

Emergency Decision Making: A Dynamic Approach

Zhenyu Yu

School of Economics and Management
Tongji University
freshyu2002@163.com

Chuanfeng Han

School of Economics and Management
Tongji University
juanfeng12@163.com

Yefeng Ma

Institute of Public Safety Research
Tsinghua University
mayf10@mails.tsingua.edu.cn

ABSTRACT

The dynamic nature of emergency decision making exerts difficulty to decision makers for achieving effective management. In this regard, we suggest a dynamic decision making model based on Markov decision process. Our model copes with the dynamic decision problems quantitatively and computationally, and has powerful expression ability to model the emergency decision problems. We use a wildfire scenario to demonstrate the implementation of the model, as well as the solution to the firefighting problem. The advantages of our model in emergency management domain are discussed and concluded in the last.

Keywords

Emergency decision making, dynamics decision making, Markov decision process

INTRODUCTION

Emergency management consists of several phases, including prevention, mitigation, preparedness, response, and recovery. This article concentrates on the decision making during the response phase, because in this period, emergency decision making is dramatically different from traditional decision making that applies rational decision models (Flin, 2001; Klein, 1986), and effective decision making can be especially challenging under the stressful and time-pressured conditions of extreme events (Boin et al., 2005).

During a wildfire, due to the rapidly development of the fire and lack of dynamic thinking ability, decision makers tend to make decision based on their mental model. Modelling the interaction of decision effect and fire involvement, probabilistically and quantitatively, will support the decision maker under dynamic decision making environment.

In this paper, we first investigate the dynamic decision problem, and map it into a formal schema. Then the problem is formulated as a Markov decision process (MDP) model, which makes it possible to apply mathematical methods. An example of wildfire fighting is also given to demonstrate the application of MDP. Finally, we discuss the advantages of MDP in emergency management.

DYNAMIC DECISION MAKING

According to Brehmer (1992), a dynamic decision making problem has four characteristics, which are:

- A series of decisions is required to reach the goal.
- The decisions are not independent; that is, later decisions are constrained by earlier decisions, and in turn, constrain those that come after them.
- The state of the decision problem changes, both autonomously and as a consequence of the decision makers' actions.
- The decisions have to be made in real time.

Obviously, most emergency decision problems share such features, and the results of dynamic decision making theory have enormous potential in emergency management domain. Some sequential decision problems can be modeled as constrained optimization, yet, it is impossible to find analytical solutions for most dynamic decision problems (Rapoport, 1975). Stochastic control theory sheds light on these problems. A generic dynamic decision problem schema is illustrated in Figure 1 (Bertsekas, 1987). This system has a set of possible actions \mathbf{A} , a set of system states \mathbf{X} , an output set \mathbf{Y} , and a set of uncertainty factors \mathbf{E} , which is further decomposed into uncertainty about the state (\mathbf{E}_x) and uncertainty about the output (\mathbf{E}_y).

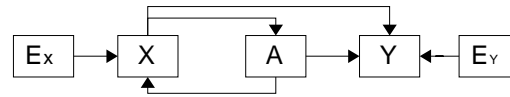


Figure 1. A generic dynamic decision problem

At a specified time step, the decision maker observes the state of a system. Based on the observed state, he/she chooses an action. The action choice yields two outcomes: the decision maker receives an output (immediate reward or cost), and the system evolves to a new state at a subsequent time step according to a probability distribution determined by the action choice. At this subsequent time step, the decision maker faces a similar problem, but the state of the system and possible actions may be different from the previous one.

MODEL FORMULATION

Markov decision process (MDP) provides a mathematical framework for modeling dynamic and probabilistic decision making problems, and it has already been used to model real world problems in a variety of disciplines, including operations research, ecology, economics, and communications engineering (Puterman, 2005).

An MDP problem can be defined as a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R})$, where

\mathcal{S} is a set of states.

\mathcal{A} is a set of actions.

\mathcal{T} is a transition function, a mapping specifying the probability $\mathcal{T}(s_i, a_i, s_j)$ of going into state s_j if action a_i is executed when the current state is s_i .

\mathcal{R} is a reward function that gives a finite numeric reward value $\mathcal{R}(s_i, a_i, s_j)$ obtained when the system goes from state s_i to state s_j as a result of executing action a_i .

Here, it should be noted that the time element is implicit in this kind of expression. Since most of the emergency response would be finished at some time, we can enumerate the time steps as $t = 1, 2, \dots, T_{end}$.

The dynamic decision problem in the previous section can be modeled with MDP. First, MDP is obviously a sequence decision process. Second, the dependence of decisions can be evaluated by the value function, in terms of accumulative utility as,

$$V_t^*(s) = \max_{a \in \mathcal{A}} [\sum_{s' \in \mathcal{S}} \mathcal{T}(s, a, s') [\mathcal{R}_t(s, a, s') + V_{t+1}^*(s')]]. \quad (1)$$

Third, the system transitions can be acquired with two components. One component consists of autonomous changes of the system in probability of $P(s_j|s_i)$, whereas the other component consists of the state transition controlled with the actions in probability of $P(a_i, s_j|s_i)$. Then, we can use Bayes's theorem to calculate the posterior probability of the system transition $\mathcal{T}(s_i, a_i, s_j)$, denoted by $P(s_j|s_i, a_i)$ in conditional probability terms,

$$P(s_j|s_i, a_i) = \frac{P(a_i, s_j|s_i)P(s_j|s_i)}{\sum_{s_k \in \mathcal{S}} P(a_i, s_k|s_i)P(s_k|s_i)}. \quad (2)$$

Fourth, we can use computers to solve the MDP problems, and therefore support the decision making in real time. The solution of an MDP problem can be calculated using value iteration (Bellman, 1957), policy iteration (Howard, 1960), and other approximate algorithms (Kocsis and Szepesvári, 2006).

A feature of MDP is that the solution of a problem, which is called the optimal policy, denoted as π^* ,

$$\pi_t^*(s) = \arg \max_{a \in \mathcal{A}} [\sum_{s' \in \mathcal{S}} \mathcal{T}(s, a, s') [\mathcal{R}_t(s, a, s') + V_{t+1}^*(s')]]. \quad (3)$$

Equation (3) explicitly gives a trajectory of actions for achieving the goal--maximizing the total expected utility or minimizing the cost. This means the decision makers know that at what time step, which action should be taken to gain an optimal outcome. This helps the decision makers gain their control of the dynamic process of

emergencies and make effective decisions. We will use a concrete example to explain how the MDP model can be formulated and support the emergency decision making in the next section.

MODEL APPLICATION

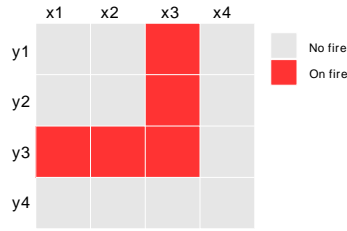


Figure 2. Initial state of a wildfire scenario

We use a supposed wildfire scenario to illustrate the application of MDP in emergency context. Considering a forest region consisted of 4 by 4 grids (as shown in Figure 2), each of which is denoted with its coordinate (x_i, y_j) . The states are the fire status of the grids, which are denoted as $s(x_i, y_j)$, see Equation (4). Each grid has at most 8 neighbors, and the neighboring relationship is as $I(x_{i_1}, y_{j_1}, x_{i_2}, y_{j_2})$, defined in Equation (5). Whether a grid on fire is affected by its neighbors, the more neighbors on fire, the larger probability it will be ignited. We suppose the probability 0.1, 0.6, 0.9 for the number of on fire neighbors 1-3, 4-6, and 7-8 respectively, as expressed in Equation (6). At each time step, a decision maker can put out an on fire grid with a successful probability of 0.6. The action put out a fire at the grid (x_i, y_j) is marked as $put(x_i, y_j)$, as in Equation (7). The reward function is the cost of the wildfire, and we assume each on fire grid at a time step costs 1 unit value, see Equation (8). For example, the cost of the state in Figure 2 is -5. The above description use a factored manner of MDP, and this can save us a huge amount of space to express the problem.

$$s(x_i, y_j) = \begin{cases} 0 & (x_i, y_j) \text{ not on fire} \\ 1 & (x_i, y_j) \text{ on fire} \end{cases} \quad (4)$$

$$I(x_{i_1}, y_{j_1}, x_{i_2}, y_{j_2}) = \begin{cases} 1 & \text{if } (x_{i_1}, y_{j_1}) \text{ and } (x_{i_2}, y_{j_2}) \text{ neighboring} \\ 0 & \text{if } (x_{i_1}, y_{j_1}) \text{ and } (x_{i_2}, y_{j_2}) \text{ not neighboring} \end{cases} \quad (5)$$

$$P(s(x_i, y_j) = 1 | s(x_i, y_j) = 0, I(x_i, y_j, x_{i_2}, y_{j_2})) = \begin{cases} 0 & \text{else} \\ 0.1 & 1 \leq \sum_{(x_{i_2}, y_{j_2}) \in \{(x,y) | I(x,y,x_i,y_i)=1\}} s(x_{i_2}, y_{j_2}) \leq 3 \\ 0.6 & 4 \leq \sum_{(x_{i_2}, y_{j_2}) \in \{(x,y) | I(x,y,x_i,y_i)=1\}} s(x_{i_2}, y_{j_2}) \leq 6 \\ 0.9 & \sum_{(x_{i_2}, y_{j_2}) \in \{(x,y) | I(x,y,x_i,y_i)=1\}} s(x_{i_2}, y_{j_2}) \geq 7 \end{cases} \quad (6)$$

$$P(s(x_i, y_j) = 0 | s(x_i, y_j) = 1, put(x_i, y_j)) = 0.6 \quad (7)$$

$$\mathcal{R}_t = -1 \times \sum_{1 \leq i \leq 4, 1 \leq j \leq 4} s^{(t)}(x_i, y_j), t = 1, 2, \dots, T_{end} \quad (8)$$

$$\min \sum_{t=1}^{T_{end}} \mathcal{R}_t \quad (9)$$

The goal of decision makers is to minimize the total cost of the wildfire during the firefighting process (Equation 9); in other words, to put out the fire as soon as possible. This is quite a typical problem, since the state of the fire is spread autonomously and affected by actions the decision maker executed, both in a probabilistic manner.

This problem is simulated with RDDDL (Sanner, 2010), and solved with PROST (Keller and Eyerich, 2012). The solution of the problem (Figure 2) is depicted in Table 1. At each time step, the RDDDL drives the state transition according to the model definition, and send the state to the PROST. The PROST receives the state and calculates the best action, and sends the action back to RDDDL to drive the next state. The action choice is based on value function. For example, the action set of the state of Figure 2 includes $put(x_3, y_1)$, $put(x_3, y_2)$, $put(x_3, y_3)$, $put(x_1, y_3)$, and $put(x_2, y_3)$. The value of each action is -92.2, -91.9, -91.4, -92.1, and -92.2 accordingly. (These values are calculated according to PROST, and only approximate the true value function.) Therefore, the optional action of this state is to put out the fire at (x_3, y_3) .

Table 1 clearly shows the state transition according to the fire spread and the effect of firefighting actions. Although the action has a probability of failure, there exists some grids worth putting out with priority [e.g.

grid(x_3, y_1)], while some not [e.g. grid(x_1, y_2)]. This is helpful, since for human decision makers it is hard to distinguish such difference. Moreover, the explicit solution of action sequence would alleviate the decision pressure, so even a less experienced decision maker can cope with such complex emergency situation.

Time	$t=0$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
State	□□■□ □□■□ ■■■□ □□□□	□□■□ □□■□ ■■○□ □□□□	□□■□ □□○□ ■■○□ □□□□	□□■□ □□○□ ■■○□ □□□□	□■○■ □□○□ ■■○□ □□□□	■■○■ ■□○□ ■■○□ □□□□
Action	$put(x_3, y_3)$	$put(x_3, y_2)$	$put(x_3, y_1)$	$put(x_3, y_1)$	$put(x_2, y_3)$	$put(x_2, y_3)$
Time	$t=6$	$t=7$	$t=8$	$t=9$	$t=10$	$t=11$
State	■■○■ ■□○□ ■○○□ □□□□	■■○■ ■■○■ ■○○□ □□□□	■■○■ ■■○■ ○○○□ □□□□	■■○■ ■■○■ ○○○□ □□□□	■■○■ ■■○○ ○○○□ □□□□	■■○■ ■■○○ ○○○□ □□□□
Action	$put(x_1, y_3)$	$put(x_1, y_3)$	$put(x_4, y_2)$	$put(x_4, y_2)$	$put(x_1, y_1)$	$put(x_1, y_2)$
Time	$t=12$	$t=13$	$t=14$	$t=15$	$t=16$	$t=17$
State	■■○■ ■■○○ ○○○□ □□□□	○■○■ ■■○○ ○○○□ □□□□	○■○■ ■○○○ ○○○□ □□□□	○■○■ ○○○○ ○○○□ □□□□	○■○○ ○○○○ ○○○□ □□□□	○○○○ ○○○○ ○○○□ □□□□
Action	$put(x_1, y_1)$	$put(x_2, y_2)$	$put(x_1, y_2)$	$put(x_4, y_1)$	$put(x_2, y_1)$	
	□ no fire	■ on fire	○ put out			

Table 1. Dynamic decision process illustration

DISCUSSION

Constrained by psychological limitations along with social and organizational factors, decision makers demonstrate bounded rationality (Simon, 1997). Decision makers make inferences about the uncertain environments of emergency events under constraints of limited time, limited knowledge, and limited computational capacities (Gigerenzer, 2004). MDP can serve as a basic tool to support the dynamic emergency decision making problems, and makes the decision problems quantitatively represented and computationally solved.

Using MDP as a formal model of dynamic decision making problems will help decision makers build up an integrated comprehension of the emergency management. The formalization of states of a system requires the consideration of all possible domains relating to the emergencies' evolvement and emergency management. The actions describe the consequences of different response tasks and missions. The uncertainty may be in the actions or in the exogenous factors, and therefore the transition probability of states can be described as a joint probability of the objective world and human intervention, for which it is easier to deal with separately. These measures reduce the uncertainty and complexity of the decision environment.

Moreover, MDP even works in some situations without complete information. The information of the disaster is impossible fully observed due to the distortion and delay of information. Partially observable Markov decision process is an extension of MDP (Cassandra, 1998), which enables the agent make inference from the partially observed information of the states and estimate the most likely states where it might lie on.

Last but not least, MDP can be implemented as software or systems for real-time decisions support. Online algorithms of MDP (Keller and Helmert, 2013) generate policies in an interactive manner, which is able to generate policies according to given states in real-time. As a result, the stress of the decision makers would be alleviated under time pressure and urgency with such decision support systems.

CONCLUSION

Dynamic decision making is inevitable in emergency management, particularly in the response phase. This article uses mathematical method to model the dynamic decision problem. MDP can essentially grasp the

dynamic features of emergency decision problems, and execute smoothly in urgent situations. Hence, MDP helps relieve the negative factors of the emergency decision environment, and supports effective decision to the emergency managers.

ACKNOWLEDGMENTS

We thank all authors, program and local committee members, and volunteers for their hard work and contributions to the ISCRAM conference. We thank Cornelia Caragea for her valuable revision and comments. This work is supported by the Natural Science Foundation of China, Grant No.91024023, 91224003.

REFERENCES

1. Bellman, R. (1957) *Dynamic Programming*, Princeton University Press, Princeton, NJ.
2. Bertsekas, D. P. (1987) *Dynamic programming: Deterministic and stochastic model*, Prentice Hall, Upper Saddle River, NJ.
3. Boin, A., Hart, P., Stern, E., and Sundelius, B. (2005) *The politics of crisis management: Public leadership under pressure*, Cambridge University Press, New York, NY.
4. Brehmer, B. (1992) Dynamic decision making: Human control of complex systems, *Acta Psychologica*, 81, 3, 211–241.
5. Cassandra, A. R. (1998) *Exact and approximate algorithms for partially observable markov decision processes*, Brown University, Providence, RI.
6. Flin, R. (2001) Decision making and leadership in crises: the piper alpha disaster, In L. K. Comfort, A. Boin, and U. Rosenthal (Eds.), *Managing crises: Threats, dilemmas, opportunities* (pp. 103–118), Charles C Thomas, Springfield.
7. Gigerenzer, G. (2004) Fast and frugal heuristics: The tools of bounded rationality, In D. J. Koehler and N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 62–88), Blackwell Publishing Ltd, Malden, MA.
8. Howard, R. A. (1960) *Dynamic Programming and Markov Processes*, MIT Press, Cambridge, MA.
9. Keller, T., and Helmert, M. (2013) Trial-based Heuristic Tree Search for Finite Horizon MDPs, *ICAPS 2013*.
10. Keller, T., and Eyerich, P. (2012) PROST: Probabilistic Planning Based on UCT, *ICAPS 2012*.
11. Kocsis, L., and Szepesvári, C. (2006) Bandit Based Monte-Carlo Planning, In *Proceedings of the 17th European Conference on Machine Learning (ECML)*, 282–293.
12. Puterman, M. L. (2005) *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley, Hoboken, NJ.
13. Rapoport, A. (1975) Research Paradigms for Studying Dynamic Decision Behavior, *Utility, Probability, and Human Decision Making*, Vol. 11, 349–369.
14. Sanner, S. (2010) *Relational Dynamic Influence Diagram Language (RDDL): Language Description*.
15. Simon, H. A. (1997) *The Sciences of the Artificial*. MIT Press, Cambridge, MA.