The following paper will be published and presented at the 2012 IEEE International Systems Conference in Vancouver, Canada, 19-23 March, 2012.

A Taxonomy of Perturbations: Determining the Ways That Systems Lose Value

Brian Mekdeci, Adam M. Ross, Donna H. Rhodes and Daniel E. Hastings
Systems Engineering Advancement Research Initiative (SEAri)
Massachusetts Institute of Technology
77 Massachusetts Avenue, Building E38-572
Cambridge, MA 02139
http://seari.mit.edu

Abstract— **Disturbances and disruptions, both internal and external to the system, are a major concern for system architects who are responsible for ensuring that their systems maintain value robustness no matter what occurs. These perturbations can have multiple causes and can affect a system in multiple ways. This paper presents a taxonomy of disturbances and disruptions to assist system architects and researchers in identifying the ways in which systems can fail to deliver value. By doing so, this taxonomy falls into a larger research effort to develop survivability design principles that will help system architects design systems that prevent, mitigate and recover from disturbances.**

Keywords-surviability; robustness; safety; reliability; systems; systems of systems; disturbances; disruptions; perturbations.

I. INTRODUCTION

Systems always operate within a particular context, which includes specific environmental conditions, resources and stakeholder expectations. For many complex systems that have long lifecycles, the context is expected to be dynamic and sometimes unfavorable to the system's mode of operation. Even if the new context is advantageous, the change itself may be problematic, particularly if the change is sudden, unexpected and/or significant. Even if the context is static, systems themselves may change (or be changed) as well, either intentionally, or unintentionally. Regardless of whether the context changes or the system changes, whether the perturbations were intentional or unintentional, stakeholders desire a value-robust system that will continue to effectively perform *no matter what*.

In the authors' research efforts, a major goal is to develop a set of design principles that will guide system architects in identifying and selecting design choices that will sustain a threshold level of value delivery *no matter what*. To do that, the *what* needs to be clearly understood, both by the system architects, who have to select design choices based on what might happen, and by researchers who seek to identify these design principles. However, understanding *what* the problem is can be quite difficult. To illustrate this, consider the following example: An exhausted pilot is flying through a thunderstorm. The rain has reduced his visibility, and he is too tired to notice that the plane's altitude is low. Suddenly, the plane clips the side of a tower, loses a wing, spirals out of control, crashes into the ground, and explodes in a fiery inferno. How can system architects make a plane survivable in such a scenario? They could design a plane that can safely land with only one wing, but will this make the plane survivable? If the architects are not including low visibility and pilot exhaustion as part of the context, then their Concept of Operations (CONOPs) for safely landing on just one wing, i.e. how the pilot and components of the plane interact with each other and the environment so that the plane lands safely, may not actually make the plane survivable. Additionally, the collision with the tower could have done more damage than just clipping the wing, and thus this particular survivability solution may only be effective in extremely rare situations. Perhaps looking at the cause of the crash would be more productive. The system architects could include windshield wipers to mitigate the rain's effect on visibility, but a similar visibility problem could arise simply from the plane flying at nighttime, as well. Either visibility problem could lead to a loss of situational awareness by the pilot, which leads to the collision with the tower. Preventing and mitigating the collision with the tower may be the best way to achieve survivability, both for this scenario and others (such as flying off course and random component failure).

This paper describes a research-derived taxonomy by which perturbations can be characterized to help both researchers and system architects identify the ways in which value delivery can be degraded, so that they can design systems to avoid, mitigate and recover from a large variety of endogenous and exogenous changes which may arise during the system's lifetime.

II. DEFINITIONS

A. Perturbations, Disruptions and Disturbances

Suppose an aircraft in flight experiences complete and sudden engine failure. Although there may be numerous causes for the engine failure and the circumstances that led to the failure may have taken a certain finite time to develop, the actual moment of failure can be considered an impulse event of zero (or near-zero) duration. In the period immediately following the failure, the aircraft is forced to make an emergency landing instead of flying to its original destination. We can define a *disruption* as an unintended, instantaneous, discontinuous state change of a system's form, operations, or context, which could jeopardize value delivery. The sudden failure of the engine is an example of a disruption. Since systems are expected to provide satisfactory value under ideal

conditions, a *disturbance* can be defined as an unintended, finite duration, continuous state change of a system's form, operations, or context, which could jeopardize value delivery. In this case, flying without an engine is an example of a disturbance. A *perturbation* is any unintended state change of a system's form, operations, or context which could jeopardize value delivery. Both disruptions and disturbances are perturbations, but they are distinct. In the limit a disturbance's duration becomes zero, the disturbance becomes a disruption. A sudden but permanent context change, such as what might occur if a new environmental law comes into effect, is an example of a disruption but not a disturbance. Disturbances typically start with a disruption, such as flying after an engine failure. However, if the changes are gradual, then a disruption does not occur. To illustrate this point, consider an example where a UAV is taking images of the ground while it flies at a high altitude. The weather may change from sunny to cloudy over time, which naturally affects the quality of images taken from the UAV. While the image quality may degrade gradually as the weather worsens, at some point the quality crosses a threshold where the overall value is below stakeholder needs. At this point, the weather impact on the UAV would be considered a disturbance (e.g., sensor degradation) even though there is no disruption per se.

B. Survivability

Survivability has been defined as *the ability of systems to minimize the impact of finite-duration disturbances on value delivery* and numerous design principles have been created to help build, implement and operate survivable systems [1]. There are three types of design principles to achieve survivability; prevention (Type I), mitigation (Type II), and recovery (Type III) (Figure 1).

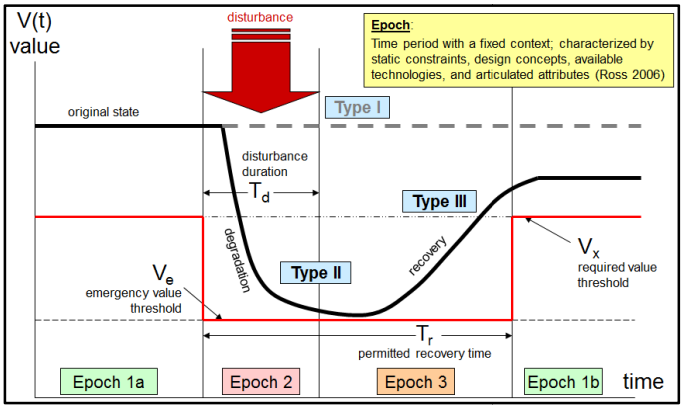

Figure 1. Definition of survivability (Richards, 2009)

This definition of survivability can be a bit misleading, as illustrated in an example where stakeholders at Ford may want to design a Mustang that "survives" low-impact collisions. A crash, however, is an instantaneous event of near-zero duration, so the application of survivability in this case is to minimizing the impact of a disruption, rather than a disturbance. Recognizing that the duration of the perturbation affects the appropriate design principles for survivability, designers can use the appropriate principles to avoid, mitigate and recover from all perturbations of interest. One of the reasons it is important to distinguish between disruptions and disturbances is that Type II survivability depends on the ability of the system to mitigate a disturbance as it happens. In some cases, a system may have time to change its mode of operation during a disturbance to help mitigate the decrease in value delivery. For disruptions, however, mitigation is not possible because the system does not have time to react. Thus, only passive Type II survivability design principles, such as increasing the system's *hardness* (i.e. increasing the intensity of the perturbation required to damage the system), will be effective in mitigating disruptive changes in the context or system.

C. Threats and Hazards

In management science, there is a saying that "nobody ever gets credit for fixing problems that never happened" [2]. Thus, it is fair to ask, if a disturbance is prevented, was it really a disturbance? Instead of preventing disturbances and disruptions from occurring, a system should try to eliminate or mitigate threats and hazards. A *threat* is an external set of conditions that exist which may cause a perturbation, but hasn't impacted value delivery, yet. Enemy occupation of an area, bad weather approaching or a change in political climate may all be threats to a system. If the system responds to the threat by changing its form or mode of operation in such a way that value has been impacted, then that threat has led to a disruption and/or disturbance (the actual system change in response to the threat). We can assume that any value-loss as a result of a system response is unintentional, even if the actions taken by the system that caused the loss were deliberate. In these cases, either the actions were accidents, or the system effectors were "forced" to perform actions because no better options were available. A *hazard* is a set of conditions inside the system that could cause a perturbation, such as poor operator training, operator fatigue, under-maintained components and outdated protocols. Eliminating or mitigating hazards is the domain of *system safety* [3].

D. Cause and Effect

Anything that causes a reduction in the value delivery of a system has at least one cause and at least one effect. Many perturbations have multiple causes and multiple effects (Figure 2). Each cause is a set of conditions that led to the perturbation. The conditions may be earlier perturbations, threats or hazards. The effects of a perturbation are the immediate changes in the context and/or system that are a direct result.

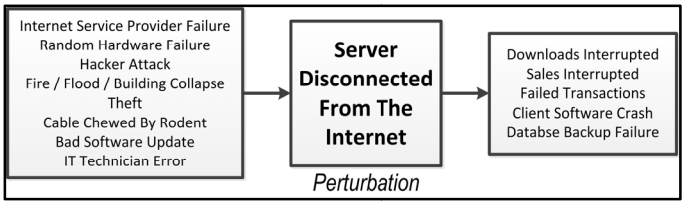

Figure 2. Multiple causes and effects of a single perturbation.

It is important to note that the effects of many incidents are often the cause of others. Fig 3 shows a chain of perturbations that occurred during the US Northeast Blackout of 2003. The initial incident which started the chain reaction was an untrimmed tree in Ohio that came too close to high voltage power lines and caused a flashover which severed a power line [4]. This link failure meant that the current had to be routed

through other lines in the network. This increase load in the lines caused some of them to fail as well. As more lines failed, more load was placed on the remaining lines until they failed as well, leading to what is referred to as a cascading failure. Eventually, entire power plant generators were shut down, which lead to a massive blackout in the Northeastern United States and some parts of Canada. In some parts, traffic lights failed as well, which lead to some fatalities including a teenage cyclist struck by a car in Ontario. Indirectly, the death in Ontario was due to an untrimmed tree in Ohio, although it was not the sole cause. There were multiple incidents and circumstances along the way that also caused this unfortunate outcome. For instance, there was a software bug in one of the Ohio power plants that prevented authorities from recognizing and mitigating the power outage earlier [5]. Also, the teenager did not have lights on his bicycle, nor was he wearing a helmet, both of which are speculated to have contributed to the fatality [6].

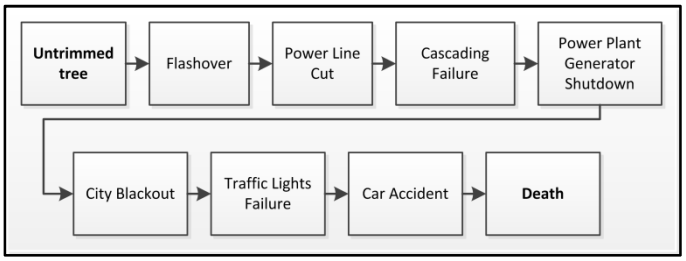

Figure 3. Chain reaction during the Northeast Blackout of 2003.

Typically, strategies for Type I survivability involve mitigating and reducing threats and hazards before they become perturbations (i.e., addressing the "cause"), whereas Type II and III survivability design principles address mitigating and recovering from the effects. Since certain effects have multiple causes and certain causes have multiple effects, separating perturbations into cause and effect is critical because it allows system architects to focus on addressing the causes and effects that will have the most impact on the system's survivability. System architects do not have the resources to address all possible perturbations. Similarly, research has suggested that some mechanisms for survivability, such as dependability and security, are incompatible [7]. Thus, by realizing that certain perturbations have multiple causes and/or effects can help system architects identify the strategies that will have the greatest impact on overall survivability and trade off design options accordingly.

Another reason it is important to separate perturbations by cause and effect is because it helps system architects focus on what they can affect and what they can't. Without thinking about cause and effect, architects of the UAV aerial imaging system may not understand how they can design a system that can survive "bad weather." Specifically, how does a system avoid "bad weather" (Type I survivability), particularly if the system is stationary? How does it mitigate "bad weather" and recover from it? Rather, system architects must realize "bad weather" is a disturbance that causes many other disruptions and disturbances, and these perturbations are the ones over which the system has more control. For example, one effect that bad weather has on the system is the blurry image caused by precipitation on the lens of the camera. In this case the effect (blurry image) and the cause (precipitation on the lens of the camera) are very clear, and appropriate design principles can be applied. Type I survivability might focus on preventing exposure to rain by sheltering the lens and preventing the precipitation from coming into contact with the lens, while Type II survivability might focus on image processing techniques that specifically correct for precipitation-induced blurry images. Thus, by clearly specifying the disturbance as a specific cause and effect, as opposed to an ambiguous statement like "bad weather," system architects can more easily develop and incorporate survivability strategies.

III. Determing a Suitable Taxonomy

There are many ways to classify a perturbation. Initial attempts at classification yielded categories such as "origin" (whether the perturbation started externally or internally) and "intent" (whether the perturbation was caused intentionally by an intelligent entity, or not) [8]. These types of classifications are very useful if the system architects want to target specific types of perturbations and ignore others. For instance, a system architect of a yogurt manufacturing plant may not feel it is necessary to protect against hacker attacks because the plant does not have any hostile adversaries. Of course, not every type of perturbation can be prevented, mitigated or recovered from and priorities have to be considered. However, to dismiss entire classes of perturbations without analysis is risky as well, since many of the largest system failures came about from events that the system architects never considered (e.g., 9/11 attacks and 2003 Northeast Blackout). Sometimes the solution to one problem is also the solution to another. For example, an authentication procedure can not only protect against unintentional purchases by legitimate users (errors), but also protect against unauthorized attacks or security compromises. If a taxonomy is developed by which perturbations are categorized independent of their effect on the system, such as "artificial" vs. "unintentional," then in the presence of time and resource constraints, there is a risk that system architects will implement solutions based on perturbations that are expected (i.e., "known unknowns"). However, what about perturbations whose very nature we do not know because they did not exist when we designed the system, or were never a problem before? It would be better to focus on the effects on a system and categorize disturbances that way, so that designers can better handle disturbances that have similar effects, even if they do not understand or cannot predict what those disturbances may be, specifically (i.e., "unknown unknowns").

IV. Analyzing Systems For Possible Perturbations

There are several ways to analyze systems for possible perturbations. Two of the more popular are Fault Tree Analysis (FTA) and the Failure Mode Effects (and Criticality) Analysis (FMEA/FMECA).

A. Fault Tree Analysis

Fault Tree Analysis is a top-down approach [9] that uses Boolean logic to determine the causes of a particular single failure. Since FTA is a deductive approach, which starts with a failure state and works backwards towards single events that

may have been responsible, it does not always find all possible initiating faults.

B. FMEA/FMECA

FMEA/FMECA looks at initiating faults, and tries to determine their immediate and subsequent effects (i.e., "failures") on the overall system. FMEA/FMECA describes "failure" as the "loss of an intended function of a device under stated conditions" [10], which addresses component / capability failures, but does not address operational perturbations. Additionally, FMEA/FMECA often does not consider human/software failures, nor does it address combined failures [11].

C. Cause and Effect Mapping

Recognizing the complex relationship between disruptions, disturbances, causes and effects, a mapping such as Fig 4 can highlight a perturbation's relationship with other perturbations, allowing system architects to recognize potential cascading failures and common problems, and prioritize survivability efforts accordingly. Only perturbations and threats that can influence the system (or for which the system has some form of control) are considered. Since the mapping is generalized, it allows system architects to generate possible perturbations of interest by thinking about what might cause a perturbation, and what other effects may it have. New causes and effects that fit or don't fit into the existing mapping are added and new relationships can be drawn to various existing causes and effects. Once the generalized cause and effect mapping is done, specific perturbations and hazards from the categories are chosen for consideration, relevant to the particular system and context under consideration. To deal with these known perturbations and hazards, a detailed hazard analysis such as FMEA/FMECA or FTA along with existing survivability design principles, can help system architects develop specific survivability strategies. An important property of Fig 4 is the non-linearity of the mapping. For example, an *operator error* is a perturbation that can easily be the initial starting point for a complex failure. For no other reason other than daydreaming, an operator can make a mistake and push the wrong button, which can start a chain reaction of events that cause multiple problems. However, an operator error can also be the result of other perturbations. If there is bad weather, a pilot may make a mistake due to the poor visibility. If communications are down, an operator may make a mistake by assuming that there are no other vehicles in the area. If a constituent system leaves a system of systems and the workload increases for the remaining systems, then an operator may make a mistake simply because they are overworked. Even worse, perturbations may be repeated and even grow in intensity. If an operator error causes an increase in workload (for instance by repeating a task that was already done), then that increase in workload may cause additional operator errors, which may in turn increases the workload, and so forth.

V. Classifying Perturbations

A first attempt at classifying perturbations for systems, such as those that perform maritime security using UAVs, is presented in Table 1. After brainstorming, the table begins with some examples of perturbations of interest, such as an aircraft colliding with another aircraft, or an increase in fuel prices (some "known unknowns"). Only the immediate effects are noted, such as physical damage to components, or cost increase. These effects, which are general in nature, are then added to the cause and effect mapping (Fig 4). Then, looking at a particular effect on the cause and effect mapping, other possible causes and effects for it is added, in a general way, to the map. These may generate further specific disruptions and disturbances, which are then added to the perturbation table, and so on. Finally, once brainstorming has not generated any additional perturbations, a final column discussing possible survivability solutions to the specific perturbations is added to the table. Not surprisingly, since some of the perturbations share similar cause and effect characteristics, these can be addressed by similar survivability design principles.

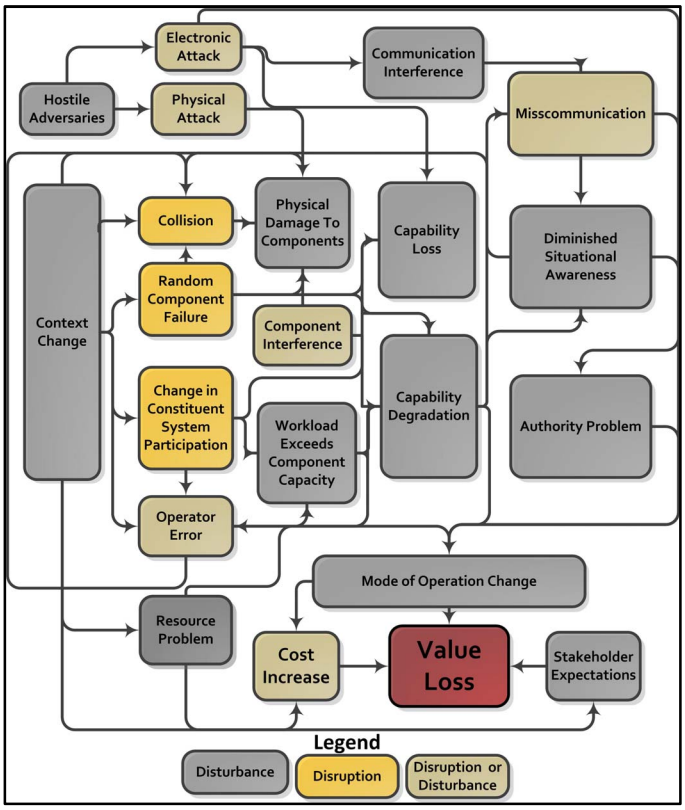

Figure 4: Cause and effect mapping.

Although the perturbations examples presented in Table 1 may seem dissimilar at first, a few commonalties emerge. Generally speaking, all perturbations of interest may eventually lead to value loss. Value loss occurs when one or more of the following main effects result from a perturbation: capability loss, capability degradation, change in mode of operations, cost increase, or change in stakeholder expectations.

TABLE 1. EXAMPLE PERTURBATION TABLE FOR A MARITIME SECURITY SYSTEM

Perturbation Example	Type	Immediate Effect	Main Effects	Causes of Perturbation	Survivability Solutions
Aircraft struck by lightning	Disruption	Physical damage to components	Capability loss, capability degradation	Context change (weather)	Decrease cross-sectional area, divert lightning away (e.g., lightning rod)
Crash between system and other mobile entity, crash between system and environment	Disruption	Physical damage to components	Capability loss, capability degradation	Collision (caused by operator error, context change, diminished situational awareness)	Decrease cross-sectional area, increase maneuverability, increase situational awareness
High winds creating turbulence	Disturbance	Mode of operation change	Change in mode of operation	Context change (weather)	Move away from disturbance, better aerodynamic design, heavier mass
Precipitation builds on lenses	Disturbance	Capability loss	Capability loss	Context change (weather)	Wipers, better image processing algorithms
Fuel price increase	Disruption	Cost increase	Cost Increase	Resource scarcity, mode of operation change	Store excess resource when not scarce, change to alternate resource
Environmental ozone regulation makes component obsolete	Disruption	Capability loss	Capability loss	Context change (environmental)	Use alternate capabilities to achieve desired outcome.
Operator gives wrong command to machine	Disruption	Capability degradation	Change in mode of operation	Context change (weather, bad working conditions), workload exceeds component capacity	Increase capacity (increase operators, increase automation), increase training, double check operator instructions
Communication interference	Disturbance	Miscommunication	Change in mode of operation	Context change (weather), electronic attack	Eliminate unnecessary communications (co-locate components) , switch to alternate communication channel, increase signal power
Noise from one UAV interferes with audio recording of another	Disturbance	Component interference	Capability degradation	Close proximity of components, tight coupling	Reduce proximity of components, reduce noise
Missile strikes aircraft	Disruption	Physical damage to components	Capability loss, capability degradation	Hostile adversaries, large cross-sectional area, enemy has capability	Use aircraft instead of land or sea vehicles, decrease cross-sectional area, increase deterrence and decrease intent of hostile adversaries (e.g., political pressure), preemptively strike
Friendly artillery unit withdraws from SoS	Disruption	Capability loss	Capability loss	Component has operational / managerial independence	Have redundant components, alternative CONOPs
Hacker attack	Disturbance	Authority problem	Capability loss, capability degradation, change in mode of operation	Hostile adversaries	Secure authentication, network analysis tools, increase deterrence and decrease intent of hostile adversaries (e.g., political pressure), preemptively strike
Random component failure	Disruption	Random component failure	Capability loss, capability degradation	Context change (weather)	Decrease exposure to hostile environments, limit use
Miscommunication between components	Disruption and/or Disturbance, depending on how long it lasts	Change in mode of operation	Change in mode of operation	Communication interference, Capability loss / degradation	Redundant communication channels, error checking, elimination of unnecessary communications (e.g., co-locate components)

A. Capability Loss

The form of the system is the current set of operational elements and capabilities that a system has. If the form changes, particularly if a component is functionally removed (i.e., no longer functioning in the system), then value delivery is likely impacted. There are many reasons a component can be lost, such as if it is voluntarily or involuntarily removed, critically damaged by an external force, fails randomly, does not have the resources to function, or is unwilling to participate (a problem associated with autonomous component systems in a larger SoS).

B. Capability Degradation

Capability degradation is when a component is still performing according to the CONOPs, but not as well as it should be. For instance, a CPU under load may not be able to respond to the tasks it needs to in a timely manner. Degradation can be the result of a number of circumstances including physical damage, insufficient resources and excessive demand.

C. Change in Mode of Operation

Perturbations can cause the system to operate differently, either to a viable mode of operation specified in its system architecture [12], or to other some mode of operation. For example, turbulence may force an aircraft to fly a different route and altitude than it normally would fly, increasing the flight time and burning more fuel. Naturally, value delivery of the aircraft to the stakeholders is decreased as a result. In systems with autonomous, or semi-autonomous constituent systems (such as most systems of systems), coordination of the capabilities must be maintained at all times. For many reasons, such as miscommunication, authorization problems, and diminished situational awareness, coordination errors occur. These errors include more than one entity attempting to perform the same task, no entity performing a required task, tasks being performed out of order, and so forth.

D. Cost Increase

All engineered systems have an associated cost, which typically is to be minimized. If a disturbance increases the cost of the system, without increasing the performance as well, then it will likely decrease the value it provides to stakeholders.

E. Change in Stakeholder Expectations

The value of a system depends upon stakeholder expectations. If these expectations change, then the value of the system can change, even if cost and performance remain static.

VI. Discussion and Future Work

The eventual goal of this research is to develop a set of design principles that will guide system architects in recognizing and evaluating system design options, both in form and mode of operation, which will increase the system's ability to deliver value *no matter what*. This paper works toward that goal by clarifying that disruptions and disturbances have different event durations, so that system architects can recognize their impact on potential survivability design choices. This paper also shows that by using causal chains and working backwards from value impact, a system architect can begin to determine appropriate locations for intervention, based on what is within their control and resource constraints. Finally, this paper shows that by using a cause and effect mapping, there are general categories of effects that can be useful as a taxonomic basis, especially for dealing with known unknowns and potential unknown unknowns.

References

[1] M. G. Richards, "Multi-Attribute Tradespace Exploration for Survivability," PhD, Engineering Systems Division, Massachusetts Institute of Technology, Cambridge, MA, 2009.

[2] N. P. Repenning and J. D. Sterman, "Nobody ever gets credit for fixing problems that never happened," California Management Review, vol. 43, pp. 64-88, 2001.

[3] N. G. Leveson, Safeware: system safety and computers vol. 12: Addison-Wesley Boston, MA, 1995.

[4] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, N. Martins, J. Paserba, P. Pourbeik, and J. Sanchez-Gasca, "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," Power Systems, IEEE Transactions on, vol. 20, pp. 1922-1928, 2005.

[5] M. Zhivich and R. K. Cunningham, "The real cost of software errors," Security & Privacy, IEEE, vol. 7, pp. 87-90, 2009.

[6] C. Greeno, "Blackout crash injuries claim teen.," in Daily Mercury, ed. Guelph, Ontario, 2003, p. 1.

[7] J. Rushby, "Critical system properties: survey and taxonomy," Reliability Engineering & System Safety, vol. 43, pp. 189-219, 1994.

[8] Mekdeci, B., Ross, A.M., Rhodes, D.H., and Hastings, D.E., "Examining Survivability of Systems of Systems," INCOSE International Symposium 2011, Denver, CO, June 2011.

[9] P. Fenelon, J. A. McDermid, M. Nicolson, and D. J. Pumfrey, "Towards integrated safety analysis and design," ACM SIGAPP Applied Computing Review, vol. 2, pp. 21-32, 1994.

[10] J. W. Langford, Logistics: Principles and Applications. : McGraw Hill, 1995.

[11] F. A. Administration, "Research Development Accomplishments FY 2004.," Washington, DC2004.

[12] B. Mekdeci, A. M. Ross, D. H. Rhodes, and D. Hastings, "System Architecture Pliability and Trading Operations in Tradespace Exploration," presented at the IEEE International Systems Conference, Montreal, PQ, 2011.