \in \tau \quad (28)$$

$$y_j^{kt} \geq 0 \text{ inteiro}, \forall j \in W, \forall t \in \tau, k = \{1,2\} \quad (29)$$

$$x_i^{wkt} = \{0,1\}, \forall i \in V, \forall t \in \tau, k = \{1,2\}, w = \{0,1,2,\dots, M_i^{kt}\} \quad (30)$$

$$z_j = \{0,1\}, \forall j \in W \quad (31)$$

$$r_{jj'}^{kt} \geq 0 \text{ inteiro}, \forall (j, j') \in W, \forall t \in \tau, k = \{1,2\} \quad (32)$$

A função objetivo (18) busca a maximização da cobertura esperada para os pontos de demanda em todos os períodos de tempo e ao mesmo tempo busca minimizar as realocações de veículos de maneira proporcional à distância de realocação, sendo a constante de proporcionalidade igual ao parâmetro β . O primeiro termo da função objetivo representa a cobertura esperada referente a uma configuração de bases e veículos em todos os períodos. O produto da demanda d_i^{kt} pelo coeficiente C_i^{wkt} indica a demanda incremental do tipo k que se espera cobrir com w veículos do tipo k posicionados na vizinhança do ponto i no período t . O segundo termo da função objetivo representa o custo global de realocações em todos os períodos. Para um detalhamento sobre o parâmetro de proporcionalidade do tempo total de realocação β sugere-se consultar Schmid e Doerner (2010).

A restrição (19) assegura o nível de serviço mínimo do sistema, ou seja, garante que em todos os períodos, todos os pontos de demanda devem ter pelo menos uma ambulância BLS alocada em uma base a menos de um raio de cobertura r_1 , e também pelo menos uma ambulância ALS alocada em uma base a menos de um raio de cobertura r_2 .

As expressões (20) e (21) garantem consistência das definições das variáveis de decisão x_i^{wkt} e y_j^{kt} . Mais especificamente, em (20) assegura-se que a quantidade de viaturas posicionadas em pontos j na vizinhança do ponto i não seja superior à quantidade de viaturas que cobre o ponto i . em (21) assegura-se que um ponto i só seja coberto por w veículos do tipo k se ele o for por $(w-1)$ veículos. As restrições (22) estabelecem que veículos só podem ser alocados a pontos candidatos que contenham bases localizadas neles.

As restrições (23) e (24) são equivalentes a equações de balanceamento de fluxo de ambu-

lâncias numa base. Elas garantem a consistência na definição das realocações, de modo que em um determinado período t , a quantidade de viaturas do tipo k alocadas em uma base localizada em um ponto j é igual a quantidade de viaturas k neste ponto no período anterior, mais a quantidade de viaturas do tipo k realocadas de outras bases para essa base j , menos o número de viaturas do tipo k realocadas dessa base j para outras bases no período anterior. Vale ressaltar que as restrições (24) garantem uma continuidade do plano de operação resultante da solução do modelo matemático, de maneira que a realocação do último período $t=T$ deve resultar na alocação do primeiro período do plano de operação $t=1$.

As restrições (25), (26) e (27) são, respectivamente, as restrições da quantidade de bases que devem ser localizadas, das ambulâncias básicas e das avançadas que devem ser alocadas ao longo dos períodos. As restrições (28) limitam, para todos os períodos, a quantidade de veículos que pode ser alocada em uma base. O domínio das variáveis de decisão é definido pelas equações (29), (30), (31) e (32).

Na expressão (1) o grafo G foi definido considerando os tempos de deslocamento como grandezas determinísticas e conhecidas *a priori*. Uma abordagem alternativa, como apresentado em Marianov e ReVelle (1996), é a consideração de tempos de deslocamento como variáveis probabilísticas com distribuição conhecida; dessa forma os tempos de deslocamento entre os pontos do grafo podem ser definidos considerando um nível de confiança δ . Pode-se ilustrar essa definição probabilística dos tempos de deslocamento assumindo que cada variável s_{ij}^t segue uma distribuição normal com média \bar{s}_{ij}^t e desvio padrão σ_{ij}^t , de tal forma que os tempos de deslocamento podem ser definidos de acordo com a expressão (33), em que z_δ é o valor da fun-

ção cumulativa normal de probabilidade que satisfaz o nível de confiança δ .

$$s^{*t}_{ij} = \bar{s}_{ij}^t + z_{\delta} \cdot \sigma_{ij}^t \quad (33)$$

Essa definição dos tempos de deslocamento é estendida à definição dos conjuntos W_i^{kt} , V_j^{kt} e N_i^{kt} , conforme expressões (34), (35) e (36).

$$W_i^{kt} = \left\{ j \in W \mid s^{*t}_{ij} = \bar{s}_{ij}^t + z_{\delta} \cdot \sigma_{ij}^t \leq r_k \right\} \quad (34)$$

$$V_j^{kt} = \left\{ i \in V \mid s^{*t}_{ij} = \bar{s}_{ij}^t + z_{\delta} \cdot \sigma_{ij}^t \leq r_k \right\} \quad (35)$$

$$N_i^{kt} = \left\{ z \in V \mid s^{*t}_{ij} = \bar{s}_{ij}^t + z_{\delta} \cdot \sigma_{ij}^t \leq r_k \right\} \quad (36)$$

A consideração de tempos de deslocamento determinísticos ou probabilísticos não altera o restante do modelo, impactando somente no cálculo de s^{*t}_{ij} e na definição dos conjuntos.

O modelo pode ser considerado original no sentido em que não há outro idêntico na literatura. Porém, ele pode também ser visto como uma extensão do modelo Q-MALP proposto por Marianov e ReVelle (1996), utilizando alguns dos conceitos apresentados em Schmid e Doerner (2010) relativos à realocação das ambulâncias, apresentando as seguintes contribuições: (i) consideração de múltiplos períodos de planejamento e consequente consideração do problema de realocação entre períodos subsequentes, (ii) consideração da característica dinâmica da questão, no sentido em que as demandas e tempos de deslocamento são diferentes para cada período, (iii) distinção entre a localização de bases e a alocação de viaturas, (iv) consideração de múltiplos tipos de veículos e diferentes raios de cobertura para cada um e (v) consideração de restrições de capacidade nas bases. Maiores detalhes sobre o modelo matemático podem ser encontrados em Andrade (2012).

Este modelo matemático pode também ser visto como uma extensão do modelo Q-MALP que considera o dinamismo dos parâmetros do problema estudado, em especial, as variações dos tempos de transporte e das demandas. Esses dois fatores são importantes no estudo de grandes metrópoles com problemas de congestionamentos e alta variação na densidade demográfica da cidade ao longo do dia, como é o caso do município de São Paulo. Essa é a principal contribuição da metodologia proposta.

Outros modelos matemáticos para o pro-

blema, principalmente os baseados no Modelo Hipercubo, descrevem de maneira mais apropriada as características estocásticas do atendimento emergencial, em especial, a cobertura esperada. Outros estudos, como por exemplo o realizado por Iannoni et al. (2011), utilizam o Modelo Hipercubo em conjunto com modelos de otimização probabilísticos; o Modelo Hipercubo é utilizado para avaliar configurações fornecidas por modelos de localização e em seguida para determinar as vizinhanças de atendimento de cada ambulância (*districting*), fornecendo soluções para implantação de sistemas de ambulância através de um algoritmo que combina as duas metodologias. Uma abordagem como esta poderia ser um avanço considerável para o modelo de proposto neste artigo, que tem como foco o posicionamento de bases e ambulâncias considerando um planejamento em múltiplos períodos, o primeiro estágio do projeto de sistemas de ambulância.

5 APLICAÇÃO DO MODELO

O modelo matemático proposto foi aplicado para avaliação e melhoria do Sistema de Atendimento Móvel Pré-hospitalar de Urgência do município de São Paulo (SAMU-SP), sendo sua solução realizada por um algoritmo de solução proposto por Andrade (2012), baseado na meta-heurística de Colônia Artificial de Abelhas. O município apresenta uma alta densidade demográfica nas regiões centrais durante os períodos diurnos, sendo esse adensamento distribuído nos períodos noturnos, além disso, a malha viária da cidade diariamente apresenta congestionamentos de veículos.

O SAMU-SP conta com 140 viaturas divididas entre viaturas básicas (BLS) e avançadas (ALS). São empregadas bases fixas e bases móveis de atendimento. As bases fixas são edificações alugadas espalhadas na cidade ou cedidas por outros órgãos públicos como estações do corpo de bombeiros e hospitais. As bases móveis, ou bases modulares, são edificações de montagem e desmontagem rápida (cerca de dois dias) que ficam localizadas em geral em praças ou qualquer local público. Uma das finalidades básicas das bases móveis é assegurar atendimento a eventos especiais com grande concentração de pessoas como, por exemplo, eventos esportivos; contudo, sua rapidez de montagem e desmontagem, faz com que sejam também uma

opção para as bases fixas.

Nesta aplicação, a cidade foi dividida em 96 distritos, todos candidatos a receberem bases e viaturas, que representam os pontos de demanda e conseqüentemente os nós da rede de atendimento, sendo que 47 desses distritos contêm bases fixas de atendimento e outros sete distritos contêm bases móveis; contudo existem ao todo 13 bases móveis que são empregadas pelo SAMU-SP. Ressalta-se que nesse estudo de caso foram identificados distritos contendo mais de uma base, resultando que o número total de bases do SAMU-SP é diferente do número de distritos que contêm bases considerando a configuração atual. Foi considerado um horizonte de planejamento de uma semana dividido em 21 períodos (3 períodos por dia ao longo de 7 dias).

O procedimento de recebimento e triagem de chamados do SAMU-SP não distingue entre chamados que necessitam de viaturas do tipo básico e chamados que necessitam de viaturas do tipo avançado; assim, as demandas foram definidas apenas em relação a um tipo de veículo; o mesmo foi feito com relação aos tempos de cobertura r_1 e r_2 do modelo matemático, ou seja, foi feita uma simplificação do modelo considerando apenas um parâmetro de cobertura t_c tendo em vista a indisponibilidade de dados das demandas de chamados por tipo.

O SAMU SP, no momento de condução da pesquisa, conta com um sistema de *Computer Aided Dispatch* (CAD) que auxilia no recebimento, triagem e despacho dos chamados, e conta também com um sistema de localização automática de veículos (*Automatic Vehicle Location* - AVL). Os dados provenientes desses sistemas foram utilizados para a condução das análises. Foram extraídos relatórios de frequência de chamados em cada região e em cada período de planejamento. As velocidades de transporte das viaturas em atendimento foram levantadas do sistema AVL.

Inicialmente, foi feita uma avaliação da configuração atual do sistema de atendimento do SAMU-SP, no que diz respeito à localização de bases. Foram realizadas tentativas de solução do problema considerando a configuração atual de bases, variando o tempo de cobertura, entre 15 e 30 minutos, e assumindo valores de duas, três e quatro horas para o tempo de atendimento. Constatou-se que a configuração atual apresenta soluções viáveis apenas a partir de um tempo de cobertura de 27 minutos.

Uma possível melhoria seria um melhor emprego das 13 bases móveis que o SAMU-SP já possui. Foram realizados testes considerando as 47 localizações das bases fixas atuais e as localizações das sete bases móveis atuais, sendo que a localização das outras seis bases móveis foi determinada pela solução do modelo matemático. Os resultados são apresentados na Figura 1 e mostram que o reposicionamento de seis bases do SAMU-SP pode diminuir o tempo máximo de cobertura do sistema de 27 para 16 minutos com probabilidade superior a 99%. Isso significa uma melhoria de desempenho apenas com o melhor emprego dos recursos atuais.

Além dessa avaliação do sistema atual, foram realizadas outras análises variando a quantidade de bases e ambulâncias do sistema e os tempos de cobertura e de atendimento. Para essas análises, foram considerados dois tipos de cenário, um que considera como restritas as 47 localizações de bases fixas atuais, denominado de instâncias de teste não livres; e outro, denominado de instâncias de testes livres, que considera 100% das bases como móveis podendo estas serem posicionadas livremente pela solução do modelo matemático. Foram testadas instâncias considerando tempos de cobertura de 15, 10 e cinco minutos, e tempos de atendimento de duas, três e quatro horas; neste artigo são apresentados apenas os resultados mais relevantes.

A Figura 2 apresenta os resultados considerando a situação de testes não livres, e um tempo de cobertura de 15 minutos. Observa-se que existem soluções viáveis para o problema mesmo considerando pequenas quantidades de bases e ambulâncias; porém, com cobertura esperada da ordem de 80% a 85%, dependendo do número de bases. Também é possível verificar que, para todas as curvas apresentadas, os ganhos marginais de cobertura obtidos com o aumento da quantidade de ambulâncias no sistema são decrescentes com a quantidade de viaturas. Essa constatação está de acordo com as afirmações de Daskin (1983). Além disso, é possível verificar que a partir de 100 ambulâncias no sistema, independentemente da quantidade de bases, o aumento no número de viaturas contribui pouco para o aumento de cobertura esperada.

A Figura 3 apresenta uma comparação entre os resultados das instâncias de testes livres e não livres considerando um tempo de cobertura de 10 minutos e um tempo de atendimento de duas horas. Pode-se observar que soluções viáveis

para as instâncias não livres são encontradas com uma quantidade de bases a partir de 80 e uma quantidade de ambulâncias a partir de 70. No caso ds instâncias livres, é possível encontrar soluções viáveis com uma menor quantidade de

bases e ambulâncias; são encontradas soluções com 70 bases e 70 ambulâncias. Esse resultado evidencia que existem vantagens em considerar 100% das bases móveis, podendo ser posicionadas em qualquer distrito.

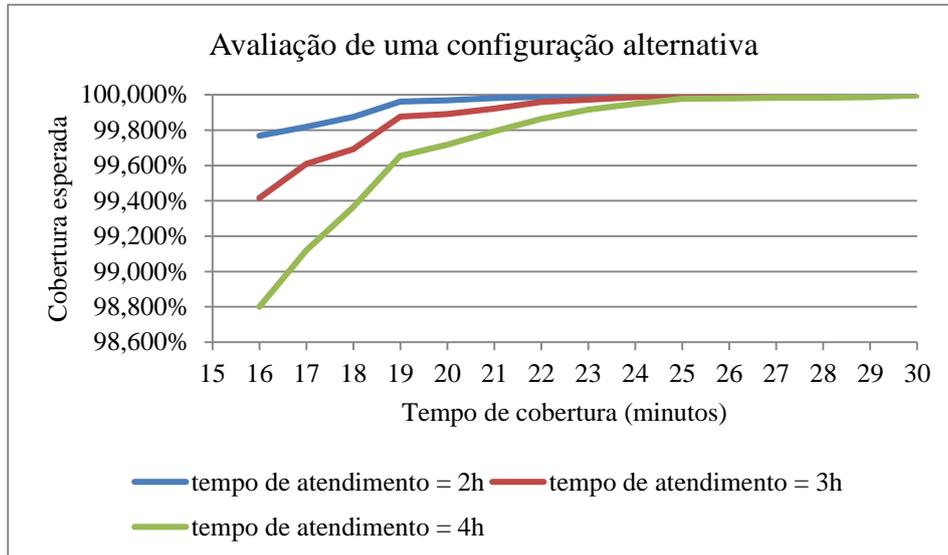


Figura 1 - Resultados da avaliação de melhoria da configuração atual de bases do SAMU-SP

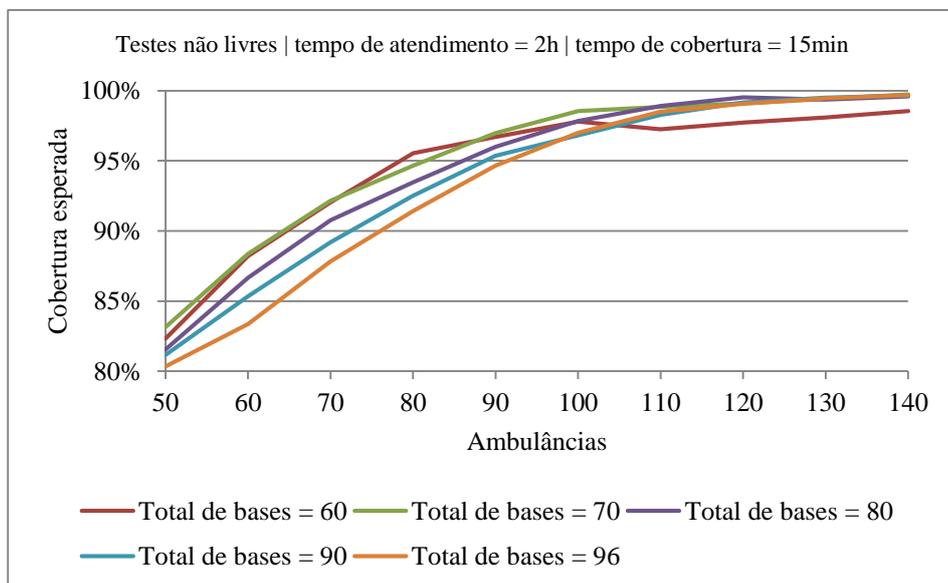


Figura 2 - Resultados das instâncias de teste não-livres considerando 15 minutos como tempo de cobertura e duas horas como tempo de atendimento

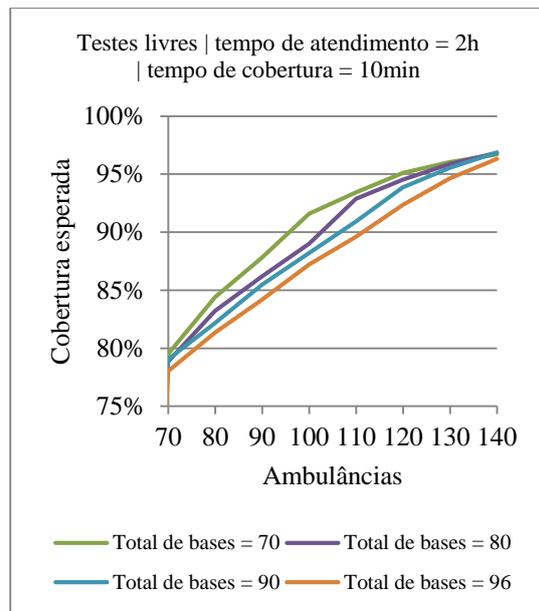
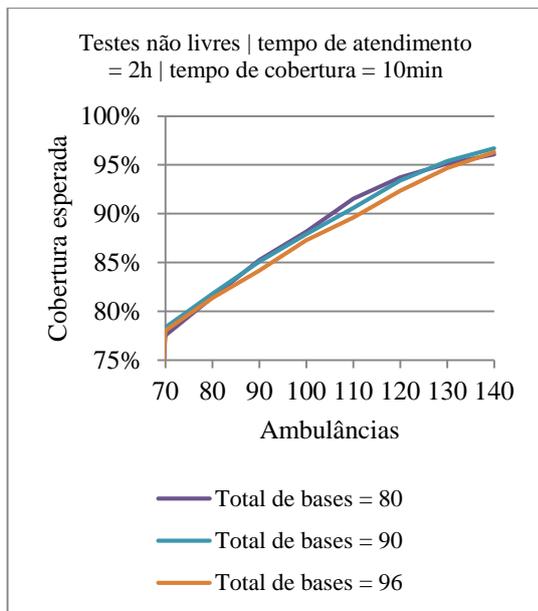


Figura 3 - Comparação entre os resultados das instâncias de testes livres não-livres considerando 10 minutos como tempo de cobertura e duas horas como tempo de atendimento

Os resultados considerando um tempo de cobertura de 5 minutos apresentam soluções viáveis apenas com 96 distritos cobertos por bases, ou seja, soluções em que todos os distritos contêm bases; dessa forma, independe se a instância considera 100% das bases como móveis ou não. Os resultados mostram que é possível encontrar soluções, considerando um tempo de atendimento igual a duas horas e com 140 ambulâncias, com cobertura superior a cerca de 92%.

6 CONSIDERAÇÕES FINAIS

Neste trabalho foi proposto um modelo matemático inédito para o problema denominado neste artigo de “Problema de localização de bases, alocação de veículos em múltiplos períodos, e realocação entre períodos subsequentes”. Esse modelo abrange os seguintes pontos do problema de planejamento de sistemas de atendimento emergencial: determinação do posicionamento de bases e da correspondente alocação de veículos feita de forma independente, possibilidade de considerar mais de um tipo de veículo para diferentes tipos de demanda, diferentes tempos de cobertura para cada tipo de veículo, consideração de capacidade de acomodação de viaturas nas bases e disponibilidade finita de recursos de atendimento (bases e viaturas), modelagem probabilística da cobertura, consideração dos padrões de variação espaço-temporal da demanda e consideração das variações temporais dos tempos de deslocamento entre os

diversos locais de uma região. Pode-se considerar que esse é um modelo bastante abrangente podendo ser aplicado a diversas situações.

Uma primeira e relevante recomendação para uma extensão futura desta pesquisa é a validação das premissas de modelagem adotadas, em particular a estacionariedade do modelo de filas *M/G/s-loss* que representa a vizinhança de um ponto do grafo, ou mesmo o processo de chegada dos chamados de emergência como sendo um processo poissoniano. Outro aprimoramento do modelo seria a sua combinação com metodologias baseadas no Modelo Hipercubo, como em Iannoni et al. (2011). O Modelo Hipercubo seria utilizado para avaliar os resultados do modelo de posicionamento de bases e ambulâncias em múltiplos períodos; isso faria com que as características estocásticas do atendimento emergencial fossem descritas de maneira mais apropriada, fornecendo melhores resultados para o planejamento de sistemas de ambulâncias em grandes metrópoles.

O estudo de caso do município de São Paulo foi realizado considerando o SAMU-SP, suas bases e viaturas. Os testes foram feitos levando em conta a variação de diversos parâmetros: número de bases e ambulâncias, tempo de atendimento de cada chamado e tempo de cobertura desejado. Os resultados mostram que é possível chegar a um tempo de cobertura de 16 minutos com probabilidade acima de 95% considerando a quantidade de recursos existentes, desde que melhor empregados em relação à configura-

ção atual. Pode-se ainda, com um acréscimo do número de bases até um total de 96, chegar a um tempo de cobertura de cinco minutos com probabilidade próxima de 95%. Os resultados também mostram que existem vantagens em operar com bases móveis, que podem ser reposicionadas em pouco tempo, ao invés de bases fixas.

Na sua forma atual, o modelo matemático apresentado neste artigo pode ser implementado pelos gestores de sistemas de atendimento emergencial como o primeiro passo de um processo de planejamento, fazendo com que os recursos sejam otimizados, maximizando o nível de serviço para os usuários e assim aumentando as chances de salvamento de vidas.

REFERÊNCIAS

- Andrade, L. A. C. G. (2012) *Heurística baseada em colônia artificial de abelhas para o problema de localização de bases, alocação e realocação de ambulâncias*. 2012. 250p. Dissertação (Mestrado) – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Logísticos, São Paulo.
- Bianchi, G. e Church, R. L. (1988) A hybrid fleet model for emergency medical service systems design. *Social Sci. Med.* 26, 163-171.
- Batta, R.; Dolan, J. e Krishnamurthy, N. (1989) The maximal expected covering location problem: revisited. *Transport. Science.* 23, 277-287.
- Brotcorne, L.; Laporte, G. e Semet, F. (2003) Ambulance Location and Relocation Models. *European Journal of Operations Research* 147, 451-463.
- Channouf, N., L'Ecuyer, P., Ingolfsson, A., & Avramidis, A. N. (2007). The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta. *Health Care Management Science* 10 (1), 25-45.
- Chiyoshi, F., Iannoni, A. P., Morabito, R. (2011) A tutorial on hypercube queueing models and some practical applications in emergency service systems. In: Simpósio Brasileiro de Pesquisa Operacional, XLIII, 2011, Ubatuba. Anais do XLIII Simpósio Brasileiro de Pesquisa Operacional.
- Church, R. L. e ReVelle, C. (1974) The maximal covering location problem. *Papers of the Regional Science Association* 32,101-118
- Daskin, M. S. (1983) A maximum expected location model: Formulation, properties and heuristic solution. *Transportation Science* 7, 48-70.
- Daskin M. (1995) *Network and discrete location: models, algorithms, and application*. John Wiley&Sons.
- Galvão, R. D., Morabito, R. (2008) Emergency service systems the use of the hypercube queueing model in the solution of probabilistic location problems. *International Transactions in Operations Research* 15 - 525-549.
- Galvão, R. D., Chiyoshi, F. Y., Morabito, R. (2005) Towards unified formulations and extensions of two classical probabilistic location models. *Computers & Operations Research* 32 - 15-33.
- Gendreau, M.; Laporte, G. e Semet, F. (2001) A dynamic model and parallel tabu search heuristic for real-time ambulance relocation. *Parallel Computing* 27, 1641-1653.
- Hillier, F. S., Lieberman, G. J. Introduction to Operations Research. 8th Edition. McGraw-Hill College, 2004.
- Iannoni, A. P., Morabito, R. e Saydam, C. (2011) Optimizing large-scale emergency medical system operations on highways using the hypercube queueing model. *Socio-Economic Planning Sciences* 45, 105-117.
- Iannoni, A. P., Morabito, R. (2007). A multiple dispatch and partial backup hypercube queueing model to analyze emergency medical systems on highways. *Transportation Research E*, 43(6): 755-771.
- Iannoni, A. P., Morabito, R. e Saydam, C. (2007) An optimization approach for ambulance location and the districting of response segments on highways. *European Journal of Operational Research*, v.145, n.2, p.528-542.
- Jyaraman, V. e Srivastava, R. (1995) A Service Logistics Model for Simultaneous Siting of Facilities and Multiple Levels of Equipment. *Computers & Operations Research* 22 (2), 191-204.
- Larson, R.C. (1974) A hypercube queueing model for facility location and redistricting in urban emergency services. *Computers & Operations Research* 1, 67-95.
- Marianov V. e ReVelle C. (1996) The Queueing Maximal Availability Location Problem: A model for the siting of emergency vehicles. *European Journal of Operations Research* 93, 110-120.
- Matteson, D. S., McLean, M. W., Woodard, D. B., Henderson, S. G. Forecasting Emergency Medical Service Call Arrival Rates. School of Operations Research and Information Engineering. Cornell University (2010).
- Medina, A. C. (1996) *Modelos para dimensionamento de frota e localização de embarcações para atendimento de acidentes marítimos*. 1996. 240p. Dissertação (Mestrado) – Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Naval e Oceânica, São Paulo.
- Rajagopalan, H.K.; Saydam, C. e Xiao, J. (2008) A multi-period set covering location model for dynamic redeployment of ambulances. *Computers & Operations Research* 35 (3), 814-826.
- ReVelle, C. S. e Hogan, K., (1989) The maximum availability location problem. *Transportation Science* 23, 192-200.
- Revelle, C. S. e Marianov, V. (1991) A probabilistic FLEET model with individual vehicle reliability requirements. *European Journal of Operations Research* 53, 93-105.
- Schilling, D.; Elzinga, D. J.; Cohon, J.; Church, R. e ReVelle, C. (1979) The Team/Fleet Models for Simultaneous Facility and Equipment Siting. *Transportation Science*, v. 13, n. 2, p. 163-175.
- Schilling, D. A.; Jayaraman, V. e Barkhi, R. (1993) A review of covering problems in facility location. *Location Science* v 1 n 1, 25-55.
- Schmid, V. e Doerner, K. F. (2010) Ambulance location and relocation problems with time-dependent travel times. *European Journal of Operations Research* 207, 1293-1303.

Singer, M. e Donoso, P. (2008) Assessing an ambulance service with queuing theory. *Computers & Operations Research* 35, 2549-2560.

Takeda, R. A.; Widmer, J. A. e Morabito, R. (2007) Analysis of ambulance decentralization in an urban emergency medical

service using the hypercube queueing method. *Computers & Operations Research* 34 - 727-741.

Toregas, C.R.; Swain, R.; ReVelle, C.S. e Bergman, L., (1971) The location of emergency service facilities. *Operations Research* 19, 1363-137.