JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 46, No. 2, March–April 2009

Incorporating Safety Risk in Early System Architecture Trade Studies

Nicolas Dulac* and Nancy Leveson†
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

DOI: 10.2514/1.37361

Ideally, safety should be a part of the early decision making used in conceptual system design. However, effectively evaluating safety risk early enough to inform the early trade studies is not possible with current technology. [We define safety broadly in this paper, as is traditional in the space and defense communities: it is not limited to human death or injury, but also includes equipment and mission losses. Risk is a broader concept than safety, as there are also risks of overrunning budgets and schedules or losing money (in a commercial environment). Risk also implies an analytical analysis whereas safety is a broad term to denote absence of unacceptable losses. To distinguish that we are talking about safety analysis, we use the term "safety risk" in this paper where confusion would result.] This paper presents a new approach to preliminary hazard analysis that can be performed before system design selection and thus can influence key architectural decisions that will be impossible to change later in the system life cycle. The approach is illustrated through a concept evaluation and refinement study for the new NASA space exploration.

Introduction

TRADITIONAL system engineering activities recognize the need for trade studies early in the architecture concept formation phase [1]. Attempts have been made at evaluating some properties of candidate architectures before a system is implemented. The system properties (and associated evaluation techniques) are very different depending on the application domain, problem formulation, requirements, and system development phase.

The field of computer and software architecture, for example, has a rich history of architecture evaluation attempts that dates from the 1970s [2]. Techniques such as the architecture tradeoff analysis method and active reviews for intermediate designs are used to evaluate quality attributes (including performance, availability, security, modifiability, etc.) of software architectures [3,4]. There are many difficulties associated with the use of these evaluation techniques [5,6]. Among others, current evaluation techniques usually require a fair amount of detail before they become effective. The earlier evaluation attempts are made, the more uncertainty in the result. Although the uncertainty may be greater, it should not prevent system architects from attempting early evaluations.

Another problem is that architecture evaluation attempts often focus on the most salient properties of a system, such as cost (for example, function points, constructive cost model [7,8], and others [6] in software engineering), while leaving out other properties such as system safety as a problem to be addressed later in the development life cycle. This is a mistake because many architecture decisions have a significant and lasting impact on safety and may not be reversible after an architecture is selected. For example, the early decision not to add a crew escape system on the space shuttle was based on early architectural decisions and has been impacting shuttle safety for over 20 years [9,10].

Similarly, during the development of large space systems, early trade studies focus on cost (often using mass as a proxy) and performance as the main properties to evaluate potential system architecture and design alternatives. Incorporating safety risk into the decision making at this stage is an important goal: If information about risk were available early, it could be used in the architectural selection process and hazards could be designed out of the system or mitigated early when the cost of doing so is much less than later in the system life cycle. Making basic design changes downstream becomes increasingly costly as development progresses and, often, compromises in safety must be accepted that could have been eliminated if safety had been considered earlier. The problem is that information about the likelihood of particular hazardous events is usually unknown before an architecture and a system design are selected. Although it is relatively easy to identify hazards at system conception, performing a hazard analysis or risk assessment before a design is available is more problematic.

Risk is usually treated as a combination of severity and likelihood. For safety risk, the events considered are the identified hazards. Classic preliminary hazard analysis is performed using a risk matrix [11,12] which provides a combination of these two hazard properties. Although formats can differ slightly, the general form of such a risk matrix is shown in Fig. 1. High-level system hazards are first identified and, for each identified hazard, a qualitative risk evaluation is performed by classifying the hazard according to its severity and likelihood.

Although severity can usually be evaluated using the worst possible consequences of that hazard, the likelihood of the hazard before any system design is performed or even earlier when a system architecture has not yet been selected, is unknown and, arguably, unknowable in most complex space systems. Some probabilistic information is available about physical events, of course, and historical information is theoretically available. Spacecraft designs, however, often include new technology and design features that limit the accuracy of historical information. For example, historical information about the likelihood of propulsion-related losses in previous spacecraft may not be accurate for new designs using nuclear propulsion. In addition, the use of software and digital systems is introducing new ways for hazards to occur that cannot be analyzed using standard hazard analysis techniques that assume accidents are caused by system component failures or using statistical techniques that assume randomness. The difficulty in predicting hazard likelihood is especially great at the very beginning of conceptual studies, where virtually no design information is available. Inaccurate a priori evaluations of hazard likelihood inevitably lead to incorrect risk assessments that can compromise the safety of the system. Discounting the risk associated with potential hazards due to an overoptimistic initial evaluation of likelihood can lead to unnecessary losses.

Received 29 February 2008; revision received 18 August 2008; accepted for publication 18 August 2008.

*Postdoctoral Researcher, Department of Aeronautics and Astronautics, 77 Massachusetts Avenue.

†Professor, Department of Aeronautics and Astronautics and Engineering Systems Division, Room 33-334, 77 Massachusetts Avenue.

LIKELIHOOD	SEVERITY I Catastrophic	II Critical	III Marginal	IV Negligible
A Frequent	I–A	II–A	III–A	IV–A
B Moderate	I–B	II–B	III–B	IV–B
C Occasional	I–C	II–C	III–C	IV–C
D Remote	I–D	II–D	III–D	IV–D
E Unlikely	I–E	II–E	III–E	IV–E
F Impossible	I–F	II–F	III–F	IV–F

Fig. 1 Standard risk matrix.

The analysis described in this paper takes into account the randomness of some events such as micrometeroid strikes, solar flares, and some mechanical failures, but it also recognizes that complex aerospace systems often fail in nonrandom ways. For example, the root causes of the Challenger and Columbia losses included inadequate management decision making and evaluation of the safety risk of identified and documented hazards due to political, economic, and performance pressures [9,13,14].

Likelihood estimation must also account for losses resulting not from component failure but from dysfunctional interactions among components. The loss of the Mars Polar Lander, for example, has been attributed to noise generated when the landing legs were deployed during descent [15]. This noise was normal and expected and did not represent a failure in the landing leg system. The onboard software interpreted these signals as an indication that landing occurred (which the software engineers were told they would indicate) and shut the engines down prematurely, causing the spacecraft to crash into the planet surface. In this loss, and in many other recent spacecraft losses related to software [16], no component "failed"—the landing legs and the software performed correctly (i.e., as specified in their requirements), but the loss occurred due to dysfunctional interaction among these spacecraft components.

As digital components proliferate in spacecraft, this type of component interaction accident will increase. Hazard analyses that assume accidents are caused by random component failures will miss this type of accident, which is typical for systems including digital components. Classic hazard analysis techniques such as fault tree analysis and failure modes and effects analysis do not work well in these types of system interaction accidents not involving component failures and alternatives are needed [17]. This topic is beyond the scope of this paper, however, which focuses on the early system architecture selection process when the system design information needed for these hazard analysis techniques is not available anyway.

We know of no existing rigorous or scientific way to obtain probabilistic or even subjective likelihood information using historical data or analysis in the case of nonrandom failures and system design errors, including unsafe software behavior. When forced to come up with such evaluations, engineering judgment is usually used, which in most cases amounts to pulling numbers out of the air. Selection of a system architecture on such a basis is questionable and perhaps one reason why risk is usually not used in the early architectural trade process.

In this paper, we propose a new way of performing likelihood analysis as part of a standard preliminary hazard analysis that can be started at the beginning of the system life cycle, before architecture selection, and used to inform the early architecture trade studies. Later, after an architecture is selected, the information generated in the analyses can be used to design hazards out of the system during the detailed design process as the original analyses are revised and refined. In this paper, we cover only the incorporation of risk into the architectural design selection and trade studies. Safety-driven design is described elsewhere [18,19].

The new analysis technique uses the hazard mitigation potential of multiple candidate architectures to estimate hazard likelihood. Hazards that are more easily mitigated in the design and operations are less likely to lead to accidents, and similarly, hazards that have been eliminated during system design simply cannot lead to an accident. Thus the goal of the analysis process described in this paper is to assist in selecting an architecture with few serious hazards and inherently high mitigation potential for those hazards that cannot be eliminated, perhaps because eliminating them would reduce the potential for achieving other important system goals.

We chose mitigation potential as a surrogate for likelihood for two reasons: 1) the potential for eliminating or controlling the hazard in the design has a direct and important bearing on the likelihood of the hazard occurring (whether traditional or new designs and technology is used), and 2) mitigatibility of the hazard can be determined before an architecture or design is selected—indeed, it helps in the design selection process.

We acknowledge up front the difficulty of providing an evaluation of our approach. Waiting until a complex space system using this approach has been built and operated for a reasonable amount of time could take decades and is impractical. A carefully controlled experiment is not feasible. Comparing the results obtained to alternatives in an informal way would be possible if there were alternatives. The only alternative we know that has been suggested is simply expert judgment, which is actually a part of our approach (but augmented and guided) and thus is not independent of it. We start with expert judgment and add information and analysis. Therefore, we are left only with an argument based on our experience in performing or reviewing dozens of preliminary hazard analyses in a variety of systems and industries. We hope that the proposal in this paper will spark more interest in coming up with alternatives that could later be compared and evaluated.

The new process is demonstrated using a Massachusetts Institute of Technology (MIT)/Draper Labs project to perform an early concept evaluation and refinement for the NASA space exploration mission. The goal was to develop a space exploration architecture that fulfills the needs of the many stakeholders involved in the exploration enterprise. Safety was defined up front as one of the most critical criteria for a successful space exploration enterprise. Because safety is an important property to many of the stakeholders, using it to influence early architectural decisions was important as most of these architectural decisions would be very costly or impossible to change later in the development process. Although the new hazard analysis methodology was used for the entire space exploration architecture, this paper for practical reasons includes only the transportation subset, bringing humans from Earth orbit to the moon/Mars surface and returning them back to Earth orbit. Earth launch and reentry, as well as moon/Mars surface operations are omitted but can be found in the NASA final reports. Note that we use this project only as an example of the new safety/risk assessment procedure and do not evaluate their overall technology and approach to architecture generation in this paper.

The next section briefly describes important elements of the space exploration example used throughout the paper. In the rest of the paper, the methodology is described and sample results presented.

Space Exploration Example

The new U.S. space exploration vision involves a return to the moon as a stepping stone for the future human exploration of Mars. During the concept evaluation and refinement performed jointly by MIT and Draper Labs, 1162 possible Earth-to-moon/Mars-and-back transportation architectures were generated. The architectures were generated by selecting transportation vehicles and functions based on a set of combination rules and constraints (see [20] for a description of the object-based architecture generation framework used). Although the risk analysis procedure described in this paper could have been used up front as one of the initial architecture filtering criteria, it was decided to initially filter out highly inefficient

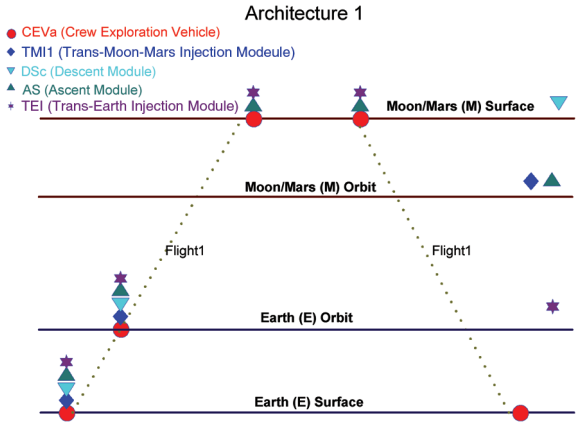


Fig. 2 Sample one-flight transportation architecture.

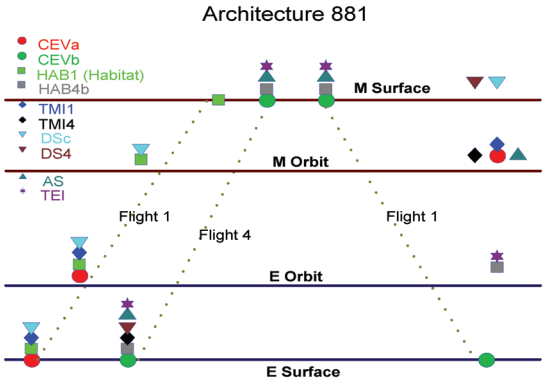


Fig. 3 Sample two-flights transportation architecture.

architectures from a mass and feasibility perspective and then perform a risk evaluation on the remaining architectures. Because of the large number of architectures considered in the architectural generation approach used, this choice seems reasonable but the risk analysis approach proposed in this paper applies either way.

In this context, an architecture can be defined as the combination of a transportation architecture with a list of parameters and options related to technology utilization, policy, and operations. (The completeness and appropriateness of the specific options considered in the MIT/Draper Labs project is not relevant to the safety/risk analysis approach being demonstrated in this paper; any selection of parameters and options would still work.) A transportation architecture includes the following:

1) The number and type of vehicles and modules used to send humans and cargo to the moon/Mars surface and return them to Earth.

2) The role and activities for each vehicle/module, including:

a) dockings and undockings,

b) trajectories and orbit insertions,

c) assembly of vehicle/modules stacks,

d) discarding of vehicles/modules,

e) prepositioning of vehicles/modules in orbit and on the planet surface.

A sample transportation architecture is shown in Fig. 2. In this simple architecture, a single flight (flight 1) is used to transport crew and cargo from the Earth (E) to the moon (M) surface and back. Flight 1 includes a crew exploration vehicle (CEVa), a trans-moon–Mars injection (TMI) module, surface descent (DSc), and ascent modules (AS), and a trans-Earth injection (TEI) module for the return. Modules on the right of Fig. 2 are discarded at various stages of the mission. For example, the surface descent module (DSc) is left on the moon's surface.

Figure 3 shows a more complex architecture where two outbound flights are required. Flight 4 is used to preposition cargo and assets such as a surface habitat (HAB4b), an ascent propulsion module (AS), and a return crew exploration vehicle (CEVb) to the M surface. Once asset prepositioning is complete, flight 1 brings the crew to the surface using an outbound CEVa and transfer habitat (HAB1).

In addition to a transportation architecture, a complete exploration architecture includes a set of parameters related to areas such as technology, propulsion, policy, and operations. Tables 1–3 provide a list of some parameters and options used in the architecture definition and associated safety analysis. The total architectural space can be theoretically obtained by taking the cross product of all the available architectural options, including the transportation architecture and additional options.

Hazard-Based Safety/Risk Analysis Methodology

The hazard-based safety/risk analysis developed is a three-step process as follows:

1) Identify the system-level hazards and associated severities.

2) Identify mitigation strategies and associated impact.

3) Calculate safety/risk metrics for a given transportation architecture.

The first two steps are performed only once, at the beginning of the process. They may have to be repeated if the architectural design space changes or if additional hazards are identified. The third step is repeated to evaluate as many transportation architectures and variations as necessary. The following sections discuss each of the three steps in more detail.

Step 1: Identify System-Level Hazards and Severities. Just as in typical hazard analyses, the first step in the methodology is to identify the system-level hazards. A hazard identification worksheet can be used to streamline this process and to ensure proper tracking of the hazards. Figure 4 shows the worksheet used for the space exploration example. This worksheet includes the standard information usually included in a hazard log during a preliminary hazard analysis [21] but it augments it with the mitigation-related information needed for the new analysis approach.

As with any preliminary hazard analysis activities, identifying system-level hazards involves 10% creativity and 90% experience.

Table 1 Technology options used in the exploration architecture definition

Technology options:	Option 1	Option 2
ISRU (in situ resource utilization)	No	Yes
Aerocapture	No	Yes
Nuclear thermal rockets	No	Yes
Solar electric propulsion (for cargo)	No	Yes
Nuclear electric propulsion	No	Yes
Nuclear surface power	No	Yes
Level of autonomy	Low	High
Highly elliptical orbital rendezvous	No	Yes
Rendezvous in transit	No	Yes
Artificial gravity	No	Yes
High-closure environmental control and life-support system (ECLSS) (water, oxygen)	No	Yes
Low boil-off propellant storage	No	Yes
In-space propellant transfer	No	Yes

Table 2 Propulsion options used in the exploration architecture definition

Propulsion options:	Option 1	Option 2	Option 3	Option 4	Option 5
Transfer to M	Hydrogen (H2)/liquid oxygen (LOX)	Methane (CH4)/LOX	Hypergolic	Nuclear	Electric
Arrival to M	H2/LOX	CH4/LOX	Hypergolic	Nuclear	——
Descent and ascent	H2/LOX	CH4/LOX	Hypergolic	Nuclear	——
Return to Earth	H2/LOX	CH4/LOX	Hypergolic	Nuclear	Electric

Table 3 Policy and operational options used in the exploration architecture definition

Policy/operational options:	Option 1	Option 2	Option 3
Heavy lift launch vehicle (HLLV)	No	Yes	
Crew size	0	1	2+
Habitable modules during TMI	——	1	2+
Habitable modules on surface	——	1	2+
Human/cargo transfer	Separate	Coupled	
Nuclear	No	Yes	
Deinvesting in the moon	No	Yes	
Level of international involvement	Low	High	
Level of commercial involvement	Low	High	
Free-return trajectory	No	Yes	
Initial Mars mission duration	Short	Long	
Level of abort options	Low	Medium	High
Mars landing sites	Single	Multi	Chain
Surface elements reusability	No	Yes	
Transportation elements reusability	No	Yes	

Consequently, domain experts should be closely involved in identifying hazards for each mission phase.

Once the hazards are identified, the severity of each hazard is evaluated by considering the worst-case loss associated with the hazard. In the example, the losses were evaluated for each of three categories: humans (H), mission (M), and equipment (E). Initially, potential damage to the Earth and planet surface environment was included in the hazard log. In the end, the environment component was left out of the analysis because project managers decided to replace it with mandatory compliance with NASA's planetary protection standards. A more comprehensive analysis should include the inherent potential of a system to mitigate environmental hazards. The methodology presented in this paper can be easily extended to do so. A custom severity scale (see Table 4) was defined to account for the losses associated with each category. Some hazards identified, such as fire, explosion, or loss of life support, span multiple (if not all) mission phases. These hazards were grouped under the label "general" hazards to simplify the analysis. However, the mitigation strategies associated with these hazards depend on the mission phase to which they apply.

Table 5 shows a summary of the identified hazards and severities, organized by the mission phase. Again, we relied on the judgment of the experts involved in the project to assess severity as is common practice in preliminary hazard analysis. Assessment of severity is not usually difficult or controversial.

Step 2: Identify Mitigation Strategies and Associated Impact. The second step of the methodology involves the identification and assessment of possible mitigation strategies for each hazard. The architectural space has a tendency to change very rapidly at the beginning of a project when different options are being explored at a fast pace. Fortunately, the methodology is highly flexible and allows for rapid reevaluations of architectures when changes occur. The key to this analysis is to determine the impact of each architectural option on each system-level hazard. To do so, a four level hazard mitigation

Hazard Name: Nuclear reactor overheating

Mission Phase: (circle all appropriate) Prelaunch | To-Space Launch | In-Space Assembly | To M Transfer | To M Descent | **(Surface Exploring)** | From M Ascent | To E Transfer | In E Orbit Arriving | On E Landing | On E Recovery

Operation/Event: Ex: docking, lift-off, etc.
Power generation for surface exploration activites

Vehicle(s)/ System(s) Affected: Ex: CEV, DAV, rover, etc.
Surface nuclear power generator, and all systems used on M surface. (HAB, DAV, rover(s), powered equipment)

Subsystem(s) Affected: Ex: engine, heat shield, etc.
Nuclear reactor, cooling subsystem,...

Severity (1-4):**

Human	Mission	Equipment	Environm.
4	4	3	1

Accident/Effect Description: What potential losses could result form the hazard occurence? What are the worst potential effects, assuming no mitigation stgrategies are implemented? What damage could result? Explain the severity ratings provided above.
Nuclear reactor core meltdown would cause loss of power, and possibly radiation exposure. Surface operations must abort mission and evacuate. If abort is unsuccessful or unavailable at the time, the crew could be lost. All surface equipment is lost. No environmental impact on Earth.

Hazard Description: Describe the hazard as a system state? What other environmental conditions could influence the effect of the hazard occurence?
Nuclear reactor operating at temperature above design limits.

Causal Factors / Assumptions What conditions allowed the hazard to occurr? Why was the system allowed to get into the hazardous state?
TBD. Possible causes include: thermal control system malfunction, solar radiation protection inadequate, insufficient radiator heat rejection,...

Mitigation Strategy:

	Cost/ Difficulty (L,M,H)	Mitigation Priority (1-4)
1. Surface power generation does not rely on nuclear technology	M	4
2. Back-up power generation system is available for surface operations	H	1

Fig. 4 Hazard identification worksheet.

Table 4 Custom hazard severity scale

Severity level	General description		
	Human	Mission	Equipment
4	Loss of life	Mission abort or mission loss	System loss
3	Severe injury or illness	Major mission objectives incomplete	Major system damage
2	Minor injury or illness	Minor mission objectives incomplete	Minor system damage
1	Less than minor injury or illness	All mission objectives completed	Less than minor system damage

Table 5 System-level hazards and associated severities

ID no.	Phase	Hazard	Severity		
			H	M	E
G1	General	Flamable substance in presence of ignition source (fire)	4	4	4
G2	General	Flamable substance in presence of ignition source in confined space (explosion)	4	4	4
G3	General	Loss of life support (includes power, temperature. oxygen, air pressure. CO2, food, water, etc.)	4	4	4
G4	General	Crew injury or illness	4	4	1
G5	General	Solar or nuclear radiation exceeding safe levels	3	3	2
G6	General	Collision (micrometeroids, debris, with modules during rendevous or separation maneuver, etc.)	4	4	4
G7	General	Loss of attitude control	4	4	4
G8	General	Engines do not ignite	4	4	2
PL1	Prelaunch	Damage to payload	2	3	3
PL2	Prelaunch	Launch delay (due to weather, prelaunch test failures, etc.)	1	4	1
LI	Launch	Incorrect propulsion/trajectory/control during ascent	4	4	4
L2	Launch	Loss of structural integrity (due to aerodynamic loads, vibrations, etc.)	4	4	4
L3	Launch	Incorrect stage separation	4	4	4
E1	EVA in space	Lost in space	4	4	1
A1	Assembly	Incorrect propulsion/control during rendezvous	4	4	4
A2	Assembly	Inability to dock	1	4	3
A3	Assembly	Inability to achieve airlock during docking	1	4	3
A4	Assembly	Inability to undock	4	4	3
T1	In-space transfer	Incorrect propulsion/trajectory/control during course change burn	4	4	3
D1	Descent	Inability to undock	4	4	3
D2	Descent	Incorrect propulsion/trajectory/control during descent	4	4	4
D3	Descent	Loss of structural integrity (due to inadequate thermal control, aerodynamic loads, vibrations, etc.)	4	4	4
A1	Ascent	Incorrect stage separation (including ascent module disconnecting from descent stage)	4	3	3
A2	Ascent	Incorrect propulsion/trajectory/control during ascent	4	3	3
A3	Ascent	Loss of structural integrity (due to aerodynamic loads, vibrations, etc.)	4	3	3
S1	Surface operations	Crew members stranded on M surface during extra-vehicular activity (EVA)	4	3	3
S2	Surface operations	Crew members lost on M surface during EVA	4	3	3
S3	Surface operations	Equipment damage (including related to lunar dust)	2	3	3
NP1	Nuclear power	Nuclear fuel released on Earth surface	4	4	2
NP2	Nuclear power	Insufficient power generation (reactor does not work)	4	3	3
NP3	Nuclear power	Insufficient reactor cooling (leading to reactor meltdown)	4	3	3
RE1	Reentry	Inability to undock	4	3	3
RE2	Reentry	Incorrect propulsion/trajectory/control during descent	4	3	3
RE3	Reentry	Loss of structural integrity (due to inadequate thermal control, aerodynamic loads, vibrations, etc.)	4	3	4
RE4	Reentry	Inclement weather	4	2	2

impact scale is used (Table 6). This scale is based on typical system safety hazard mitigation priority scales [21] and is used to determine the impact (if any) of a given architecture option on a given hazard.

In the space exploration example, a database of mitigation impact was generated with the help of domain experts and was recorded in a spreadsheet. It contains the mitigation impact (1–4) of each architectural option, for each hazard, for each category (human, mission, or equipment).

Figure 5 shows only a small part of the database created to record hazard mitigation information and evaluate architectures. The system-level hazards and their associated severities (human, mission, and equipment) are listed in the top rows. The architectural and technology options are listed in the column on the left. The architectural space is divided into parameters with alternatives, for example, the parameter of launch vehicle type includes a binary alternative for the use of a heavy launch vehicle (HLLV). Many of the available architectural options are hidden to make the table more readable (e.g., chemical propellant options).

The effects of architectural parameters and technology options on each hazard are recorded in the database according to the 1–4

Table 6 Hazard mitigation scale and priority

Mitigation impact Level	General description	Detailed description
4	Eliminate	Complete elimination of the hazard from the design
3	Prevent	Reduction of the likelihood that the hazard will occur
2	Control	Reduction of the likelihood that the hazard results in an accident
1	Reduce damage	Reduction of damage to the system if an accident does occur

Hazard ID -->		G1			G2			G3			G4			G5			G6			G8			G9		
Hazard Name -->		Fire			Explosion			Loss of Life Support			Crew Injury or Illness			Collision			Loss of Structural Integrity			Loss of Attitude Control			Incorrect Propulsion / Control		
Design/Architecture Parameter	**1**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**3**	**1**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**4**
ISRU - Yes	1	1	1	1	1	1	1	2	2	2															
ISRU - No									3	3															
Aerocapture - Yes	1	1	1	1	1	1	1																		
Aerocapture - No																	3	3	3						
Nuclear Thermal Rockets - Yes		1	1	1	1	1	1																		
Nuclear Thermal Rockets - No	1																								
Solar Electric Propulsion - Yes		1	1	1	1	1	1																		
Solar Electric Propulsion - No	1																						3	3	
Nuclear Electric Propulsion - Yes		1	1	1	1	1	1																		
Nuclear Electric Propulsion - No	1																						3	3	
Rendevous in transit - Yes																									
Rendevous in transit - No	1													3	3	3				3	3	3	3	3	3
Artificial gravity - Yes											3	3													
Artificial gravity - No	1																								
High-closure ECLSS (H2O, O2) - Yes																									
High-closure ECLSS (H2O, O2) - No	1							3	3	3															
In-space propellant transfer - Yes																									
In-space propellant transfer - No	1	3	3	3	3	3	3																		
HLLV - Yes	1																3	3	3						
HLLV - No																									
Nuclear - Yes																									
Nuclear - No	1																								
Free-return trajectory - Yes	1	3			3			3			3												2		
Free-return trajectory - No																									
Initial Mars mission duration - Long	1																								
Initial Mars mission duration - Short								2	2		2														
Level of abort options - High								3	3		1	1	1												
Level of abort options - Moderate	1																								
Level of abort options - Low																									
Crew size - 0		4			4			4			4			4			4			4			4		
Crew size - 1+	1																								

Fig. 5 Sample hazard mitigation database.

mitigation scale. For example, not performing rendezvous in transit and/or highly elliptical orbital rendezvous reduces the likelihood of a collision between modules and vehicles (mitigation level 3). Similarly, the use of an unmanned architecture (crew size = 0) completely removes the potential for human loss (mitigation level 4), but does not directly impact potential mission or equipment losses. As in any hazard analysis process, documenting the rationale and assumptions for each hazard, mitigation strategy, and impact is critical. For complex system architectures, the mitigation database can be very large, which makes it impossible for analysts to remember the inputs of every domain expert. Consequently, database updates and changes can be very difficult unless the mitigation strategies and impact were carefully documented and linked to source material. Once the hazard mitigation database has been populated, it is possible to start evaluating the overall mitigation potential of various exploration architectures.

Step 3: Evaluate Architectures and Calculate Safety/Risk Metrics. As previously mentioned, a complete exploration architecture is defined as the union of the transportation architectures with a set of technology/policy parameter options. To evaluate the risk associated with a specific architecture, an architecture vector is created that includes all of the parameters for that architecture. This vector is in the form of a large string of binary numbers. A sample architecture vector can be found in the second column of Fig. 5. The "1" values in the vector indicate that the corresponding option is selected. In this example, the vector shows that the selected architecture includes the following: 1) in situ resource utilization (ISRU), 2) aerocapture, 3) no nuclear thermal rockets, 4) no solar electric propulsion, 5) no nuclear electric propulsion, 6) no rendezvous in transit, 7) etc.

An architecture vector has to be created for each architecture evaluation. The architecture vector generation process can easily be automated if a large number of evaluations have to be performed. Once an architecture vector has been defined, the risk evaluation and metrics computation proceeds as follows:

1) For each hazard and each hazard category (human, mission, or equipment), that is, for each column of the spreadsheet, the algorithm scans for the option that provides the maximum hazard mitigation for each architectural parameter. These *maximum mitigation factors* are added to obtain the sum-total maximum hazard mitigation factor for each hazard and each category (H, M, E). The maximum mitigation factors and *their sum total* are architecture independent. They only depend on the architectural option space and the hazard mitigations identified. The process of searching for maximum mitigation factors should be automated to provide the flexibility necessary to modify the architectural space and to make the tool evolvable. The maximum mitigation factors obtained in this step are almost never achievable in real life because of potential impracticality or cost or because that would require unacceptable tradeoffs with other system requirements and constraints such as mass, performance, or development schedule. For example, selecting an unmanned mission (crew size = 0) reduces human risk considerably, but it directly conflicts with the essence of the space exploration vision. Nevertheless, the maximum hazard mitigation factor is important because it provides an architecture-independent theoretical absolute upon which all the other architectures can be compared.

2) To evaluate a specific architecture, the algorithm matches the selected options with their respective hazard mitigation impact and computes a sum of the mitigation factors obtained for the options selected in the architecture under evaluation. This process is repeated for each hazard and category (H, M, E). The result is a set of *hazard mitigation indices* obtained for a particular architecture.

3) A *relative residual risk index* is calculated for each hazard (h) and each category (c) using the following formula:

$$\text{relative residual risk index } (h,c) = (1 - (\text{hazard mitigation index } (h,c)/\text{maximum mitigation factor } (h,c)))$$

4) If a hazard and/or category is completely eliminated (mitigation level = 4) by a selected architectural option, the relative residual risk index for this hazard is automatically set at zero.

5) A postmitigation *relative severity index* for each hazard and category is then calculated as follows:

$$\text{relative severity index } (h,c) = (\text{relative residual risk index } (h,c)) * \text{original hazard severity } (h)^2$$

The "squared" severity is used to provide heavier weighting on the higher severity indices (see the weighting factors used in the MIT/Draper project). Weighting factors are system-dependent and have to be discussed with analysts and project managers because they have a significant impact on the final analysis results. Relative severity rating is always subjective and a decision that is usually made at the project management or organizational level. We chose a weighting factor that seemed appropriate for this particular project and was acceptable to the engineers participating, but other weightings are possible and easily implemented.

6) Three *relative risk metrics* (human, mission, or equipment) are obtained by averaging the relative severity indices for each category (H, M, E) across all hazards.

7) As needed, an *overall residual safety/risk metric* (ORSRM) for an architecture can be obtained using a weighted average of the relative severity indices for each category (H, M, E) with custom weighted factors (H, M, E) selected by the project team. In this project, the weighting factors selected were as follows: human: 9, mission: 3, equipment: 1. Again, different weighting factors can be used, depending on the judgment and goals of the particular project or organization (see Fig. 6).

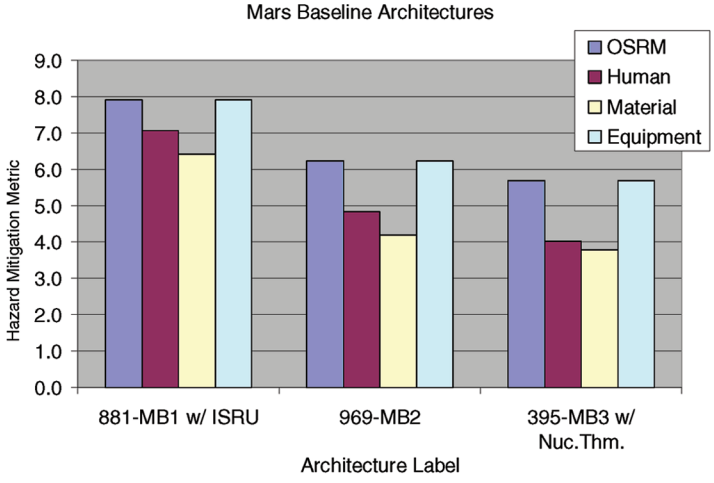
The hazard mitigation metrics are used to evaluate and rank potential transportation architectures. By systematically selecting and deselecting options in the architecture description, it is possible to perform a first-order assessment of the relative importance of each architectural option in determining the overall residual safety/risk metric.

Sample Results

The results from the analysis provide a ranking of the selected transportation architectures based on the hazard mitigation potential of each. Hundreds of parameters are considered in the safety/risk analysis, but the analysis allowed the identification of major contributors to the hazard mitigation potential of selected architectures. These contributors include the use of heavy module and equipment prepositioning on the surface of Mars and the use of minimal rendezvous and docking maneuvers. Prepositioning modules allows for pretesting and mitigates the hazards associated

		Human	Mission	Equipment
Hazard Severity	Weight	9	3	1
4.Catastrophic	16			
3.Critical	9			
2.Major	4			
1.Marginal	1			

Fig. 6 Weighting factors used in calculating ORSRM.



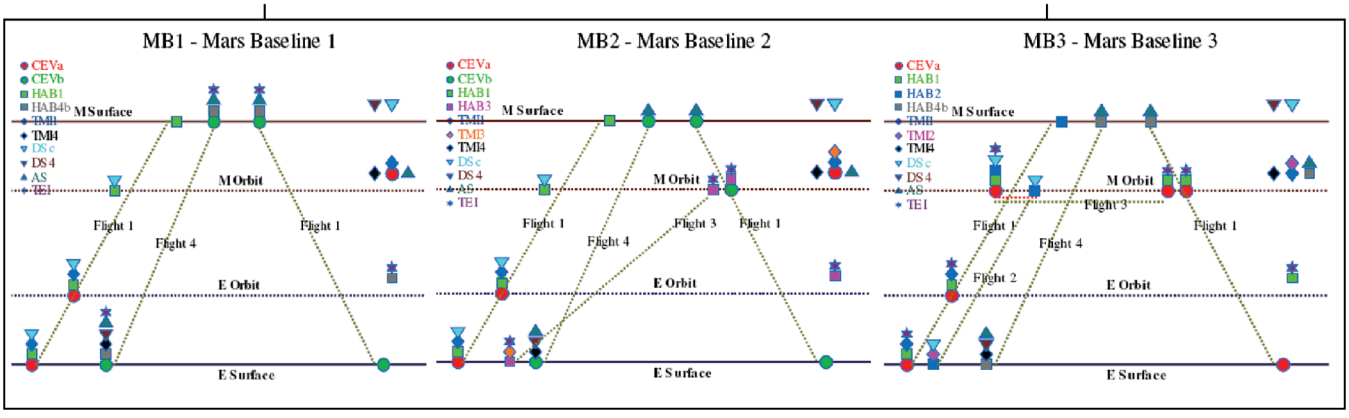


Fig. 7 Evaluation results for three Mars baseline architectures.

with loss of life support, equipment damage, and so on. On the other hand, prepositioning modules increases the reliance on precision landing to ensure that all landed modules are within range of each other. Consequently, using heavy prepositioning may require additional mitigation strategies and technology development to reduce the risk associated with landing in the wrong location. As another example, a transportation architecture requiring no docking at Mars orbit or upon return to Earth inherently mitigates hazards associated with collisions or failed rendezvous and docking maneuvers. On the other hand, having the capability to dock during an emergency, even though it is not required during nominal operation, provides additional mitigation to loss of life support, especially in Earth orbit.

Hundreds of architectures were evaluated for their inherent hazard mitigation potential. After multiple down selections, three very different Mars baseline architectures (MB1, MB2, MB3) were selected as final candidates (see Fig. 7). In this analysis, MB1 was evaluated with the use of ISRU, because a direct return from the Martian surface is not possible using the MB1 transportation architecture unless propellant is extracted from local resources. MB3 was evaluated with nuclear thermal rocket propulsion, which shortens the transit time but requires additional mitigation for hazards associated with the use of nuclear power (radiation, contamination, overheating, inadequate propulsion/control, high structural loads, etc.).

An automated tool was created to perform multiple evaluations based on the needs of the team responsible for designing the transportation architecture. As mentioned previously, hundreds of different architecture evaluations have been performed with little maintenance and data input efforts. The analysis started at the very beginning of the conceptual design phase and the methodology proved flexible and extensible enough to carry the team from day 1 of conceptual design up to the beginning of the detailed design phase, at which point, a more detailed hazard analysis methodology such as systems-theoretic accident modeling and processes (STPA) [17,18] will be necessary and safety-driven design of the system and its components can be started [18,19].

Conclusions

The methodology described in this paper was developed to assist in performing structured preliminary hazard analysis during early system architecture trade studies. Such a methodology allows considering the inherent hazard mitigation potential of candidate system architectures early in system development when safety can be increased relatively cheaply without the need for costly downstream development changes or compromises. As shown in the example, the methodology handles highly complex, broad-scoped, multivehicle, time-dependent systems and is flexible, extensible, and adaptable to other types of complex systems.

The example in this paper focused on describing the methodology using a very structured transportation architecture generation scheme. However, the methodology was applied with equal success to perform a structured preliminary hazard analysis of the surface operations mission architecture of the same project. The information available for the surface architecture evaluation had very different format, content, and context, but the hazard mitigation analysis was similar. The application of the methodology in the early system architecture trade studies for this project demonstrated that a highly structured preliminary hazard analysis process using mitigation potential as an estimator of likelihood in a safety/risk evaluation is feasible and practical in a project of this size and complexity.

As discussed in the Introduction, the accuracy of the results cannot be directly evaluated and rests on the assumption that being able to mitigate hazards affects the likelihood of their occurrence. There is no way to scientifically prove this assumption at this time, but it seems a reasonable assumption for spacecraft designers to make and, indeed, underlies all attempts to increase safety through engineering design.

Acknowledgments

This work was partially supported by a NASA Concept Exploration and Refinement (CE&R) contract to the Massachusetts Institute of Technology (MIT) and Draper Laboratories and by a grant (NAG2-1543) from the NASA Engineering for Complex Systems Program. We would like to thank all the Draper Laboratories engineers and MIT faculty and students who contributed to the risk analysis described in this paper.

References

[1] Blanchard, B. S., and Fabrycky, W. J., *Systems Engineering and Analysis*, 3rd ed., Prentice–Hall, Englewood Cliffs, NJ, 1998.

[2] White, J. R., and Booth, T. L., "Towards an Engineering Approach to Software Design," *Proceedings of the Second International Conference on Software Engineering (ICSE)*, IEEE Computer Society, San Francisco, CA, 1976, pp. 214–222.

[3] Barbacci, M., Carriere, J., Kazman, R., Klein, M., Lipson, H., Longstaff, T., and Weinstock, C., "Steps in an Architecture Tradeoff Analysis Method: Quality Attribute Models and Analysis," Software Engineering Institute, CMU, TR CMU/SEI-97-TR-029, May 1998.

[4] Clements, P., Kazman, R., and Klein, M., *Evaluating Software Architectures: Methods and Case Studies*, Addison–Wesley Publishers, Boston, MA, 2002.

[5] Chatman, V. V., "Change Points: A Proposal for Software Productivity Measurement," *Journal of Systems and Software*, Vol. 31, No. 1, 1995, pp. 71–91.
doi:10.1016/0164-1212(94)00088-5

[6] Kemerer, C. F., "An Empirical Validation of Software Cost Estimation Model," *Communications of the ACM*, Vol. 30, No. 5, 1987, pp. 416–429.
doi:10.1145/22899.22906

[7] Boehm, B. W., *Software Engineering Economics*, Prentice–Hall PTR, Englewood Cliffs, NJ, 1981.

[8] Boehm, B. W., Abts, C., Brown, A. W., Chulani, S., Clark, B. K., Horowitz, E., Madachy, R., Reifer, D. J., and Steece, B., *Software Cost Estimation with COCOMO II*, Prentice–Hall PTR, Englewood Cliffs, NJ, 2000.

[9] Gehman, H., (Chair), *Columbia Accident Investigation Report*, NASA, Washington, D.C., 2003.

[10] McCurdy, H., *Inside NASA: High Technology and Organizational Change in the U.S. Space Program*, Johns Hopkins Univ. Press, 1994.

[11] NASA General Safety Program Requirements, NASA NPR 8715.3C, March 2008.

[12] Standard Practice for System Safety, Department of Defense MIL-STD-882D, 10 Feb. 2000.

[13] Rogers, W. R., (Chair), *Report of the Presidential Commission on the Space Shuttle Challenger Accident*, U.S. Government Accounting Office, Washington, D.C., 1986.

[14] Leveson, N. G., "Technical and Managerial Factors in the NASA Challenger and Columbia Losses: Looking Forward to the Future," *Controversies in Science and Technology, Vol. 2: From Chromosomes to the Cosmos*, edited by D. L. Kleinman, K. A. Cloud-Hansen, C. Matta, and J. Handelsman, Mary Ann Liebert, Inc., New Rochelle, NY, 2007, pp. 237–261.

[15] JPL Special Review Board, Report on the Loss of the Mars Polar Lander and Deep Space 2 Missions, NASA Jet Propulsion Laboratory, 22 March 2000.

[16] Leveson, N. G., "The Role of Software in Spacecraft Accidents," *Journal of Spacecraft and Rockets*, Vol. 41, No. 4, July–Aug. 2004, pp. 564–575.
doi:10.2514/1.11950

[17] Leveson, N. G., *System Safety Engineering: Back to the Future*, MIT Press, Cambridge, MA (to be published) (draft available at http://sunnyday.mit.edu/book2.pdf).

[18] Owens, B., Herring, M., Leveson, N. G., Ingham, M., and Weiss, K. A., "Application of a Safety-Driven Design Methodology to an Outer Planet Exploration Mission," *IEEE Aerospace Conference*, IEEE, Piscataway, NJ, March 2008, pp. 174–193.

[19] Leveson, N. G., and Dulac, N., "Safety and Risk Driven Design in Complex Systems of Systems," *NASA/AIAA First NASA Space Exploration Conference*, AIAA, Reston, VA, 2005, pp. 57–74; also AIAA Paper 2005-2558.

[20] Koo, H. B., "A Meta-Language for System Architecting," Ph.D. Dissertation, Engineering Systems Division, Massachusetts Institute of Technology, Cambridge, MA, 2005.

[21] Leveson, N. G., *Safeware: System Safety and Computers*, Addison–Wesley, Reading, MA, 1995.

J. Martin
Associate Editor