

Management of Emergency Response Teams under Stochastic Demands

Iliya Markov¹ and Sacha Varone²

¹*HES, Haute Ecole de Gestion de Genève, Campus de Battelle, Bâtiment F, 7 route de Drize, 1227, Carouge, Switzerland*

²*Decision Sciences, Haute Ecole de Gestion de Genève, Campus de Battelle, Bâtiment F, 7 route de Drize, 1227, Carouge, Switzerland*

Keywords: Contractual Constraints, Decision Support Tool, Emergency Response Teams, Rare Skills, Stochastic Optimization.

Abstract: We propose a stochastic optimization model for the composition of emergency response teams. An emergency intervention requires first an evaluation of the situation which results in the need of different skills. People involved in the response team must therefore comply with the required skills, be available, with a past and future workload respecting contractual compliance. In addition, we must also anticipate the possibility of future interventions that will require rare skills. It is the uncertain future demand for these skills that introduces stochasticities to the system. Since shifting agents between emergencies may be impossible or impractical, we would like to ensure that rare skills are not wasted but assigned to the emergency that most needs them. We model this with a mixed integer linear program implemented in AMPL and capable of being solved in real-time on common solvers.

1 INTRODUCTION

The composition of emergency response teams is a task which is often done manually by a team lead who evaluates the nature and seriousness of the accident at hand. The team lead has a list of people on standby that she can choose from. She knows their availability, qualifications and working hours constraints, and must compose the team in real time taking care that scarce resources in terms of rare skills might be required in future interventions. The success of the intervention thus rests on the knowledge of the team lead. Her absence, in the case of illness for example, may hinder the fluent course of emergency response operations. This highlights the importance of developing a decision support tool to aid the composition of response teams and reduce the risk of disturbances in the absence or replacement of the team lead. To our knowledge, there is no existing tool for the real-time composition of emergency response teams in optimal or near-optimal way.

Emergencies are very often unpredictable. Their critical nature, however, means that they must be evaluated and responded to very quickly. The key factor in responding to emergency situations is not one of monetary costs but of consistently meeting important benchmarks of performance. Examples include incident response times, sufficient personnel with appropriate skills, efficient distribution and management of the equipment, etc.

Road safety, particularly after a traffic accident, is a prime example of an emergency that requires an immediate response. In addition to clearing the road of vehicles and debris, which may require heavy machinery, a response team must close lanes, manage traffic and sometimes investigate accident causes, particularly when criminal behavior is suspected. Road safety therefore requires a near immediate response to prevent danger but also the presence of appropriate skill levels in terms of forensics, first aid, road maintenance, the operation of special machinery, etc. In addition, the availability of agents, their contractual working hours, overtime pay and resource utilization must all be taken into account as they themselves have an affect on what the optimal team composition is.

The stochastic nature of emergencies also affects the optimal agent assignments, resource utilization, etc. In the case of rare skills, for example, we would have to consider future emergencies that occur while the current emergency is being handled. Since shifting agents between emergencies may be impossible or impractical, we would like to ensure that rare skills are not wasted but assigned to the emergency that most needs them. This means that it could be optimal for an agent to remain idle and wait for the next emergency rather than being assigned now.

In this article we propose a mixed integer linear programming model whose objective function and constraints accommodate the stochastic aspect of emergencies. The team composition in terms of agent

assignments must take into account the possibility of future emergencies that may require, among other factors, rare agent skills. Our model is implemented in AMPL and solved to optimality with a common solver. The model is tested on randomly generated data instances of realistic size and solves to optimality in real time. Future practical extensions and implementations of the model may necessitate the development of appropriate heuristics.

This article is structured as follows. After this introduction, section 2 offers a review of related literature from several domains that our research touches on. Section 3 develops a base stochastic model for emergency team composition and then offers several extensions to it. Section 4 describes the results of the experimental testing. Finally, section 5 concludes and stresses on further developments and applications.

2 LITERATURE REVIEW

Scheduling and assignment problems are some of the most studied and applied problems in combinatorial optimization. They are encountered in almost all businesses with multiple employees especially where large and complex task structures need to be efficiently executed. These problems are very often modeled as integer or binary programs, which makes it relatively easy to express the objectives and constraints. On the other hand, they are very often NP-hard and thus difficult to solve. (Cattrysse and Wassenhove, 1992) present a survey of algorithms to solve reasonable instances of the generalized assignment problem (GAP) to optimality. Large instances, however, may necessitate the use of heuristics. (Pentico, 2007) presents the most useful variations of the assignment problem since the publication of the seminal work of (Kuhn, 1955) on the Hungarian method.

A host of authors have contributed to the development of variations of the assignment problem, whether in terms of modeling or solution techniques. (Caron et al., 1999) and (Volgenant, 2004) look at the assignment problem with seniority and task priority constraints. (Volgenant, 2004) offers an improvement to the coefficient scaling approach of (Caron et al., 1999) by successively re-optimizing assignment problems of increasing size. The latter method has the complexity of a standard linear assignment problem and can be successfully applied to a bottleneck problem. (Felici and Mecoli, 2007) propose a version of the GAP where agents exhibit join and split preferences. The split preference problem can be formulated as a minimum cost flow on a suitable graph. It is a polynomial time algorithm and can be solved with a

common solver in reasonable time. On the other hand, the join preference problem is NP-hard and needs to be solved by heuristics. (Gilberti and Righini, 2007) design an integer programming model for the multi-period allocation of duties in rapid intervention teams with the objective of minimizing the maximum intervention time. Moreover, the assignment has to be equitable so that the effort is evenly distributed among the available teams.

Stochasticities in the assignment problem have also been addressed in the literature. (Albareda-Sambola and Fernández, 2000) study the GAP with independent and identically distributed demands for agents following a Bernoulli distribution. Their model is based on probabilistic constraints and is solved as a transportation problem. It proves to be the best in terms of the value of the objective function and in avoiding infeasibilities compared to the alternative approaches they consider. (Spoerl and Wood, 2004) model uncertainty in the elastic GAP¹ when the amount of resources used by the agents follow a normal distribution. They penalize resource overuse and find substantial value of the stochastic solution. (Albareda-Sambola et al., 2006) study a two-stage stochastic GAP where task availability is distributed following a Bernoulli distribution. In the first stage, initial assignments are made before task availability is known with certainty. After task availability is revealed, reassignments may be made at a cost. The authors construct a convex approximation of the non-convex objective function and solve the problem using exact algorithms.

(Toktas et al., 2004) study stochastic resource-constrained GAPs where the resource capacities follow an unknown probability distribution which can be sampled. They develop exact and approximate solution techniques based on Lagrangian relaxation to solve the collectively capacitated GAP (CCGAP)² with three alternative recourse functions. The stochastic approaches perform significantly better than those using expected values of the capacities in a deterministic setting. In addition, approximate stochastic programming techniques are as efficient as deterministic techniques. (Toktas et al., 2006) review alternative methodologies for dealing with capacity constraints under uncertainty of the resource capacities when stochastic programming is infeasible or un-

¹The elastic generalized assignment problem (EGAP), addressed by (Brown and Graves, 1981), is an extension of the GAP that treats capacity constraints as soft constraints, with capacity overuse penalized in the objective function.

²The collectively capacitated generalized assignment problem is a version of the GAP with multiple resources collectively capacitated for all agents, without the requirement for one-to-one matching between tasks and agents.

desirable, for example when it is very difficult to define a recourse function.

The emergency team composition problem belongs to the vast array of literature on disaster and emergency planning. (Green and Kolesar, 2004), (Altay and Green, 2006), and (Simpson and Hancock, 2009) offer detailed discussions on the development of OR methodologies in emergency planning and disaster response. They describe the City of New York through its fruitful cooperation with the RAND Corporation as the principal arena for much of the early research on police, fire and ambulance services allocation and deployment. (Kolesar and Walker, 1974), for example, develop a model and heuristic to determine the location of fire trucks in New York. OR applications began in the context of improving performance in organized environments. Emergencies, on the other hand, represent the disruption of such designs and highlight the practical needs of many organizations.

One of the first influential papers outside the urban services group was (Brown and Vassiliou, 1993), which presents a decision support system for the real-time tactical allocation of response units in the context of disaster relief following a major earthquake in Greece. Similar research is carried out by other authors as well. (Rolland et al., 2010) include workload and labor requirements, precedence constraints, resource availability, and critical deadlines. Skills are also considered by modeling them as team-task mismatch costs in the objective function. (Barbarosoğlu and Arda, 2004), (Beraldi et al., 2004), (Yi and Kumar, 2007), and (Mete and Zabinsky, 2010), among others, apply stochastic programming techniques to the management of disasters of various natures. These studies, however, do not necessarily belong to the assignment problem realm.

Emergency response planning, as briefly described above, generally has a space dimension and involves routing problems. In terms of the management of emergency teams, (Rolland et al., 2010), for example, deal with the time-periodic scheduling of assignments in the context of disaster response situations. Their contribution consists in proposing a hybrid meta-heuristic which allows for the real-time solution of the disaster recovery scheduling problem. Unlike us, however, they work in a deterministic environment and view teams as already existing. Our model on emergency team composition and management therefore is a step forward in a new direction. It allows us to both position ourselves in the area of emergency response and develop our contribution of a new approach in terms of team composition.

Even though optimal team composition is an area

of interest to OR, the focus has generally not been on emergency response. (Boon and Sierksma, 2003), for example, develop an assignment problem with more than 20 additional constraints which is applied to team composition in volleyball.

(Shipley and Johnson, 2009) treat the question of choosing project team members who meet the project's goals for the preferred cognitive style. The paper presents a fuzzy logic and an algorithm based on belief in the fuzzy probability of a cognitive style fitting a defined goal. Thus the paper addresses uncertainty in decision making and facilitates the member selection process. It illustrates how OR can adapt behavioral research techniques. (Woodruff, 2010) also treats the problem of choosing optimal teams under predefined policies. It treats the players in terms of their specific skills rather than interactions and demonstrates that if satisfaction of the policy is difficult then simply choosing a set of individually best members will often not result in the best policy conformance. In other words, a team of the best will not be the best team.

The model we propose in this article considers workers with different skills. In this respect it touches on a significant body of literature that treats the problem of cross-trained workers from various perspectives. (Campbell and Diaby, 2002) develop heuristics for the model of (Campbell, 1999) which assigns cross-trained workers across different departments at the beginning of a shift. The departmental utility function is non-linear and the resulting model can be viewed as a variant of the GAP. It models cross-trainings by specifying an agent's capability as a percentage of a fully qualified agent in that task. (Campbell, 2011) takes the problem of scheduling and assignment of cross-trained workers to a stochastic environment with variable demands. The second-stage is a recourse on realized demands.

(Sayın and Karabatı, 2007) also build a framework based on the non-linear nature of the departmental objective function of (Campbell, 1999). They propose a two-stage model, where two objective functions, departmental utility and skill improvement, are considered in the assignment of cross-trained workers. (Fowler et al., 2008) develop heuristics for the model of (Wirojanagud et al., 2007) which determines different staffing decisions, such as hiring, firing and cross-training of workers, in order to minimize workforce related costs over multiple periods. It is a multi-period model where cross-training is done with the purpose of satisfying labor requirements in each period. In this article, we attempt to find the best assignment of agents with multiple skills by taking into account the need for rare skills in the future. The train-

ing of agents is not an objective of our model.

3 PROBLEM FORMULATION

The methodology we propose here is based on a two-stage stochastic linear programming framework. In such a framework, the first stage consists of a structural component which is fixed while the second stage varies with uncertainty. In general, the second stage is represented by a collection of distinct future scenarios which are governed by stochastic parameters. In other words, the stochastic parameters have a discrete distribution or are represented by a finite sample from a continuous distribution. A number of studies concern the selection of scenarios in stochastic programs. (Jenkins, 2000), for example, proposes a mathematical programming model for the selection of a limited set of scenarios that have the maximum similarity with all potential scenarios.

The first-stage decision variables in a stochastic program represent decisions that must be made now, before the future state is known. These are proactive decisions often concerned with planning. They are independent of the values of the stochastic parameters in the second stage. This property is known as non-anticipativity, i.e. our actions now cannot take advantage of knowledge of the future. The second-stage decision variables, commonly referred to as recourse or reactive variables, represent the actions we take after the values of the stochastic parameters are revealed. For every scenario, there is a set of recourse variables linked to the first-stage decision variables by scenario-dependent constraints. Solving the stochastic program thus yields an optimal first-stage decision under all scenarios, i.e. all considered realizations of the stochastic parameters.

In a team composition framework, let's consider an emergency that occurs now. This emergency is composed of a set of tasks, each of which requires a set of skills to be simultaneously present in every agent assigned to it. In the simplest case, an emergency may only consist of one task. Moreover, a task may only require one skill. An agent, in turn, may possess a number of skills. Each task requires a number of skilled agents to be assigned to it.

We consider a set of tasks I indexed by i , a set of agents \mathcal{J} indexed by j , a set of skills \mathcal{S} indexed by s , and a set of resources \mathcal{R} indexed by r . Moreover, we consider a set of future emergencies \mathcal{F} indexed by f that develop while the current emergency is being handled. Any emergency, be it the current one or the future ones, is composed of a subset of I . The number of tasks and the number of needed agents in the fu-

ture is not known a priori. It is stochastic and depends on the particular emergency scenario that develops. Therefore, we have to decide on the most appropriate assignments so that the staffing requirements of both the current and the future emergencies are satisfied with agents possessing the necessary skills and at the lowest possible cost. The assignment of agents with rare skills becomes critical here. Assigning such an agent to a trivial task in the current emergency means that his rare skills will not be available if a future emergency needs them.

The index f designates the types of scenarios or, in other words, the types of emergencies that can develop while still handling the current emergency. In a generalized framework, we assume that the lapse between two consecutive emergencies of the same type f follows an exponential distribution. The probability of an occurrence during the handling of the current emergency therefore does not depend on the past due to the memoryless properties of the exponential distribution. These probabilities can be normalized to add up to one and used in the objective function. Once a future emergency occurs, the system is re-optimized taking care of any parameter changes in terms of worked hours, availability, etc. Due to the continuous nature of the exponential distribution we can assume that two emergencies cannot occur at the same time. The decision maker will thus need to optimize with the occurrence of every emergency to find out the most appropriate assignments. We introduce the following notation in terms of parameters:

c_{ij}	cost of assigning agent j to task i
q_{js}	skill qualification of agent j (1 if skill s is present, 0 otherwise)
e_{is}	skill requirement of task i (1 if skill s is required, 0 otherwise)
n_i	number of agents required for task i in the current emergency
n_{fi}	number of agents required for task i in future emergency f
u_{ir}	use of resource r per agent assigned to task i
t_r	total available amount of resource r
a_j	availability of agent j
d	duration of the current emergency
d_f	duration of future emergency f
h_j	number of already worked hours by agent j
H_j	contractual number of working hours of agent j
p_f	probability of occurrence of emergency f

All of these parameters are evaluated at the time of every optimization (re-optimization). For example, the available agents at any given optimization

step will not include those that are out handling an emergency. In a realistic environment, therefore, this optimization algorithm will be part of a larger system that tracks the status of all agents and updates their relevant parameters. We should also mention that the costs in the above notation should not necessarily be understood as monetary costs. These could be utilities, hierarchy levels, execution efficiencies, response times (depending on where the agents are located) etc.

A more detailed discussion of p_f is in order. We can imagine that there is an indefinite number of different emergencies that can occur considering all the different parameters that describe them. This is why the model we describe below assumes that we have already performed data mining on the historical data we have and extracted several emergency types (in terms of duration, tasks, etc.) that are closest to our complete sample of emergencies. Looking at the frequency of these emergencies, we can calculate the probability that they develop while the current emergency is being handled. Normalizing these probabilities to add up to one, we obtain the parameters p_f .

The decision variables in the current and future emergencies are:

- x_{ij} assignment of agent j to task i in the current emergency (binary)
- x_{fij} assignment of agent j to task i in future emergency f (binary)

The stochastic optimization model thus takes the following form:

$$\min \sum_{i \in I} \sum_{j \in \mathcal{J}} c_{ij} x_{ij} + \sum_{f \in \mathcal{F}} p_f \sum_{i \in I} \sum_{j \in \mathcal{J}} c_{ij} x_{fij} \quad (1)$$

$$\text{s.t.} \quad \sum_{i \in I} (x_{ij} + x_{fij}) \leq 1, \quad f \in \mathcal{F}, j \in \mathcal{J} \quad (2)$$

$$e_{is} x_{ij} \leq q_{js}, \quad i \in I, j \in \mathcal{J}, s \in \mathcal{S} \quad (3)$$

$$e_{is} x_{fij} \leq q_{js}, \quad f \in \mathcal{F}, i \in I, j \in \mathcal{J}, s \in \mathcal{S} \quad (4)$$

$$\sum_{j \in \mathcal{J}} x_{ij} \geq n_i, \quad i \in I \quad (5)$$

$$\sum_{j \in \mathcal{J}} x_{fij} \geq n_{fi}, \quad f \in \mathcal{F}, i \in I \quad (6)$$

$$(h_j + d) \sum_{i \in I} x_{ij} + (h_j + d_f) \sum_{i \in I} x_{fij} \leq H_j, \quad f \in \mathcal{F}, j \in \mathcal{J} \quad (7)$$

$$\sum_{i \in I} \sum_{j \in \mathcal{J}} (u_{ir} x_{ij} + u_{ir} x_{fij}) \leq t_r, \quad f \in \mathcal{F}, r \in \mathcal{R} \quad (8)$$

$$x_{ij} \leq a_j, \quad i \in I, j \in \mathcal{J} \quad (9)$$

$$x_{fij} \leq a_j, \quad f \in \mathcal{F}, i \in I, j \in \mathcal{J} \quad (10)$$

$$x_{ij}, x_{fij} \in \{0, 1\}, \quad f \in \mathcal{F}, i \in I, j \in \mathcal{J} \quad (11)$$

Expression (1) designates the minimization of the

cost of all assignments, i.e. the costs of assigning agents to the current emergency and all future emergencies according to their probability. Constraints (2) impose that an agent is assigned to at most one task either in the current emergency or (any or all of) the future emergencies (as we do not know which one will occur first). As explained above, it may be optimal for the agent to remain on standby and wait for the next emergency. Constraints (3) ensure that an agent is assigned to a task in the current emergency only if he has all the necessary skills. Constraints (4) extend this to the future emergencies. Constraints (5) and (6) require a minimum number of agents per task in the current and future emergencies. Constraints (7) establish that an agent can be assigned to the current or future emergencies only if his contractual working hours do not expire during the emergency to which he is assigned. We have the usual assumption here that working time only accrues during the handling of an emergency. Standby is not considered as working time in our model. Moreover, we assume that the agent should be present during the whole duration of the emergency.

Constraints (8) impose a limit on resources. We have multiple resources collectively capacitated for all agents, an approach that appears as early as in (Mazzola and Neebe, 1986). Given the fact that we do not require a one-to-one matching between agents and tasks, our model in its essence resembles the CC-GAP proposed by (Toktas et al., 2004). In our model, however, resource consumption is not agent specific. It depends on the number of agents assigned but not the specific agents that are assigned. Constraints (9) and (10) specify agent availability and, finally, constraints (11) establish the variable domains.

The strict separation between the parameters relating to the current emergency and those relating to the future emergency types allows for a more flexible model. Thinking about the emergencies that may develop, we think of the emergency types that we defined above. Once an emergency develops, however, it is no longer an emergency type but an actual emergency. A quick assessment of it will allow us to attribute the appropriate parameters in terms of what tasks it is composed of, how many agents each task needs and what the emergency's duration is.

Tasks in the above model are in essence task types. The set I describes all possible task types that can occur³, such as road maintenance, forensics, etc. Stochasticity in our model is manifested through the types of emergency events in the scenarios. Based on the event that occurs, some tasks will not exist (they

³The definition of I can be based on historical analysis of the emergency data.

do not need any agents) or may need a certain number of agents. More precisely, therefore, stochasticity in the model is governed by the values of the parameters n_{fi} . In addition to the base model presented above, some of the constraints can be substituted or supplemented by the alternatives below. [task types]

Overtime: We can modify the working hours constraints (7) to allow overtime. Agent overtime will then be penalized in the objective function. The idea is similar to that of (Brown and Graves, 1981), cited in (Spoerl and Wood, 2004), where they penalize agent capacity under- and over-use. Working time, in this sense, is agent capacity. We define the following parameters:

c_j^+ cost of overtime of agent j
 α, β weighting parameters

The decision variables associated with this approach are:

y_{fj}^+ overtime hours of agent j (linear)
 y_{fj}^- under-time hours of agent j (linear)

Introducing the above notation, the model becomes:

$$\min \alpha(\text{Objective 1}) + \beta \sum_{f \in \mathcal{F}} p_f \sum_{j \in \mathcal{J}} c_j^+ y_{fj}^+ \quad (12)$$

s.t. (2), (3), (4), (5), (6)

$$h_j + d \sum_{i \in \mathcal{I}} x_{ij} + d_f \sum_{i \in \mathcal{I}} x_{fij} - y_{fj}^+ + y_{fj}^- = H_j, \quad (13)$$

$$f \in \mathcal{F}, j \in \mathcal{J}$$

$$y_{fj}^+, y_{fj}^- \geq 0, \quad f \in \mathcal{F}, j \in \mathcal{J} \quad (14)$$

(8), (9), (10), (11)

In addition to minimizing assignment costs, the objective function (12) also minimizes the penalty associated with the overtime hours in constraints (13).

Capacitated overtime: We can also add maximum allowable overtime on constraint (13). In the model below, h_j^+ stands for the maximum allowable overtime hours for agent j . The capacitated overtime model thus becomes:

min Objective 12

s.t. (2), (3), (4), (5), (6), (13)

$$y_{fj}^+ \leq h_j^+, \quad f \in \mathcal{F}, j \in \mathcal{J} \quad (15)$$

(8), (9), (10), (11), (14)

Shared resources: Constraints (8) model task-based resource consumption per agent, i.e. adding an additional agent to any task increases that task's resource consumption. We may call these *individual* resources because they are allocated to each agent

individually. An example of such resources are face masks. However, we may want to model an alternative resource consumption pattern that is not linked to every incremental agent. The model below is appropriate for resources such as vehicles that drive the agents to the emergency. One vehicle should be dispatched for every “n” agents regardless of which task of the emergency they are assigned to. We call these *shared* resources. To model them, we only need to supply one additional parameter and two decision variables:

k_r number of agents per resource r
 v_r use (consumption) of resource r in the current emergency (integer)
 v_{fr} use (consumption) of resource r in the future emergency f (integer)

The alternative resource consumption model thus becomes:

min Objective 1

s.t. (2), (3), (4), (5), (6), (7)

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} x_{ij} \leq k_r v_r, \quad r \in \mathcal{R} \quad (16)$$

$$\sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} x_{fij} \leq k_r v_{fr}, \quad f \in \mathcal{F}, r \in \mathcal{R} \quad (17)$$

$$v_r + v_{fr} \leq t_r, \quad f \in \mathcal{F}, r \in \mathcal{R} \quad (18)$$

(9), (10), (11)

$$v_r, v_{fr} \in \text{integer}, \quad f \in \mathcal{F}, r \in \mathcal{R} \quad (19)$$

Constraints (16) and (17) increment the resource consumption variables v_r and v_{fr} based on how many agents are assigned to each emergency⁴. Constraints (18) then ensure that the resource availability is respected. There could be different variations of this constraint based on the specific circumstances or the preferences of the decision maker.

Soft constraints: Depending on the policy of the decision maker, some of the constraints above may be softened. Whether two agents may be able to carry out a task that requires three agents, whether these agents will be able to execute the tasks if they are not fully qualified or lack some of the necessary resources are questions that should be analyzed by each of the model's users. A van intended to carry four people may be able to carry five, but a fire cannot be put out without a fire engine.

In the above model and its extensions, most constraints (apart from overtime) must be perfectly satis-

⁴The reason we sum over the agent assignment variables (x_{ij} and x_{fij}) as opposed to the parameters specifying the number of required agents (n_i and n_{fi}), is to preserve the flexibility to soften constraints (5) and (6) as described below.

fied in order to find a feasible solution for the emergency team composition problem. However, in many practical applications the constraints do not need to be perfectly satisfied. In the case of soft constraints, the constraint may be violated and the solution is still acceptable, although not as “desirable”. A common way to specify soft constraints is by adding penalty variables to each soft constraint. These variables take the value of zero when the constraint is satisfied and then increase with the degree of constraint violation. There is a penalty parameter associated with each penalty variable. The resulting penalty function is minimized in the objective.

In the model above, some of the constraints that can be softened include the number of agents per task constraints (5, 6) and the qualification constraints (3, 4). In the first case, the step may be necessary if an insufficient number of agents are available. The model will include penalties on task under-staffing which may vary depending on task criticality. In the second case, the step may be necessary if the workforce is not sufficiently qualified yet. Given a penalty on not satisfying the qualification constraints, the model will satisfy the skill requirements of a subset of the tasks. For the remaining tasks, under-qualified agents will be assigned. Softening both groups, it would make sense to attach higher penalty to the number of agents constraint as assigning less than perfectly qualified agents to a task is usually better than assigning no agents at all.

Regarding the resource constraints (8, 18), in the model above they are infeasibility constraints. They are either satisfied or not and do not affect agent assignments. Our presumption is that a task consumes the same amount of resource regardless of which particular agent is assigned to it. Therefore, being hard our resource constraints serve two purposes. First, they ensure that an assignment is feasible only if there are enough available resources. Second, they allow us to keep track of the used resources. Softening the resource constraints will mean that agents may be under-supplied with resources in some or all of the scenarios. In particular, softening constraints (18) may be sensible if we allow for “overcrowded” vehicles, for example.

4 EXPERIMENTAL RESULTS

In order to test the performance of our model, we carried numerical experiments on randomly generated instances of the stochastic model with capacitated overtime constraints and two types of resources – the individual resources modeled by constraints (8)

and the shared resources modeled by constraints (16) through (19). That is, our model is composed of the objective function (12) with α and β equal to 1, plus constraints (2) - (6), (8) - (11), and (13) - (19). All of our constraints except working hours remain as hard constraints. Our idea was to generate problem instances of realistic size as depicted in table 1 so as to assess the model’s performance as if applied to a realistic setting. The tests were carried out on a 2.66 GHz Intel Core 2 Quad PC with 8 Gb of memory running a 64-bit Windows 7. The problem was modeled in AMPL and we used an AMPL client associated with the Gurobi 5.0 commercial solver.

Table 1: **Random data.** This table presents the key parameters used for the generation of random instances.

Parameters	Size
Task types	15
Agents	300
Skills	10
Individual resources	10
Shared resources	4
Emergency scenarios	8

We generated a set of random instances involving 300 agents and 15 tasks. Each emergency consists of a random subset of the tasks and each task requires a random number of agents.

Most of the model’s parameters are based on uniform distributions, with rounding or flooring where necessary. As an exception, the number of required agents per task is modeled using an exponential distribution with a rate of 0.3 and then floored to the nearest integer. The choice of the exponential distribution allows tasks involving no agent with a high frequency. Such tasks do not belong to the associated emergency. This choice of parameter also allows occasional large numbers. There are a total of 10 skills that we consider, of which seven are frequent skills, present in almost every agent, while the three remaining skills are rare.

There are 10 types of individual resources, i.e. resources modeled according to constraints (8), and four types shared resources – those modeled by constraints (16) through (19). Finally, we have eight future emergency scenarios. In every generated instance, the probability of an emergency of each type occurring during the handling of the current emergency is sampled from a predefined historical distribution. These probabilities are then normalized to add up to one. As explained above, in a realistic situation, the model is run iteratively with every emergency occurrence taking into consideration the latest informa-

tion on the model’s parameters.

As table 2 below shows, we generated 200 random instances of the model. Since all of our constraints are hard, seven of the instances (3.50%) turned out to be infeasible.

Table 2: **Instance statistics.** This table presents various characteristics of the randomly generated instances.

Before presolve	Size	
Number of instances	200	
Feasible instances	193	
Number of variables	45,336	
Number of constraints	452,983	
After presolve	Mean	St. Dev.
Number of variables	24,515.40	2,253.77
Number of constraints	4,803.12	38.24
Solution time (sec)	3.91	0.43

The table also shows that the original instances have 45,336 variables and 452,983 constraints. After AMPL’s presolve, however, the number of variables is almost halved and the number of constraints is reduced by a factor of 90. The instances solve for an average of 3.91 seconds. Based on a *t*-distribution, the 95% confidence interval for the mean is from 3.85 to 3.97 seconds. Using the free CBC 2.7.5 solver on the same instances, the solution times generally remain below 20 sec.

In a second experiment, we doubled the value of each parameter in table 1. The resulting instances have 325,336 variables and 6,455,890 constraints. After presolve, the number of variables is reduced approximately three times and the number of constraints is reduced to around 19,000, i.e. less than 1% of the original value. Using Gurobi 5.0, the solution times are between 20 and 30 sec, which is still a comparatively low value for the size of the problem we have.

The results therefore suggest that instances of realistic size can be solved in reasonably short time on a commercial or free solver without the use of any heuristics. Further developments of the model through the addition of client-specific constraints or the softening of some of the constraints will inevitably make the model more cumbersome and therefore impact the solution time. Future testing and possible commercial applications will show whether heuristics or meta-heuristics need to be developed.

5 CONCLUSIONS

In this article, we proposed a stochastic optimization model for the composition of emergency response teams taking into account agent skills, availability, contractual compliance and resource utilization. The experimental results showed the viability of the model as a first step of a multi-stage project. The next steps will see our model enriched with alternative objective functions and more constraints, including client specific ones in order to increase the marketability of our prospective tool. Finally, our model is intended to evolve into a web-based decision support application with on-demand access.

The potential market of our tool includes enterprises in the areas of emergency response services and call centers. Most of the medium and large IT enterprises, especially the ones working with the banking sector, ensure very short response times and should thus be able to compose expert teams in real time.

There is one more message we want to convey with our work. The generalized assignment problem is NP-hard. Being a variation and a stochastic extension of the GAP, our model can be viewed as “difficult” to solve. Researchers often tend to develop heuristics when it comes to solving difficult problems, even of modest size. Nevertheless, we have shown that before developing heuristics, it makes sense to model the problem in a mathematical programming language and try an available commercial solver, or a free solver such as those from the COIN-OR project (www.coin-or.org).

REFERENCES

- Albareda-Sambola, M. and Fernández, E. (2000). The stochastic generalized assignment problem with Bernoulli demands. *Sociedad de Estadística e Investigación Operativa Top*, 8(2):165–190.
- Albareda-Sambola, M., van der Vlerk, M. H., and Fernández, E. (2006). Exact solutions to a class of stochastic generalized assignment problems. *European Journal of Operational Research*, 173(2):465–487.
- Altay, N. and Green, W. G. (2006). OR/MS research in disaster operations management. *European Journal of Operational Research*, 175(1):475–493.
- Barbarosoğlu, G. and Arda, Y. (2004). A two-stage stochastic programming framework for transportation planning in disaster response. *Journal of the Operational Research Society*, 55(1):43–53.
- Beraldi, P., Bruni, M., and Conforti, D. (2004). Designing robust emergency medical service via stochastic programming. *European Journal of Operational Research*, 158(1):183–193.

- Boon, B. H. and Sierksma, G. (2003). Team formation: Matching quality supply and quality demand. *European Journal of Operational Research*, 148(2):277–292.
- Brown, G. G. and Graves, G. W. (1981). Real-time dispatch of petroleum tank trucks. *Management Science*, 27(1):19–32.
- Brown, G. G. and Vassiliou, A. L. (1993). Optimizing disaster relief: Real-time operational and tactical decision support. *Naval Research Logistics*, 40(1):1–23.
- Campbell, G. M. (1999). Cross-utilization of workers whose capabilities differ. *Management Science*, 45(5):722–732.
- Campbell, G. M. (2011). A two-stage stochastic program for scheduling and allocating cross-trained workers. *The Journal of the Operational Research Society*, 62(6):1038–1047.
- Campbell, G. M. and Diaby, M. (2002). Development and evaluation of an assignment heuristic for allocating cross-trained workers. *European Journal of Operational Research*, 138(1):9–20.
- Caron, G., Hansen, P., and Jaumard, B. (1999). The assignment problem with seniority and job priority constraints. *Operations Research*, 47(3):449–453.
- Cattrysse, D. G. and Wassenhove, L. N. V. (1992). A survey of algorithms for the generalized assignment problem. *European Journal of Operational Research*, 60(3):260–272.
- Felici, G. and Mecoli, M. (2007). Resource assignment with preference conditions. *European Journal of Operational Research*, 180(2):519–531.
- Fowler, J. W., Wirojanagud, P., and Gel, E. S. (2008). Heuristics for workforce planning with worker differences. *European Journal of Operational Research*, 190(3):724–740.
- Gilberti, D. and Righini, G. (2007). Optimization of duties assignment in emergency services. In *Service Operations and Logistics, and Informatics, 2007, IEEE International Conference on, Philadelphia, PA*, pages 1–6.
- Green, L. and Kolesar, P. (2004). Improving emergency responsiveness with management science. *Management Science*, 50(8):1001–1014.
- Jenkins, L. (2000). Selecting scenarios for environmental disaster planning. *European Journal of Operational Research*, 121(2):275–286.
- Kolesar, P. and Walker, W. E. (1974). An algorithm for the dynamic relocation of fire companies. *Operations Research*, 22(2):249–274.
- Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97.
- Mazzola, J. and Neebe, A. (1986). Resource-constrained assignment scheduling. *Operations Research*, 34(4):560–572.
- Mete, H. O. and Zabinsky, Z. B. (2010). Stochastic optimization of medical supply location and distribution in disaster management. *International Journal of Production Economics*, 126(1):76–84.
- Pentico, D. W. (2007). Assignment problems: A golden anniversary survey. *European Journal of Operational Research*, 176(2):774–793.
- Rolland, E., Patterson, R., Ward, K., and Dodin, B. (2010). Decision support for disaster management. *Operations Management Research*, 3(1):68–79.
- Sayın, S. and Karabatı, S. (2007). Assigning cross-trained workers to departments: A two-stage optimization model to maximize utility and skill improvement. *European Journal of Operational Research*, 176(3):1643–1658.
- Shiple, M. F. and Johnson, M. (2009). A fuzzy approach for selecting project membership to achieve cognitive style goals. *European Journal of Operational Research*, 192(3):918 – 928.
- Simpson, N. C. and Hancock, P. G. (2009). Fifty years of operational research and emergency response. *Journal of the Operational Research Society*, 60(S1):126–139.
- Spoerl, D. and Wood, R. (2004). A stochastic generalized assignment problem. Working paper. Department of Operations Research, Naval Postgraduate School, Monterey, California.
- Toktas, B., Yen, J. W., and Zabinsky, Z. B. (2004). A stochastic programming approach to resource-constrained assignment problems. *Stochastic Programming E-Print Series*, <http://www.speps.org>.
- Toktas, B., Yen, J. W., and Zabinsky, Z. B. (2006). Addressing capacity uncertainty in resource-constrained assignment problems. *Computers & Operations Research*, 33(3):724–745.
- Volgenant, A. (2004). A note on the assignment problem with seniority and job priority constraints. *European Journal of Operational Research*, 154(1):330–335.
- Wirojanagud, P., Gel, E. S., Fowler, J. W., and Cardy, R. (2007). Modelling inherent worker differences for workforce planning. *International Journal of Production Research*, 45(3):525–553.
- Woodruff, C. J. (2010). Multivariate optimisation for procurement of emergency services equipment - teams of the best or the best of teams? *European Journal of Operational Research*, 205(1):186–194.
- Yi, W. and Kumar, A. (2007). Ant colony optimization for disaster relief operations. *Transportation Research Part E: Logistics and Transportation Review*, 43(6):660–672.