17th ICCRTS
"Operationalizing C2 Agility"

Requirements for Data Mining the Decision Space

Topic(s): 4, 3, and 6

Authors: Jill L. Drury*, Gary L. Klein*, Scott Musman*, Yikun Liu†, Mark Pfaff†

POC: Jill L. Drury
The MITRE Corporation
202 Burlington Road, Bedford, MA 01730-1420
jldrury@mitre.org

Authors' affiliations:
*The MITRE Corporation
†Indiana University Indianapolis

Requirements for Data Mining the Decision Space

ABSTRACT

Our previous work showed that decision-making performance could be improved by providing a "decision space": information that summarizes the terrain of plausible outcomes resulting from a range of options. We have shown in earlier work that providing this decision space yields the first level of "option awareness," enabling decision makers to visually identify the most robust option that will provide satisfactory outcomes across the broadest swath of plausible futures. Our current work investigates a deeper level of option awareness, comprehending the conditions that tend to lead to better or worse outcomes for the options, which we believe can result from interactive data mining of a decision space. This paper documents our analysis of the requirements for a novel interactive decision support system (DSS) that will, for the first time, combine interactive data mining techniques with a frequency-formatted decision space visualization. Using mock-ups and a prototype of our initial DSS design, we illustrate how the requirements might be met.

Introduction

Command and control (C^2) operations often involve time-sensitive, safety-critical decision-making in uncertain circumstances. Accordingly, we are researching decision support techniques that have the potential to assist multiple C^2 domains, beginning with emergency response C^2 as an example. A broad consensus is emerging that emergency responders and crisis managers could benefit from this type of research because "decision support systems can be used to reduce the time needed to make crucial decisions regarding task assignment and resource allocation" (Thompson et al. 2006, p. 250). In addition, our own research (Drury et al. 2009) has shown that providing such decision spaces results in better, faster and more confident decisions.

In our experience, responders have told us "Just give us more situation awareness and we will make better decisions." Situation awareness (Endsley, 1988; Endsley, 2000) occurs when operators perceive facts about a volume of time and space, comprehend those facts, and are able to predict the state of the environment in the near future. But knowing a compendium of facts about the situation, which Hall et al. (2007) call the *situation space*, is a necessary yet insufficient prerequisite to decision making (e.g., Belton and Stewart, 2002). Decision makers must also know what options are available and analytical information that facilitates knowing which option is preferable; Hall et al. (2007) call this the *decision space*. When decision makers have sufficient decision space information to comprehend the relative quality of one option versus another, we say they have *option awareness* (Drury et al., 2009; Klein et al., 2010).

Currently, most decision support systems (DSSs) designed for emergency response concentrate on providing information about the situation space. The RimSim (Campbell et al. 2008) emergency response system, Zographos and Androutsopoulos's (2008) integrated hazardous materials routing and emergency response decision support system, the DIORAMA disaster management system (Kondaveti and Ganz, 2009), and commercial systems such as CoBRA (Defense Group, Inc., undated) all concentrate primarily on providing facts about the situation.

Provided only with a situation space, decision makers often generate an analysis of options – the decision space – in their heads. In familiar circumstances, but under time pressure and uncertainty, experienced decision makers assess the situation and respond with the first option that seems to be satisfactory (Lipshitz et al., 2001; Klein, 1999). When the relative quality of the different available options is not immediately apparent, decision makers mentally simulate sequentially the possible results of one option after another (Phillips et al., 2004), again stopping at the first apparently satisfactory option. There are limits, of course, to the number of options that can be considered under time pressure (Klein and Brezovic, 1986). As the number of possible options becomes overwhelming, unaided decision makers may simply default to the easiest option to implement rather than make an otherwise satisfactory choice (e.g., Sethi-Iyengar et al., 2004).

So, choosing a robust option from among many is difficult due to limitations in the brain's short-term memory capacity (Cantor, 2009). However, a computer-generated display of the decision space can offload this cognitive processing to the computer, which then displays the resulting range of outcomes for each potential option under various plausible environmental conditions. This visualization approach allows the decision maker to actually see the relationships between options that are otherwise obscured rather than requiring them to mentally simulate each one (Drury et al., 2009). By returning choice to a perceptual comprehension process, we enable decision makers to apply their more powerful visual, pattern matching capabilities rather than their more limited capacities for mental simulation.

In fact, our previous work showed empirically that providing a visualization of the decision space did improve decision-making accuracy, speed, and confidence (Drury et al., 2009, Liu et al., 2011). We showed that supplying this decision space yields the first level of option awareness, enabling decision makers to visually identify the most robust option that will provide satisfactory outcomes across the broadest swath of plausible futures. Our current work investigates a deeper level of option awareness, comprehending the conditions that tend to lead to better or worse outcomes for the options, which we believe can result from *interactive data mining of the decision-space.*

This paper provides a first look at how we plan to use data mining techniques to provide enhanced option awareness. It documents our analysis of the requirements for a novel interactive decision support system (DSS) that will, for the first time, combine interactive data mining techniques with a frequency-formatted decision space visualization, as described below. Using mock-ups and a prototype of our initial DSS design, we illustrate how the requirements might be met. This paper ends with our plans for completing and validating the new DSS, and our thoughts on applying decision space visualization techniques to other C^2 domains.

Visualizing the decision space

Computer-based forecasting models can assess dozens of options with hundreds of variations due to uncertainty, resulting in a landscape of plausible outcomes. Bankes (1993) terms this approach *exploratory modeling*. The variations are generated because each option has a number of variables associated with different elements of the situation space. Changing some variables results in different options, while changes in other variables (those that may interact with the options) are beyond a decision maker's control. For example of the latter case, responders may be alerted to an explosion but are initially uncertain regarding the size of the explosion. The forecasting model can recalculate the results of choosing a particular option many times, each time using a different value for the explosion's initial magnitude. Further, the forecasting model can be run multiple times to take into account the uncertainty regarding executing the course of action defined in the decision option. Consider how traffic congestion between the fire station and the site of the explosion introduces uncertainty regarding the percentage of fire trucks that will be successful in getting to the scene and also the times that may elapse before the trucks arrive. Because computing power is now much more readily available than in the past, it is possible to run many different combinations of explosion size, percentage of fire trucks arriving at the scene, and times for the trucks to arrive.

Each run of the model can be scored according to a multi-attributed utility function (Keeney and Raiffa, 1993). For example, runs to evaluate the efficacy of sending three fire trucks to the explosion can be scored using overall cost as a metric, adding the cost of sending the trucks to the property damage that ensues, the casualties suffered, and opportunity costs. Casualty cost can be calculated as the sum of medical care for injuries, lost productivity due to recovery time, and the value assigned to lives lost via insurance actuarial tables. Opportunity costs occur when resources that are allocated to the present event are needed for emergencies that happen in the near future. (Although mutual aid agreements exist to help neighboring municipalities, such a situation almost always involves sending responders from longer distances and thus slows response, which can be critical when responding to time-sensitive disasters such as large fires.) When there is a principled basis for doing so, the scoring components can be weighted and combined so that they are each given different emphases that reflect the values of the decision makers. Alternatively, each can be normalized to a common scale (e.g., z-scores) but not combined, and then plotted as described below, so that the components themselves can be compared.

The different combinations of variables over multiple runs results in a range of costs associated with each option. A frequency format approach to displaying the results depicts all of the values of the runs in a frequency distribution for each option. Prior research (Gigerenzer and Hoffrage, 1995; Hoffrage and Gigerenzer, 1998) has shown that information consumers can comprehend the results more readily when the information is presented in the form of frequency distributions instead of probabilities. Comparisons among distributions can be further facilitated by using graphical statistical summaries such as box plots (Tukey, 1977), with one box plot depicting the range of plausible results for each option. Figure 1 shows an example box-plot visualization showing the range of potential results from sending between zero and five fire trucks to an emergency event.

When scoring components are not combined, for each run a value could be plotted for each component, which could be color-coded to facilitate comparing costs among components. In this way, the z-scores for lives lost and for property damage under one option could be compared with each other and with those individual components under another option.

This exploratory modeling method of calculating multiple possible results is different from the frequently used method of calculating one run per option, based on the most likely conditions, in an attempt to find the optimal option. While an optimal plan will return the highest expected return on investment, under deep uncertainty (Lempert et. al., 2003), optimal strategies lose their prescriptive value if they are sensitive to uncertainty about situational conditions. That is, selecting an apparently optimal strategy may lead to poor outcomes if conditions diverge from the assumed values. Thus the option indicated by an optimal strategy can be a poor choice when there are multiple plausible futures for each option, as is the case in this example.

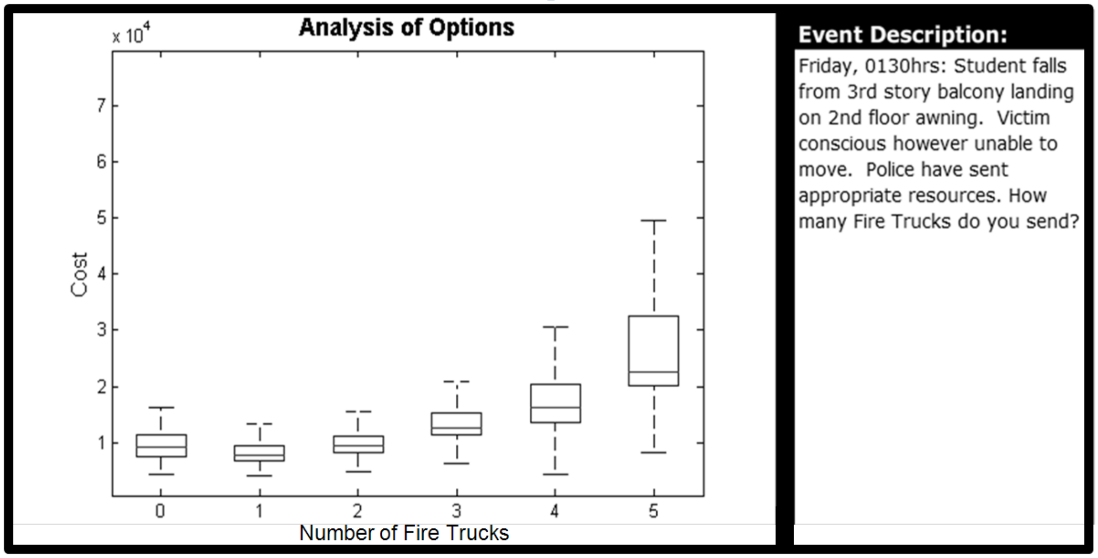

Figure 1: A surface visualization of decision space provides Option Awareness. Fire trucks are needed to get to victims on a second floor awning. The box plots reveal the distribution of resulting costs and show that over-allocating resources results in high costs.

Chandrasekaran (2005) and Chandrasekaran and Goldman (2007) note that the approach of running many forecasting model calculations using all plausible combinations of variables can reveal *robust* options. Robust options are those that result in acceptable outcomes across the broadest swath of plausible futures. In the example shown in Figure 1, sending one fire truck is forecasted to result in the lowest median cost, as indicated by a comparison of the lines bisecting each box plot. Sending one fire truck also results in the lowest-cost worst case and best case, as indicated by the top and bottom "whiskers" of each box plot. Winning on these three box-plot characterization parameters indicates that this option is the most robust: the most insensitive to variations in the elements in the situation space.

Perception of the relative robustness of alternative options is termed *level-1 option awareness* (Klein et al., 2010). We are now using data mining of the decision space as a way of investigating *level-2 option awareness*: comprehension of the relationships between factors underlying the option outcomes; we believe this can lead to *level-3 option awareness*: projection of these underlying relationships to adjusted or new options. In other words, decision makers who have attained level-3 option awareness understand the factors driving better or worse outcomes to the extent that they can take steps to minimize the factors that tend to lead to bad outcomes and maximize the factors that lead to good outcomes, thereby yielding improved options. We have conjectured that applying interactive data analysis techniques can enable C^2 operators to test hypotheses about apparent relationships between conditions and outcomes, which is key to attaining level-2 option awareness. These interactive visualizations should lead to better comprehension of these interactions, potentially enabling decision makers to modify options to develop more robust alternatives.

Data mining the decision space

In general, data mining is the analysis step of the process of discovering knowledge in data (Fayyed et al. 1996). More specifically, Wikipedia (2011) defines data mining as "the process of discovering new patterns from large data sets involving methods at the intersection of artificial intelligence, machine learning, statistics and database systems. The goal of data mining is to extract knowledge from a data set in a human-understandable structure."

Our initial efforts to apply data mining to a decision space have focused on mining the mission outcomes resulting from running an enhanced version of the NeoCITIES simulation model (Jones, 2006) for a variety of different operational conditions related to fire/rescue emergencies. NeoCITIES was developed as a team-based decision-making test environment in which test participants are presented with various emergency situations and are asked to make decisions regarding the number of resources to send to each emergency event. We took the NeoCITIES time-stepped, event-based simulation model originally developed at Pennsylvania State University and increased its fidelity.

In the example decisions that we describe in this paper, decision makers can choose to send more or less fire fighting equipment, and choose where the equipment originates (i.e., Station A or Station B). The simulation model also incorporates other characteristics associated with the emergency event such as its initial magnitude, the location of the fire, and traffic congestion. Based on these scenario characteristics the simulation computes how long it takes the fire fighting equipment to arrive on the scene, how much damage the fire causes, etc. Even though these latter scenario characteristics are outside the control of the decision maker, knowing about them helps the decision maker to understand how the conditions that characterize the emergency are affecting outcomes.

We used the Weka (University of Waikato, undated; Witten et al., 2011) open-source, general-purpose data mining tool on the NeoCITIES data to determine which conditions lead to better or worse outcomes. First, expert judgment is used to define the criterion cost that delimits good versus bad outcomes from the simulation. Then we use the J48 algorithm (University of Waikato, undated) implemented in Weka to produce a classifier tree, which is a hierarchical structure starting with one node representing a scenario attribute that provides the greatest discrimination among outcomes. At each successively lower level tree branches either lead to another discriminatory factor or to a leaf node: the leaf nodes in the tree represent a set of outcomes that occur under a set of conditions that are described by the discrimination path traversed through the tree to get to the node. While this particular tree-format will not likely be the format in which results will ultimately be presented to users, it does represent the type of data that will underlay that ultimate presentation.

The results from this type of data mining can be used as part of pre-planning, when the location of fire stations and assignment of fire resources are being considered. We also see a potential use in an Incident Command Center real-time during the emergency event. To be used in real-time, the classifier tree can be translated into a playbook: the decision maker can assess the situation, find those conditions in the playbook and see what is recommend to do to obtain good outcomes.

It is easiest to explain how we are using classifier trees via an example. The classifier tree in Figure 2 shows the attributes in oval-shaped nodes, the different attribute values on the arcs connecting the nodes, and the results (or outcomes) in the rectangular leaf (or end point) nodes. Based on the criterion set by expert judgment, the outcomes are labeled as good, moderate (i.e., satisfactory) or bad. Each leaf node in the tree becomes a part of a "rule" that is derived from the data produced by running the model under different conditions. To illustrate these data-derived rules, we have labeled three of the leaf nodes with Example 1, Example 2 and Example 3. The rules for each of these nodes are as follows:

> Rule for Example 1: when two or fewer trucks are used, the outcome is always bad.
>
> Rule for Example 2: if the initial magnitude of the fire is small (<= 3) and more than two trucks are used from Station A, then the outcome is always moderately good.
>
> Rule for Example 3: if the initial magnitude of the fire is small (<= 3), more than three trucks are sent from Station B, and they arrive in less than seven minutes, then the outcome is always good.

By noting such rules, decision makers could learn that if the fire is small, they should send three or more trucks from Station B if traffic is light (because otherwise a transit time of seven minutes is not possible); otherwise they should send three or more trucks from Station A. This data mining approach thus illuminated an otherwise obscured relationship between traffic congestion and resource allocation.

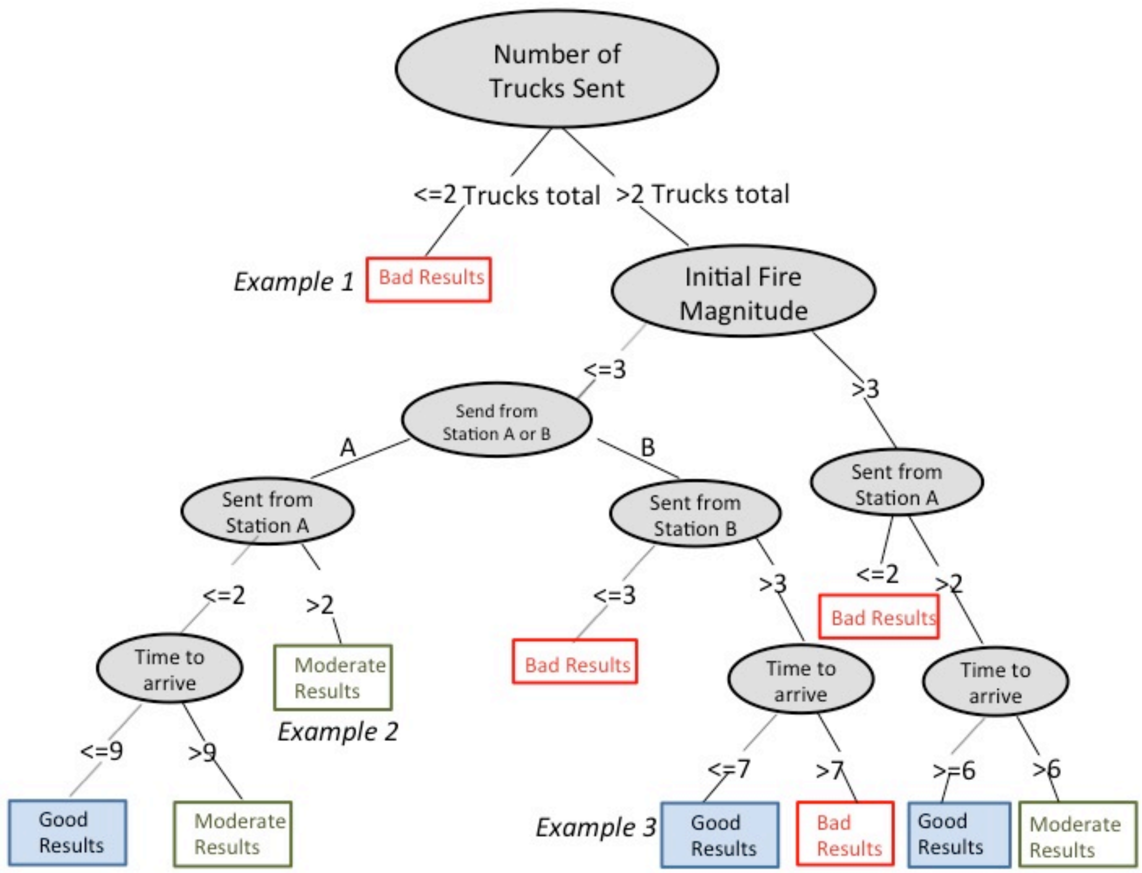

Figure 2. Classifier tree for determining how many fire trucks to send from each station. Three nodes are labeled as examples and are explained further in the text.

Requirements

As a step towards developing an interactive DSS that exploits level-2 and 3 option awareness, we defined requirements that would enable us to verify that the eventual DSS will have the intended functionality and usability. We have composed these requirements in a domain-neutral fashion to make them generically applicable.

Concurrent with requirements definition, we have been developing notional DSS interaction mockups. These informal sketches of user interfaces that incorporate data mining of the decision space serve two purposes. First, they can help to explain the ideas embodied in the requirements. Second, and perhaps more importantly, they serve as a proof-of-concept to show that the requirements can, in principle, be satisfied. The figures in this paper come from the mockups.

At the highest level, the interactive DSS should aid the C^2 operator (who we now call "the user," to be more general) in discriminating among the options to determine the one that has the most robust outcomes.

> 1.0 Overall Goal
> The system shall enable the user to see which options have more robust outcomes.
> - Rationale: to aid understanding decision choices, users will want to find options that will have a majority of good, or at least satisfactory, outcomes.

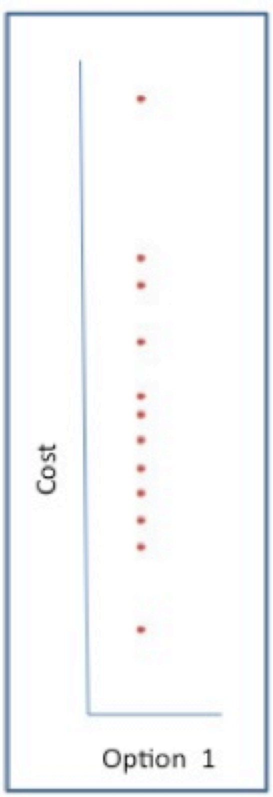

Figure 3. Example outcomes for Option 1. Note that one case has a very high cost.

For example, an option may have a distribution of outcomes, with one outcome being very costly and the rest of the cases having lower costs, as illustrated in Figure 3. If the high-cost case is due to a preventable circumstance, then, after modifying the option to prevent that outcome, all of the remaining outcomes may have satisfactory costs and therefore the option would be much more desirable. Thus there is value in being able to see individual outcomes and their conditions, so that options could be amended to be more robust.

> 1.1 Robustness Support
> Robust options are defined as options that have possible results (outcomes) that are acceptable to users over a broad range of cases developed under many different sets of conditions.
> 1.1.1 The system shall enable the user to identify options that are more or less robust.
> - Rationale: This provides option awareness level 1.
>
> 1.1.2 The system shall enable the user to see options, when they exist, that have the possibility of being made more robust after being modified to mitigate the conditions leading to bad outcomes and/or facilitate conditions leading to good outcomes.
> - Rationale: This constitutes option awareness level 2 and 3. Options that have conditions that lend themselves to being shaped are more attractive.

To see the distribution of results from cases for a particular option and determine whether the option could be improved implies that the DSS needs to support the user in exploring the data. The following are the general requirements related to data exploration.

1.2 Data Exploration Support

For the requirements in this section, “conditions” refer to the variable data parameters that were used for the simulation model run that resulted in a particular outcome. These requirements support option awareness level 2.

1.2.1 The system shall support the user in exploring the data in real-time, such that the user can immediately see the relationship between the change in their selection and the results.

- Rationale: Research shows that performance can be improved by allowing people to make dynamically changing selections and see the results of the changes in real-time (Williamson and Shneiderman, 1992).

1.2.2 The system shall support the user in identifying which conditions yield similar outcomes.

- Rationale: Users want to know what conditions will tend to lead to better or worse outcomes so that they can choose or modify an option with the result that more satisfactory outcomes will occur.

While the requirements above cover the top-level goals of data exploration, more specific requirements are needed to ensure that users have adequate support for exploring the outcomes and the conditions underlying them. We conjecture that decision makers will use one or both of two starting points:

- Start with the options' outcomes and look at the conditions underlying them. The decision maker may want to take this approach to see what conditions may be causing particularly good or bad outcomes. This situation is illustrated in Figure 4, since one option's outcomes are selected.
- Start with the conditions and see what outcomes will result from applying those conditions. This approach may be helpful when a decision maker is reasonably sure that certain conditions will occur, so he or she may want to look at only those outcomes that correspond to the likely conditions. In this way the decision maker can choose the option whose outcomes are all at least satisfactory under the expected conditions. Because the relevant outcomes are on its low end, option 1 in Figure 5 depicts this approach.

Further, decision makers may want to take either of these approaches when looking at one or multiple options.

- The conditions that yield good or bad outcomes may differ between options. Those options whose conditions can be shaped are preferable. Therefore the decision maker will explore good/bad outcomes one option at a time. Figure 4 illustrates this situation.
- Viewing multiple options may be of interest when trying to determine the conditions that lead to similar outcomes across multiple options. See Figure 5.

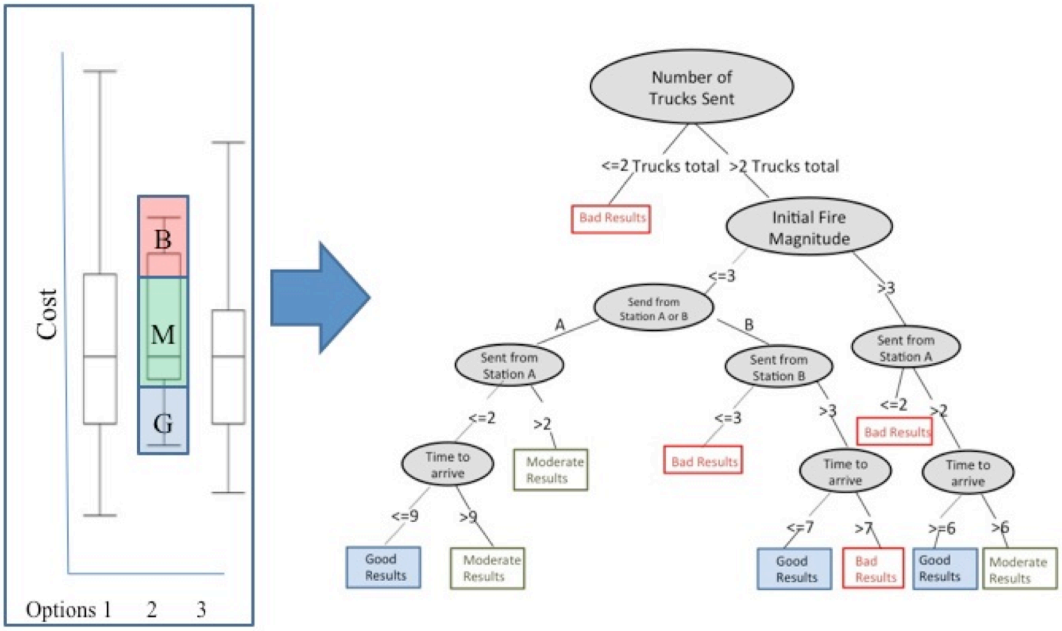

Figure 4. The user identifies on the box plots which cost ranges of outcomes are bad (B), moderate (M), and good (G), and the classifier tree (described above in Figure 2) indicates which conditions generate the outcomes that fall into those ranges. While the cost ranges are highlighted for a single option, it is also possible to highlight a single set of cost ranges across all options and a tree will be generated accordingly.

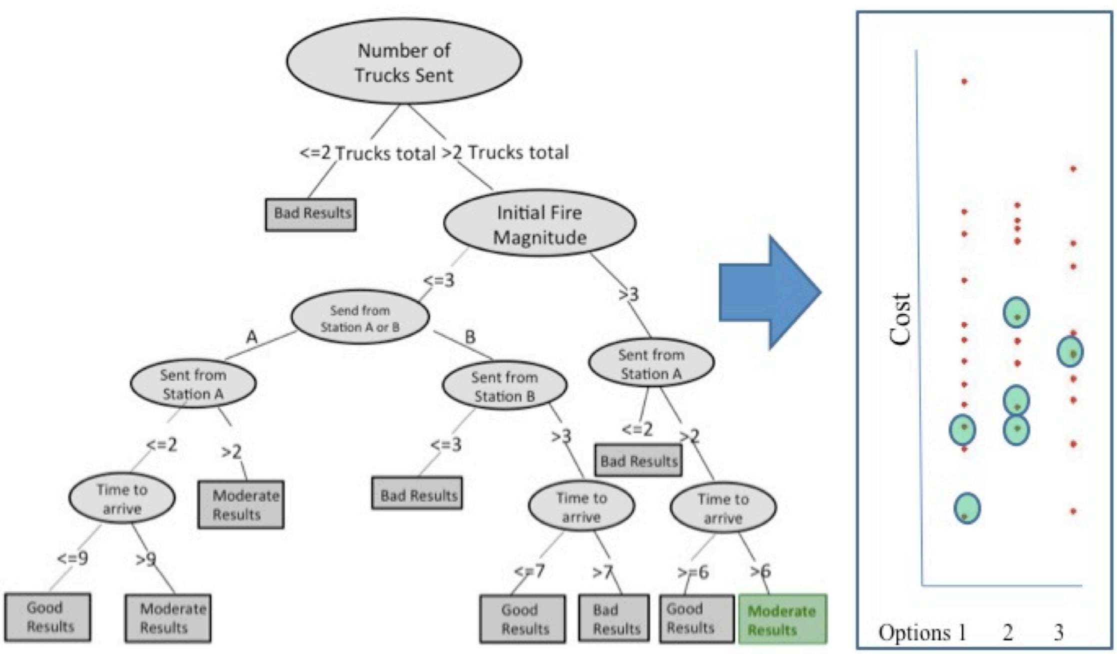

Figure 5. The user identifies a moderate condition in the classifier tree and it interactively highlights the outcomes in green that occur due to that condition. The columns of dots depict the individual outcomes that underlie, and are summarized by, the box plots. While multiple options' outcomes are highlighted, it is also possible to highlight the data for a single option.

The capabilities of starting with outcomes versus starting with conditions, and viewing one versus multiple options, yields a 2x2 matrix, as seen in Table 1.

Table 1. Combinations of Number of Options and Data Exploration Starting Points

	Number of Options Explored	
Starting Point	Outcomes for Single Option → Conditions (see Figure 4)	Outcomes for Multiple Options → Conditions
	Conditions → Outcomes for Single Option	Conditions → Outcomes for Multiple Options (see Figure 5)

Requirements 1.2.3.1 – 1.2.3.4 below are structured to account for all four of the combinations in Table 1. Because the detailed requirements related to data exploration pertain to the capabilities given to users for exploring the data, we express them as graphical user interface (GUI) support requirements.

1.2.3 GUI Support for Data Exploration

1.2.3.1 The system shall enable the user to select the outcomes that are satisfactory, better than satisfactory, and worse than satisfactory for a single option and view the conditions that would lead to those outcomes.

- Rationale: users will want to identify conditions that affect options differently.

1.2.3.2 The system shall enable the user to select the outcomes that are satisfactory, better than satisfactory, and worse than satisfactory across multiple options and view the conditions that would lead to the outcomes for those options.

- Rationale: users will want to find common conditions that work across a number of options, when they exist.

1.2.3.3 The system shall enable the user to select the conditions for a single option, and view the outcomes to which these conditions lead.

- Rationale: users may know that a condition is likely to occur; thus they would like to know the associated outcomes.

1.2.3.4 The system shall enable the user to select conditions across multiple options, and view the outcomes to which these conditions lead.

- Rationale: same as for 1.2.3.3.

Two additional requirements round out the set. One facilitates comparing simultaneously the conditions that pertain to two outcomes, and the other expresses the need for a way to view data in a summary fashion.

1.2.3 GUI Support for Data Exploration, concluded

1.2.3.5 The system shall display, upon user request, the conditions associated with at least two individual (user-specified) outcomes simultaneously.

- Rationale: Users may want to understand why an individual outcome is very bad or very good, and viewing the conditions associated with that outcome can yield that understanding. At least two outcomes should be available for simultaneous viewing for comparison purposes.

1.2.3.6 The system shall provide a visual or textual summary of the outcomes that result from the selected conditions in a manner that is comprehensible to the intended set of users.

- Rationale: Providing a summary of this information will convey the results of the data mining to the users in a way that will enable them to choose satisfactory options or modify options to become more satisfactory.

Discussion and future plans

Although we have used as an example a display of outcomes based on box plots and a display of conditions based on binary classifier trees, alternative visualizations could be substituted for these two decision space depictions. Reflecting this reality, the requirements are not specific to a particular visualization type. The requirements are also not specific to the emergency response C^2 domain, thus increasing their potential utility for use in other C^2 domains. The 2x2 matrix for options and starting points provides a structure to ensure that the requirements cover the important aspects of the user interaction functionality.

We are currently building a DSS in accordance with these requirements that incorporates the data mining and visualization approaches described in this paper. We plan to run human-in-the-loop tests on this DSS to determine the effect of adding this type of level-2 and level-3 option awareness information into the decision space.

We are also beginning to work with a more complex simulation model of state stability and insurgent recruitment. Work with this model and related intelligence analysis will provide us with further insight regarding how such decision space products can be accommodated into an analysis workflow. Additionally, we are working with a group of crisis managers at a major eastern US airport. These crisis management C^2 subject-matter experts are providing us increased insight into how crisis management works in their jurisdiction and are providing feedback regarding how a decision space can be tailored to best support them.

Acknowledgments

This work was supported by The MITRE Corporation's Mission-Oriented Investigation and Experimentation project 0712M810-AA. All product names, trademarks, and registered trademarks are the property of their respective holders.

References

Bankes, S. (1993). Exploratory modeling for policy analysis. *Operations Research,* Vol 41. No. 3 (May – June), 435 – 449.

Belton, V. and Stewart, T. J. (2002). Multiple criteria decision analysis: an integrated approach. Norwell, MA: Kluwer Academic Publishers.

Cantor, J. (2009). *Conquer Cyber Overload. Madison,* WI: CyberOutlook Press.

Chandrasekaran, B. (2005). From optimal to robust COAs: challenges in providing integrated decision support for simulation-based COA planning. Laboratory for AI Research, The Ohio State University.

Campbell, B., Mete, H. O., Furness, T., Weghorst, S., and Zabinsky, Z. (2008). Emergency response planning and training through interactive simulation and visualization with decision support. In *Proc. of the IEEE 2008 International Conference on Technologies for Homeland Security*. Waltham, MA.

Chandresekaran, B. and Goldman, M. (2007). Exploring robustness of plans for simulation-based course of action planning. *Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Multicriteria Decision Making.*

Defense Group, Inc. (undated).CoBRA Software: For the new breed of first responders. http://www.defensegroupinc.com/cobra/, accessed 24 February 2011.

Drury, J. L., Klein, G. L., Pfaff, M., and More, L. (2009). Dynamic decision support for emergency responders. *Proc. of the 2009 IEEE Technologies for Homeland Security Conference,* Waltham, MA.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the 32nd Annual Meeting of the Human Factors and Ergonomics Society*, 97-101.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: a critical review, in M. R. Endsley & D. J. Garland (eds.), *Situation Awareness Analysis and Measurement*, Lawrence Erlbaum Assoc., Mahwah, New Jersey, pp. 3 – 32.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. Available at http://www.kdnuggets.com/gpspubs/aimag-kdd-overview-1996-Fayyad.pdf; retrieved 23 February 2012.

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats, *Psychological Review*, 102 (4), 684-704, American Psychological Association.

Hall, D. L., Hellar, B. and McNeese, M. (2007). Rethinking the data overload problem: Closing the gap between situation assessment and decision making, *Proc. of the 2007 National Symposium on Sensor and Data Fusion (NSSDF) Military Sensing Symposia (MSS)*, McLean, VA.

Hoffrage, U. and Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine*, 73(5), 538 – 540.

Jones, R. E. T. (2006). *The development of an emergency crisis management simulation to assess the impact a fuzzy cognitive map decision-aid has on team cognition and team decision-making.* PhD thesis, College of Information Sciences and Technology, Pennsylvania State University, August 2006.

Keeney, R. and Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value tradeoffs*. Cambridge, UK: Cambridge University Press.

Klein, G. A. (1998). *Sources of Power: How People Make Decisions*. Cambridge, MA: MIT Press.

Klein, G. A., and Brezovic, C. P. (1986). Design engineers and the design process: Decision strategies and human factors literature. *Proceedings of the Human Factors and Ergonomics Society 30th Annual Meeting, 2*, 771-775.

Klein, G. L., Drury, J. L., Pfaff, M. S., & More, L. D. (2010). COAction: Enabling collaborative option awareness. In *Proceedings of the 15th International Command and Control Research and Technology Symposium (ICCRTS)*, Santa Monica, CA.

Kondaveti, R. and Ganz, A. (2009). Decision support system for resource allocation in disaster management. In *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS)*, Minneapolis, MN, September 2009.

Lempert, R. J., Popper, S. W., and Bankes, S. C. (2003). *Shaping the next one hundred years: New methods for quantitative, long-term policy analysis*. RAND MR-1626. Santa Monica, Calif.: The RAND Corporation.

Lipshitz, R., Klein, G. A., Orasanu, J., and Salas, E. (2001). Taking stock of naturalistic decision making. *J. Behav. Dec. Making,* **14**: 331-352.

Liu, Y., Moon, S. P., Pfaff, M. S., Drury, J. L., & Klein, G. L. (2011). Collaborative option awareness for emergency response decision making. In *Proceedings of the 8th Annual International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, Lisbon, Portugal, May 2011.

Phillips, J. K., Klein, G. A., and Sieck, W. R. (2008). Expertise in judgment and decision making: A case for training intuitive decision skills. In D. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 297-315). Malden, MA: Blackwell Publishing Ltd.

Sethi-Iyengar, S., Huberman, G., and Jiang, W. (2004). How much choice is too much? Contributions to 401(k) retirement plans. In O. S. Mitchell & S. Utkus (Eds.), *Pension Design and Structure: New Lessons from Behavioral Finance* (pp. 83-95). Oxford: Oxford University Press.

Thompson, S., Altay, N, Green, W. G. III, and Lapetina, J. (2006). Improving disaster response efforts with decision support systems. *International Journal of Emergency Management*, *3*(4), 250-263.

Tukey, J. W. (1977). *Exploratory Data Analysis.* Reading, Mass: Addison-Wesley.

University of Waikato, New Zealand, Machine Learning Group (undated). Weka. http://www.cs.waikato.ac.nz/ml/weka/. Accessed 24 February 2012.

Wikipedia (2011). Data mining. http://en.wikipedia.org/wiki/Data_mining. Accessed 22 February 2012.

Williamson, C. and Shneiderman, B. (1992). The Dynamic HomeFinder: Evaluating dynamic queries in a real-estate information exploration system. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Copenhagen, Denmark, June 1992.

Witten, I. H., Frank, E., and Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques, 3ED*. Burlington, MA: Morgan Kaufmann.

Zographos, K. G. and Androutsopoulis, K. N. (2008). A decision support system for integrated hazardous materials routing and emergency response decisions. *Transportation Systems Research Part C, Emerging Technologies*, *16*(6), 684-703.

Author Information

Dr. Jill Drury received a BA in Physics from Macalester College in 1980. She received MS degrees in Business Administration in 1986 and Computer Science in 1994, both from Boston University. A Doctor of Science (ScD) degree in Computer Science followed from the University of Massachusetts Lowell in 2002. Her research interests are in optimizing interactive technologies for team-based decision-making in safety-critical applications; particularly for work with real-time operations centers and command and control systems. She is Associate Department Head of the Collaboration and Multi-Media Department of The MITRE Corporation and an Adjunct Assistant Professor at the University of Massachusetts Lowell.

Dr. Gary L. Klein received his BA in Psychology from UCLA in 1974, and his PhD in cognitive social psychology from Texas Tech University in 1982. Dr. Klein's work has focused on modeling how people acquire and use information. He co-developed a collaboration evaluation framework that has been applied to collaboration in intelligence, command and control, and air traffic flow management. Currently, he leads a number of projects on using simulation models to improve decision makers' "option awareness" under deep uncertainty. He is the Senior Principal Scientist in cognitive science and artificial intelligence in the Command & Control Center at The MITRE Corporation.

Scott Musman received a Bachelor's degree in Electronic Engineering from the University of Sussex, U.K., in 1984 and a Masters degree in Computer Science from Johns Hopkins University in 1988. He has worked much of his career automating decision making and developing decision support systems for complex time varying and spatial decision making under uncertainty problems. His has worked on problems associated with automated target recognition, real-time planning, and a number of cyber security related problems. He was Director of R&D at Integrated Management Services Inc., Head of Enterprise Security Research at BAE Systems, AIT, and is now a Principal Engineer at the MITRE Corporation.

Dr. Mark Pfaff received a Bachelor's degree in visual arts from Pennsylvania State University in 1995, a master's degree in multimedia technology from Duquesne University in 2001, and a Ph.D. in 2008 from the College of Information Sciences and Technology at the Pennsylvania State University. Dr. Pfaff is currently an Assistant Professor in the Indiana University School of Informatics (Indianapolis). His research explores the intersections of people, information, and technology in computer-supported cooperative work (CSCW) environments through the use of experimental simulations and mixed-methodological approaches. Mark is formerly an instructor in the Interactive Media department at Duquesne University where he taught multimedia development, Web design, human-computer interaction (HCI), and sound design. At Penn State he managed the User Science and Engineering Lab at IST and was a research assistant for both the Multidisciplinary Initiatives in Naturalistic Decision Systems (MINDS) Group and the Center for Network-Centric Cognition and Information Fusion (NC2IF).

Yikun Liu received a Bachelor's degree in Electronic Engineering (Automation) and a master's degree in Systems Engineering both from Xi'an Jiaotong University of P.R. China. Yikun Liu is currently a Doctoral Candidate in Indiana University School of Informatics (Indianapolis). He works as Research Assistant in the USER LAB with Dr. Mark Pfaff. His previous research activities include computer vision, network security and simulation task environment for decision making. His most recent research interests include awareness display design and work-unrelated interruption behaviors in office working environments.