Information Relaxation Bounds for Evaluating Ambulance Dispatch and Allocation Policies

Lavanya Marla
University of Illinois at Urbana-Champaign, lavanyam@illinois.edu

Achal Bassamboo
Kellogg School of Management, Northwestern University, a-bassamboo@kellogg.northwestern.edu

Milind Sohoni
Indian School of Business, milind_sohoni@isb.edu

Motivation. Emergency medical services (EMS) provide quick and timely first response to callers. Most of these systems are structured as public-private partnerships. In these cases, often the government agency is faced with the question of deciding whether the private operator can serve a certain population at the required service level wit the given fleet size or ambulance allocation irrespective of the tactical strategies the operator employs. On the other hand, the operator too needs to decide which tactical or strategic measure is likely to serve best the requirement of the population and the government; without first solving for allocation, redeployment or dispatch policies (all of which are hard problems with numerous suboptimal methodologies). For example, it is useful to find if modifying the dispatch policy will give more benefit as opposed to dynamic redeployment without employing an array complex methods available for these problems.

Recently, therefore, Yue, Marla, and Krishnan (2012), Maxwell et al. (2014) have suggested two types of bounds for such systems - omniscient information relaxations and resource relaxations. We attempt to improve upon the omniscient information relaxations in Yue, Marla, and Krishnan (2012) to obtain tighter bounds, which Maxwell et al. (2014) describe as an open problem.

Problem Statement. Let R be a set of (historical or simulated) calls or requests from the service region, Q_i be the set of feasible bases for request i (including null), ordered from nearest to farthest. Let x_{js} be the (binary variable) defined for $s \in Q_i$ and $i \in R$, is 1 iff assignment to i is from base s (s can be null). Let P_i^s be the set of parents of request i via base s ($j \in P_i^s$, that is, j arrives before i, j completes after i arrives, and s is in both Q_j and Q_i). Let L_{is} be the cost of assigning ambulance from base s to request i and a_s be the number of ambulances allocated to base s (with the allocation vector $A = a_1,..., a_{|A|}$).

The most common dispatch policy adopted by EMSs is a myopic greedy policy, to assign the nearest free ambulance as a call arrives, or drop the call (assuming it behaves as a loss system). The 'best possible' (but unachievable) dispatch policy can be found by an omniscient dispatcher,

who knows all the calls a priori, and finds an assignment of ambulances to calls that minimizes the total cost; and is a lower bound to any primal dispatch policy. In our experiments, the omniscient bound is too loose to determine if a new dispatch policy that is better than greedy can be devised for a given call log. We tighten this bound building on theory from Brown, Smith, and Sun (2010).

Approach An information filtration is the order in which information is revealed to the decision-maker. For example, in a natural non-anticipative filtration, calls to an EMS are known only upon arrival; whereas they are known apriori in an anticipative information filtration. Let $\mathbb{G} = (\mathcal{G}_0, ..., \mathcal{G}_T)$ be a relaxation of the natural filtration $\mathbb{F} = (\mathcal{F}_0, ..., \mathcal{F}_T)$, such that $\mathcal{F}_t \subseteq \mathcal{G}_t$ for all t. We denote $\mathcal{A}_\mathbb{G}$ as the set of policies that are adapted to $\mathbb{G}$ and $\mathcal{A}_\mathbb{F}$ as the set of policies that are adapted to $\mathbb{F}$, where $\mathcal{A}_\mathbb{F} \subseteq \mathcal{A}_\mathbb{G}$.

Consider a DP with a minimization objective, such as dynamic dispatch/redeployment, as:

$$J_t(a_0, ..., a_{t-1}) = \inf_{a_t \in A_t(a_0,...,a_{t-1})} \{c_t(a_0, ..., a_t) + \mathbb{E}[J_{t+1}(a_0, ..., a_{t-1})|\mathcal{F}_t]\} \tag{1}$$

The omniscient optimization problem, where all future calls in the problem are known apriori is

$$\min_{a_t \in A_t(a_0,...,a_{t-1}), \forall t} \sum_t \hat{c}_t(a_0, ..., a_t) \tag{2}$$

We now define a penalty function p. p is said to be feasible if it is measurable and satisfies $\mathbb{E}(p) \leq 0$ under any non-anticipative filtration. Weak duality holds in the information relaxation setting, as follows. $\mathbb{E}[c(\alpha_F)] \geq \mathbb{E}[c(\alpha_F) + p(\alpha_F)] \geq \inf_{\alpha_G \in \mathcal{A}_\mathbb{G}} \mathbb{E}[c(\alpha_G) + p(\alpha_G)]$.

Now we define the structure of the penalty. Let $w_0(a, \omega), ..., w_T(a, \omega)$ be a series of generating functions where each w_t depends only on the first $t+1$ actions $(a_0, ..., a_t)$ of a. Let $p_t(a) = p_t^1(a) - p_t^2(a)$, where $p_t^1(a) = \mathbb{E}[w_t(a)|\mathcal{F}_t]$ and $p_t^2(a) = \mathbb{E}[w_t(a)|\mathcal{G}_t]$ and $p(a) = \sum_{t=0}^{T} p_t(a)$. This penalty satisfies: (i) For all $\alpha_F \in \mathcal{A}_\mathbb{F}$, i.e., non-anticipative actions, we have $\mathbb{E}[p_t(\alpha_F)|\mathcal{F}_t] = 0 \forall t$, and (ii) $(p_0(a), ..., p_T(a))$ is adapted to $\mathbb{G}$ and p_t depends only on the first $t+1$ actions $(a_0, ..., a_t)$ of a.

To capture p as the value function, we can write $p_t^1(a)$ as $\min \mathbb{E}[c(a_{t+1}, ..., a_T)|\mathbb{F}]$ and $p_t^2(a)$ as $\min \mathbb{E}[c(a_{t+1}, ..., a_T)|\mathbb{G}]$. Because $\mathbb{G}$ is non-anticipative, we will always get a feasible penalty. The bound formulation is:

$$J_t^\mathbb{G}(a_0, ..., a_{t-1}) = \inf_{a_t \in A_t(a_0,...,a_{t-1})} \left\{c_t(a_0, ..., a_t) + p_t^1(a_0, ..., a_t) - p_t^2(a_0, ..., a_t) + \mathbb{E}[J_{t+1}(a_0, ..., a_{t-1})|\mathcal{G}_t]\right\} \tag{3}$$

$$= \inf_{a_t \in A_t(a_0,...,a_{t-1})} \Bigg\{c_t(a_0, ..., a_t) + \min_{a'} \mathbb{E}[c(a'_{t+1}, ..., a_T)|\mathbb{F}] - \min_{\bar{a}} \mathbb{E}[c(\bar{a}_{t+1}, ..., \bar{a}_T)|\mathbb{G}] + \mathbb{E}[J_{t+1}(a_0, ..., a_{t-1})|\mathcal{G}_t]\Bigg\} \tag{4}$$

$$\geq \inf_{a_t \in A_t(a_0,...,a_{t-1})} \Bigg\{c_t(a_0, ..., a_t) + \min_{a'} \mathbb{E}[c(a'_{t+1}, ..., a_T)|\mathbb{F}] - \mathbb{E}[c(\bar{a}_{t+1}, ..., \bar{a}_T)|\mathbb{G}] + \mathbb{E}[J_{t+1}(a_0, ..., a_{t-1})|\mathcal{G}_t]\Bigg\} \tag{5}$$

We replace p_t^2 by a feasible solution under the non-anticipative filtration $\mathbb{G}$ instead of the optimal solution, as shown in (5). This will still be a lower bound, which could be, in some cases, worse than the anticipative bound. Therefore, it is important to ensure that $\min_{a'} \mathbb{E}[c(a'_{t+1},...,a_T)|\mathbb{F}] \geq \mathbb{E}[c(\bar{a}_{t+1},...,\bar{a}_T)|\mathbb{G}]$, so that we have a positive penalty and hence a meaningful bound. Expanding (5) recursively, we write this as a single, large, integer program.

First, we revisit some existing notation and introduce some new notation. As defined earlier, R is the original call log, consisting of calls indexed by i. Q_i is the priority queue of bases that can serve each call, ranked from the least response time to the highest. A null element is added to the end of each call's priority queue, indicating that the call cannot be served and is dropped. Let s be the index of the bases; and s_l^i denote the lth ranked base in the priority queue Q_i of call i. Let the call log R span the time horizon $[0,T]$ and t be the time index. The call log is divided into time intervals along which the penalty function is formulated. Let R_{pt} be the pth sample path ($p \in P$) of calls (sampled by a look-ahead algorithm), satisfying: (i) R_{pt} has exactly the same calls as R from $[0,t]$, and (ii) calls from $(t,T]$ are sampled according to the same underlying arrival process as R but are not the same as in R. Let $\bar{R}_{pt}$ be the component of the call log R_{pt} that spans the interval from $(t,T]$.

The variables in the formulation are as follows. $\boldsymbol{x'_{pt}}$ represents the solutions obtained by implementing a non-anticipative, myopic nearest-free-ambulance policy in $[t,T]$ on call log R_{pt}. Correspondingly, $\boldsymbol{x''_{pt}}$ represents the solutions obtained by implementing an anticipative policy (defined below) in $[t,T]$ on call log R_{pt}. The variable β_{is} captures the number of ambulances free at base s at the time of arrival of call i.

Additionally, let $\mathbf{x^*_{pt}}$ be the optimal solution obtained after solving the omniscient optimal dispatch (7)-(8) on the call log $\bar{R}_{pt}$. Note that the cost of the omniscient optimal dispatch solution $\sum_{i\in R_t}\sum_{s\in Q_i} x^*_{pt,is}L_{is}$ is always less than the cost of the same set of calls for any other dispatch policy, specifically, an $x_{pt,is}$ obtained by an assignment over the entire call log R_{pt}. We utilize this property to find a good feasible solution in (18).

$$\min \sum_{i\in R}\sum_{s\in Q_i} x_{is}L_{is} + \frac{1}{|P|}\left\{\sum_t\sum_{p\in P}\sum_{i\in\bar{R}_{pt}}\sum_{s\in Q_i} x'_{pt,is}L_{is} - \sum_t\sum_{p\in P}\sum_{i\in\bar{R}_{pt}}\sum_{s\in Q_i} x''_{pt,is}L_{is}\right\} \tag{6}$$

$$\text{s.t.} \sum_{s\in Q_i} x_{is} = 1, \qquad \forall\, i\in R \tag{7}$$

$$x_{is} + \sum_{j\in P_i^s} x_{js} \leq a_s \qquad \forall\, i\in R \tag{8}$$

$$\sum_{s\in Q_i} x'_{pt,is} = 1 \qquad \forall\, i\in\bar{R}_{pt}, \forall\, p\in P, \forall\, t \tag{9}$$

$$\beta_{pt,is} = a_s - \sum_{k\in P_i^s} x'_{pt,ks} \qquad \forall\, s=1,...,Q_i, \forall\, i\in R_{pt}, \forall\, p\in P, \forall\, t \tag{10}$$

$$x'_{pt,is_1^i} \leq \beta_{pt,is_1^i} \qquad \forall i\in\bar{R}_{pt}, \forall\, p\in P, \forall\, t \tag{11}$$

$$x^{'}_{pt,is^i_1} \geq \frac{1}{M}\beta_{pt,is^i_1} \qquad \forall i \in \bar{R}_{pt}, \forall\, p \in P, \forall\, t \qquad (12)$$

$$x^{'}_{pt,is^i_l} \geq \frac{1}{M}\beta_{pt,is^i_l} - \sum_{n=1}^{l-1}\beta_{pt,is^i_n} \qquad \forall l = 2,...,|Q_i|, \forall\, i \in \bar{R}_{pt}, \forall\, p \in P, \forall\, t \qquad (13)$$

$$x^{'}_{pt,is^i_l} \leq 1 - \frac{1}{M}\sum_{n=1}^{l-1}\beta_{pt,is^i_n} \qquad \forall\, l = 2,...,|Q_i|, \forall\, i \in \bar{R}_{pt}, \forall\, p \in P, \forall\, t \qquad (14)$$

$$\sum_{i\in\bar{R}_{pt}}\sum_{s\in Q_i} x^{''}_{pt,is}L_{is} \leq \Gamma_t \sum_{i\in\bar{R}_{pt}}\sum_{s\in Q_i} x^{*}_{pt,is}L_{is} \qquad \forall\, p \in P, \forall\, t \qquad (15)$$

$$\sum_{s\in Q_i} x^{''}_{pt,is} = 1, \qquad \forall\, i \in \bar{R}_{pt}, \forall\, p \in P, \forall\, t \qquad (16)$$

$$x^{''}_{pt,is} + \sum_{j\in P^s_i} x^{''}_{pt,js} \leq a_s \qquad \forall\, i \in R_{pt}, \forall\, p \in P, \forall\, t \qquad (17)$$

$$x_{is} = x^{'}_{pt,is} = x^{''}_{pt,is} \qquad \forall\, i \in R_{[0,t]}, \forall\, s \in Q_i, \forall\, p \in P, \forall\, t \qquad (18)$$

$$x_{is} \in \{0,1\} \qquad \forall\, s \in Q_i, \forall\, i \in R \qquad (19)$$

$$x^{'}_{pt,is} \in \{0,1\} \qquad \forall\, i \in R_{pt}, \forall\, s \in Q_i, \forall\, p \in P, \forall\, t \qquad (20)$$

$$x^{''}_{pt,is} \in \{0,1\} \qquad \forall\, i \in R_{pt}, \forall\, s \in Q_i, \forall\, p \in P, \forall\, t \qquad (21)$$

x is the variable representing the ambulance allocations of the omniscient problem. (7)-(8) and (19) are the constraints for the pure omniscient problem. $x^{'}$ is the solution corresponding to the myopic dispatch problem, and spans $[t, T]$ for each time interval t. (9)-(14) and (20) are the constraints corresponding to the non-anticipative (myopic) solution at each time t, and for each sample path p. $x^{''}$ is the solution to the omniscient problem from time t to T. (15)-(17) and (21) are the constraints corresponding to a 'good' anticipative (feasible omniscient) solution at each time t. To ensure that the solution from (15)-(17) is a 'good enough' feasible solution, for each sample path p and time point t, we pre-compute the optimal solution $\mathbf{x^*_{pt}}$ over the call log $\bar{R}_{pt}$, and use that to constrain that the second term of the penalty does not exceed the pre-computed optimal solution by some factor Γ. Constraints (18) ensure that the solutions in the true path, and the non-anticipative (myopic) sample path and the anticipative sample path match between $[0, t]$.

For large values of Γ, the solution to this integer program produces loose or meaningless bounds that might be worse than the omniscient bound. As we lower Γ, we will find tighter bounds, and the integer program can become infeasible at values of Γ close to 1.

Results and Conclusions. Our results show that for a large Asisn city, given the current ambulance allocation, the greedy policy is near-optimal and no dispatch policy can improve the service level by more than 1%. Therefore, other actions of reallocation or increasing fleet size need to be used to improve service. By contrast, in other queueing settings modeled as an N-network (Harrison 1988), greedy policies are suboptimal but our lower bound provides a tight bound on the possible improvement achievable by any assignment policy.

References

Brown DB, Smith JE, Sun P, 2010 *Information relaxations and duality in stochastic dynamic programs. Operations Research* 58(4, Part 1 of 2):785–801.

Harrison JM, 1988 *Heavy traffic analysis of a system with parallel servers: Asymptotic optimality of discrete-review policies. The Annals of Applied Probability* 8(2):822–848.

Maxwell MS, Ni EC, Tong C, Henderson SG, Topaloglu H, Hunter SR, 2014 *A bound on the performance of an optimal ambulance redeployment policy. Operations Research* 62(5):1014–1027.

Yue Y, Marla L, Krishnan R, 2012 *An efficient simulation-based approach to ambulance fleet allocation and dynamic redeployment. Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, 398–405, AAAI'12 (AAAI Press), URL `http://dl.acm.org/citation.cfm?id=2900728.2900786`.