A Stochastic Programming Model for Casualty Response Planning during Catastrophic Health Events

Aakil Caunhye, Xiaofeng Nie

A Stochastic Programming Model for Casualty Response Planning during Catastrophic Health Events

Catastrophic health events (CHEs) are natural or man-made incidents that result in numbers of ill or injured people sufficient to overwhelm the immediate capabilities of emergency response and healthcare systems [1]. These events overpower hospitals even when the latter trigger contingency capacity surges that are usually sufficient to deal with regular disasters. Examples of CHEs include nuclear detonations, pandemics, and earthquakes. This work tackles one of the main issues of healthcare response to CHEs, which is to build a tactical response plan that incorporates 1) community-based surge capacity addition via alternative care facilities (ACFs), 2) casualty triage, and 3) self-evacuee movement, all in a single framework.

The framework is implemented using a three-stage stochastic programming model. With the passage of time, the initial chaos settles, casualties are discovered, some communication lifelines are restored, and information exchanges become more accurate. The three-stage stochastic programming methodology allows us to use each stage to make decisions that apply to a different time period after a CHE has struck. In the first stage, the model locates ACFs, in the second stage it allocates casualties for triage, and in the third stage it allocates casualties for treatment based on triage decisions. Allocation decisions are generated for different disaster scenarios, whilst location decisions are made irrespective of and robust to these scenarios. This allows emergency responders to setup ACFs right after the disaster and wait for more information on actual damages to choose the allocation plan for the scenario that most closely matches these damages.

In disaster situations where time is of the essence, large solution times are unacceptable and detrimental to the practicality of the resulting response plan. To address the large solution time issue, we propose an algorithm, based on Benders' decomposition (see [2]), to obtain good response plans fast. The algorithm is an iterative process where binary decision variables (for ACF location), the problematic decision variables, are isolated in a master problem and solved separately from other decision variables. The master problem generates ACF location solutions which are then fed to a dual subproblem to generate cuts. The challenge of constructing the algorithm is to make it generate fast improving solutions.

We address three major issues in Benders' decomposition. The first is that feasibility cuts, generated when the master location solution is infeasible, are weak. We reduce the number of possible location solutions that are infeasible, by further constraining the master problem. This is done by adding valid inequalities that ensure that new ACFs located provide enough capacity to triage all casualties and treat all low-priority ones when all existing ACFs are maximally utilized. The second issue is that optimality cuts can also be weak. To improve their strength, we add flow cover inequalities to the deterministic equivalent to tighten the polyhedron of its linear relaxation. Magnanti and Wong [3] prove that tightening the polyhedron of the linear relaxation yields stronger cuts. The third issue is degeneracy, meaning that there may be multiple possible cuts for a location solution. To tackle this issue, we perturb the dual subproblem so that it generates maximal nondominated Benders' cuts (see [4]), which are stronger cuts, in degenerate cases.

Summary of results

We implement the model and algorithm in the case study of an earthquake situation in California. We make several remarks from sensitivity analyses. The first remark is that having extra response teams at one's disposal to set up more ACFs improves the optimal objective value, but with decreasing marginal improvement. A limit is reached beyond which the optimal value remains constant, no matter how many more ACFs are available. Knowledge of this limit helps avoid an over-supply of ACFs. In the case where the total resource available for ACFs is fixed, a bigger budget of smaller ACFs tends to improve the optimal solution, with a decreasing rate. The choice of a bigger budget of smaller ACFs is favourable as it allows the placement of ACFs close to hospitals or other ACFs to shorten the travel times of high-priority and low-priority casualties for treatment.

Another interesting remark of practical importance is that the triage-to-treatment capacity ratios for both ACFs and hospitals exhibit similar effects on the optimal objective value. Our facilities tend to give the best optimal plans when they have triage-to-treatment capacity ratios whereby they can treat all their triaged casualties. At this point, there is no need to transport casualties out for treatment.

With regards to the algorithm, we have tested the effectiveness of each modification we have made to the original Benders' decomposition algorithm. In every case where neither CPLEX nor the conventional Benders' algorithm is able to obtain an initial feasible solution within 1 hour, our modified Benders' decomposition finds a near-optimal solution (actually reaching close to optimal in most cases within 20 minutes). The performance of the algorithm shows that initial cuts are very strong and solutions improve very fast compared to the traditional Bender's decomposition or CPLEX.

References

[1] Toner, E., R. E. Waldhorn, C. Franco, A. Norwood, B. Courtney, K. Rambhia, M. Watson, T. V. Inglesby. 2010. The next challenge in healthcare preparedness: catastrophic health events. Center for Biosecurity of UPMC, Baltimore, MD.

[2] Benders, J. F. 1962. Partitioning procedures for solving mixed-variables programming problems. Numer. Math. 4(1) 238-252.

[3] Magnanti, T. L., R. T. Wong. 1981. Accelerating Benders decomposition: algorithmic enhancement and model selection criteria. Oper. Res. 29(3) 464-484.

[4] Sherali, H. D., B. J. Lunday. 2013. On generating maximal nondominated Benders cuts. Ann. Oper. Res. 210(1) 57-72.