A Systemic Approach to Risk Analysis of Infrastructure Systems of Systems

---

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment

of the requirements for the degree

Doctor of Philosophy

by

Zhenyu Guo

May

2014

APPROVAL SHEET

The dissertation

is submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Zhenyu Guo

AUTHOR

The dissertation has been read and approved by the examining committee:

Dr. Yacov Y. Haimes

Advisor

Dr. Barry Horowitz

Dr. Donald E. Brown

Dr. Zongli Lin

Dr. Garry M. Jacyna

Accepted for the School of Engineering and Applied Science:

James H. Aylor

Dean, School of Engineering and Applied Science

May

2014

Abstract

Many of the nation's large-scale physical infrastructure systems are commonly composed of interconnected and intra- and interdependent subsystems, which in their essence constitute systems of systems (SoS) with multiple functions, operations, and stakeholders. Their complexity is characterized by the highly interconnected and interdependent physical, cyber, organizational and economic subsystems through shared resources, decisions and states, which constitute a major source of systemic risks inherent to the system and pose great challenges in their risk modeling, assessment, and management. To meet the increasing needs of reliable services provided by these infrastructure systems, system owners and decision makers need tools to foresee potential emergent forced changes from within and outside the system and to understand their impacts so that efficient risk management strategies can be developed.

Risk analysis of complex SoS requires a systemic and holistic approach that integrates multiple perspectives, models and tools. The focus of this dissertation is to develop a systemic framework of precursor analysis, which supports the design of an effective and efficient precursor monitoring system having the ability to i) identify indicators or warnings of dynamic and evolving risks to system failure; (ii) monitor critical precursors to system failure through continuously tracking and observing triggering changes in the states of the system; and (iii) reduce the hindsight bias frequently observed between pre- and post- accident risk assessment when using precursors. This pro-active and dynamic anticipatory analysis is supported by meta-modeling the functional components and subsystems of the SoS, and their relationships in

a control structure and is achieved through a process of precursor identification, prioritization, detection, and evaluation.

The identification of precursors to system failure requires an understanding of system failure mechanism. This dissertation explores potential sources of systemic risks in complex SoS through analyzing a unique failure mode of the system in a nonlinear dynamic multi-objective decision process. It demonstrates that the decision maker's inappropriate preference among multiple competing objectives and the interdependencies between uncoordinated subsystems contribute to the failure of complex SoS even though all its components are functioning correctly. The results also suggest that an optimal decision strategy doesn't necessarily guarantee system safety. Through quantifying the level of subsystem interdependency caused by common states, this dissertation develops a method to decompose interconnected subsystems within SoS and a method to coordinate multiple subsystems in a decentralized way.

This dissertation demonstrates the theories and methodologies with a case study on the US highway bridge system. Highway bridges, which constitute large- and multi-scale physical infrastructure systems, and are essential elements of transportation networks, have a large number of interconnected and interdependent sub-systems, with broad social and economic consequences from bridge failure. The precursor analysis framework allows examining the impacts of current bridge inspection, maintenance, and decision practices on the overall reliability of bridge infrastructure systems; enables decision makers to make more timely and informed decisions to efficiently allocate limited risk management resources; and thus, prevent future severe consequences resulting from future bridge failures.

Acknowledgements

First of all, I would like to express my sincere appreciation to my advisor, Dr. Yacov Haimes, an admirable educator and a wonderful mentor, for his guidance, encouragement, patience, and continuous support throughout the course of my study. The extensive knowledge, vision, and creative thinking of Dr. Haimes have been the source of inspiration for me throughout the development of my dissertation. I have learnt a great deal from him, not only in the field of risk analysis, but also in how to become a better researcher, learner, and thinker. In addition to a good advisor, he and his wife, Sonia, are also good friends who helped me and my family through these years in many aspects of our lives. I will never forget their warmness and kindness.

I would also like to extend my appreciation to my committee members: Dr. Barry Horowitz, Dr. Donald Brown, and Dr. Zongli Lin, for their insightful advice and continuous support as I moved from an idea to a completed study. Special thanks is also extended to Dr. Garry Jacyna, whose amazing ideas inspired and shaped my work in many aspects and who provided thought-provoking feedback and comments to my dissertation.

My special gratitude goes to Dr. Kenneth Crowther who encouraged me to embark this academic journey, Dr. Steven Chase for providing his expertise and guidance in the field of transportation and structural engineering, Dr. Michael Smith for his support for the AMP program, Dr. James Lambert as an excellent project leader, and Erika Evans for her invaluable support and assistance at the Center for Risk Management. I was also fortunate enough to have Dr. Eva Andrijcic as one of my closest colleagues and she has been very helpful and supportive in the face of numerous obstacles.

I owe my thanks to all the faculty, staff, and fellow graduate students at the Department of Systems and Information Engineering for providing the abundant resources and wonderful academic environment. They made my years at UVa to be an exciting experience. My sincere thanks also goes to the many current and former graduate students I have met and worked together at the Center for Risk Management during these years.

I am especially grateful to Hua Hui, my beloved wife, for the innumerable sacrifices made by her in shouldering far more than her fair share of the parenting and household burdens. Without her encouragement, support, and of course the delicious food she prepared, the completion of this dissertation would be impossible. She was always there cheering me up and stood by me through the good times and bad. We enjoyed our simple life in Charlottesville with a lot of sweat memories. To my two little kids, David and Joanna, who missed out on a lot of Daddy time while I sought intellectual enlightenment, and to my parents for giving me a chance to thrive.

List of Figures

List of Tables

Acronyms and Abbreviations

ASAP – Aviation Safety Action Program
ASCE – American Society of Civil Engineers
ASIPS – Applied Strategies for Improving Patient Safety
ASP – Accident Sequence Precursor
ASRS – Aviation Safety Reporting System
COA – Couse of Action
DOT – Departments of Transportation
FHWA – Federal Highway Administration
FMEA – Failure Mode and Effect Analysis
FMECA – Failure Mode Effect and Criticality Analysis
FO – Functionally Obsolete
FTA – Fault Tree Analysis
HHM – Hierarchical Holographic Modeling
HOT – Highly Optimized Tolerance
UNISDR – United Nations International Strategy for Disaster Reduction
NBI – National Bridge Inventory
NBIS – National Bridge Inventory System
NTSB – National Transportation Safety Board
OL – Overload
PHA – Preliminary Hazard Analysis
PRA – Probabilistic Risk Analysis
PSM – Phantom System Model
PSRS – Patient Safety Reporting System
RFRM – Risk Filtering and Ranking Method
SAE – Society for Automotive Engineers
SCADA – Supervisory Control and Data Acquisition
SD – Structurally Deficient
SoS – Systems of Systems
STAMP – Systems-Theoretic Accident Model and Processes
TSS – Theory of Scenario Structuring

UNEP – United Nations Environment Program

USDOT – The US Department of Transportation

List of Symbols

A, B – coefficient matrices of state transition equation of a system or subsystem

c – control objective of the output variable of a system or subsystem

$C(x,t)$ – the chloride ion concentration at a distance of x cm from the concrete surface after t seconds of exposure of the chloride source

D_c – the chloride diffusion coefficient expressed in cm^2/sec

C_0 – the equilibrium chloride concentration on the concrete surface

Es – error between the inputs and output between two subsystems

Et – error between consecutive time stages for the same subsystem

F – objective function of a system or subsystem

k – index of time

Pr – probability

S – vector of state variables of a system or subsystem

t – time

u – vector of control/decision variable of a system or subsystem

u^* – vector of optimal control/decision variable of a system or subsystem

x – vector of input variables from one subsystem to another subsystem

y – vector of output variables of a system or subsystem

$\tilde{y}$ – vector of output variables adjusted by constraints of a system or subsystem

Zs – vector of output variables from one subsystem to another subsystem

Zt – vector of output variables between consecutive time stages for the same subsystem

β – coefficient matrices of state difference equation

Δ – perturbation to the subsystem

θ – weighting factor between the two objective functions of a specific system or subsystem, $0 \leq \theta \leq 1$

λ, μ – vector of Lagrangian multipliers

1. Introduction

1.1. Motivation

Physical infrastructures, serving as the foundations of society's wellbeing, encompassing the entire private and public sectors, are the driving force of today's social and economic development. By virtue of their multifarious reach in our lives, they have also become the focal interest of diverse stakeholders that span local, state, and federal governments. The need to better understand and improve current status of the U.S. physical infrastructure systems is evidenced in reports spanning the last two and a half decades. As early as 1988, a national commission issued a report titled Fragile Foundations, citing: "*The National Council on Public Works Improvement … has found convincing evidence that the quality of America's infrastructure is barely adequate to fulfill current requirements, and insufficient to meet the demands of future economic growth and development*" [National Council on Public Works Improvement 1988]. This sentiment was recently revisited in reports issued by the American Society of Civil Engineers (ASCE), including a report titled "Can We Come Back From the Brink," which concludes that the U.S.

surface transportation infrastructure is “failing to keep pace with the expanding needs of a burgeoning population” [ASCE, 2009b]. In a series of ongoing report cards, the American Society of Civil Engineers [ASCE 1998, 2003, 2005, 2009a, 2013] gives a sober picture of the nation’s infrastructures, highlighting maintenance needs and addressing the “fragile foundations” of the infrastructures as evidence that citizens and governing bodies are increasingly recognizing the desperate state of our nation’s infrastructure systems. The most recent issue of the ASCE Infrastructure Report Card gives America’s civil infrastructures a grade D+ (poor), and estimates that a $3.6 trillion investment will be required by 2020 to bring the U.S. infrastructure systems to an acceptable condition, of which $1.7 trillion will be needed to improve the condition of the surface transportation infrastructure (including roads and bridges) [ASCE, 2013]. Furthermore, the recent stimulus package by Congress, which includes substantial funds for bridges and other physical infrastructures, highlights the importance of principled and strategic investment. The practice of persistent infrastructure underinvestment, coupled with a significant growth in commercial and non-commercial transportation demand, has left U.S. transportation infrastructure “*stuck in the last century and ill-equipped for the demands of a churning global economy*” [Building America’s Future Educational Fund, 2011]. Meanwhile, the likelihood and potential adverse consequences of infrastructure failures caused by emergent forced changes, which connotes external or internal sources of risk that may adversely affect the states of the system [Haimes, 2008, 2012], both natural and intentional, are increasing. Examples of these forced changes originate from both within and outside of the system, including global climate change, global and national economic crisis, terrorist activities, cyber war and crimes, increasing user

demands, physical deterioration, and the lack of maintenance resources. Thus, the need to better maintain US infrastructure systems is real and urgent. However, the complexity of the infrastructure systems and the lack of theories and methodological approaches to understand and analyze the risks associated with their behaviors significantly complicate this effort.

Many of the nation's critical infrastructure systems fall within the category of complex systems. Complex systems are commonly composed of interconnected and intra- and interdependent subsystems, which in their essence constitute systems of systems (SoS) with multiple functions, operations, and stakeholders [Haimes, 2012]. The emergent, large-scale engineering systems, such as ground transportation, aviation, supply chains, power grid, and cyber infrastructure systems, pose great challenges in their risk modeling and management. The complexity of SoS is characterized by the highly interconnected and interdependent physical, economic, and social components, which constitutes a major source of emergent forced changes to the system. To meet the increasing needs of reliable services provided by these infrastructure systems, system owners and decision makers need tools to foresee potential emergent forced changes from within or outside the system, identify interdependencies among its different system components, and understand the impacts to the systems so that efficient risk management strategies including preparedness and response plans can be developed. Risk assessment, communication, and management are indispensable tools to evaluate the states of the system, reduce its vulnerability and increase its resilience to any emergent forced changes. Developing and applying risk analysis theories and methodologies for these

complex physical infrastructures SoS is a key requirement to manage critical infrastructure systems effectively and efficiently.

The theories and methodologies developed in this dissertation are driven by the needs from risk analysis of these large-scale physical infrastructure systems. This dissertation aims to develop a systemic precursor analysis framework for complex physical infrastructure SoS, in which a pro-active, dynamic anticipatory analysis tool is designed to identify, prioritize, detect, and evaluate different sources of emergent forced changes that have the potential to cause system failure. It first explores some systemic risks of complex infrastructure SoS through analyzing a unique failure mechanism from a control perspective, and then discusses an approach to identifying precursors to system failure through systemically exploring a meta-model of system failure mechanisms. The likelihood of different failure modes of the system based on information from identified precursors is then evaluated, so that decision makers can make more timely and informed decisions to response to the risks and prevent severe consequences in the future. The proposed approach is expected to be used as a complement and supplement tool to the static, passive Probabilistic Risk Analysis (PRA) approach.

Three indispensable phases are proposed for a formal precursor analysis process: (i) system modeling; (ii) precursor identification, filtering, and prioritization; and (iii) precursor detection and evaluation. In the system modeling phase, a generic meta-model based on control theory is introduced to model the failure mechanism of complex infrastructure SoS. The precursor identification, filtering, and prioritization phase aims to select the most important precursors leading to system failure based on its likelihood and urgency, in order to design an efficient precursor monitoring system. The precursor

detection and evaluation phase focuses on the quantification of detection uncertainties and the provision of improved situation awareness to the decision makers through a holistic, system-level understanding of the risks of system failure.

This methodology is demonstrated along with a case study on highway bridge infrastructure systems. Highway bridges constitute an important part of transportation infrastructure and the lifelines of commerce. The dismal state of our nation's bridges has been well documented in the literature. The condition of highway bridges is continuously deteriorating due to the lack of appropriate maintenance, with 26% of America's bridges structurally deficient or functionally obsolete [ASCE, 2009a]. Many bridges receive insufficient inspection and maintenance due to limitations of funding, equipment, manpower, and available technology. A bridge infrastructure system is a system with many interdependent functional components, which are also managed by different decision-making organizations. The developed methodology is expected to help bridge owners to efficiently prioritize and plan for inspection, maintenance, and rehabilitation activities based on precursors, and to reduce the risk of bridge failure.

1.2. Problem Definition

Infrastructure system owners seek to understand the trends of risks associated with emergent forced changes that affect the states of their systems, in order to prevent, mitigate, or prepare for undesirable future occurrences. Unanticipated, undetected, misunderstood, or ignored emergent forced changes, whether they originate from within or from outside a system, are likely to affect a multitude of states of that system with potentially adverse consequences. Therefore, it is imperative to be able – through

scenario structuring, modeling and risk analysis – to envision, discover, track precursors to, collect data, and measure emergent forced changes to the system.

The complexity in highly interconnected and interdependent subsystems constitutes one of the major sources of emergent forced changes to the system. Challenges in risk analysis of complex physical infrastructure SoS emerge from both system modeling and risk assessment. For the modeling perspective, no single model is able to represent the multiple perspectives of complex SoS. One key issue in modeling SoS is how to identify and quantify the causal intra- and interconnected and interdependent relationships among subsystems. System models that fail to consider subsystem interdependencies are unable to capture the impacts from decisions for one subsystem on other related subsystems, thus fail to uncover intricate complex interactions and causal relationships among the myriad components and subsystems that constitute SoS.

One modeling approach to model inter-subsystem interdependency, the extrinsic modeling approach, assumes the output from one subsystem is the input to another subsystem. This input-output modeling approach [Leontief, 1951a, 1951b], [Haimes and Jiang, 2001], [Lian and Haimes, 2006] has been widely-used to model systems where the exchange of physical commodities constitutes a major source of inter-system interdependencies, such as production processes, supply chains, and the dynamics of general economy. However, the characteristics of the subsystems in SoS suggest that their interdependencies rely more on sharing common states, resources, and information as well as decisions and constraints. These interdependencies cannot be directly represented as input-output relationships. Information about these interdependencies needs to be identified and utilized to develop new approaches to modeling SoS. This

dissertation extends the intrinsic approach [Haimes 2012] to modeling subsystem interdependencies through shared state variables. A two-level decentralized control structure is also developed to integrate and coordinate objectives of various subsystems.

Probabilistic Risk Assessment (PRA) is a systematic methodology to evaluate risks associated with a complicated engineered technological system and has been successfully performed at all phases of the life cycle from concept definition and pre-design through safe removal from operation. Kaplan and Garrick [1981] introduced the Theory of Scenario Structuring (TSS) and within it the triplet questions in the risk assessment process: 1) what can go wrong? 2) what is the likelihood that it would go wrong?; and 3) what are the consequences?. Kaplan, Haimes and Garrick [2001] subsequently modified TSS by stating that the set of all scenarios cannot be a complete set. Risk assessment methodologies based on the concept of PRA systematically develop risk scenarios of the initiating events and attempt to quantify the probability and consequences for each scenario. This approach is used in event trees and event sequence diagrams. Principal examples of inductive approaches include Failure Mode and Effect Analysis (FMEA) [Stamatis, 2003], Failure Mode Effect and Criticality Analysis (FMECA) [Bouti & Kadi, 1994], and Preliminary Hazard Analysis (PHA) [Vincoli, 2006].

However, applying this inductive risk assessment approach to complex SoS faces new challenges. The term "emergent" denotes that the forced changes to SoS are usually dynamic, evolving, and possibly unexpected. Investigations of several accidents of complex systems such as Three Mile Island Accident [Perrow 1984] show that the causes of complex system failure usually include multiple component failure and their unexpected interactions. For example, multiple initiating events posing no risk to each

individual subsystem may have possibly unknown complex interdependencies and causal relationships thus cause significant adverse consequences at the SoS level. In these cases, it is practically impossible to enumerate an all-inclusive set of potential risk scenarios due to both the large number of scenarios resulted from the combination of multiple initiating events and the analyst's limited knowledge about the subsystem interdependency. An initiating event for an emergent forced change might also not be directly observable. Besides that, there might not be any historical data or expert judgment to estimate the probability of risk scenarios for emergent, large scale complex SoS, and to support efficient allocation of limited risk management resources. In addition to the above issues, short of actually observing triggering changes in the state space of the system by collecting and analyzing information and other evidence, the static and passive approach lacks the capability (i) to track and monitor different risk scenarios over time, and (ii) to forecast indicators and warnings on evolving and emergent forced changes. Risk assessment depending solely on inductive and static methods might fail to detect emergent risks to SoS, thus result in inefficient allocation of risk management resources.

We posit that a comprehensive risk assessment of complex infrastructure SoS should employ both inductive and deductive approaches to supplement and complement each other. The basic difference between deductive and inductive methods is the direction of the analysis. Inductive method is the appropriate analysis to carry out if a given set of initiating causes is identified and the goal is to determine the resulting consequences. Deductive method is the appropriate analysis to carry out if a given undesired event is defined and the goal is to determine its basic causes. The inductive approach is useful in assuring that the analysis is broad enough to encompass all possible scenarios while the

deductive approach has the benefit of focusing the analysis on the undesired event. Principal examples of inductive approaches include Fault Tree Analysis (FTA) [Vesely *et al*. 1981].

Deductive risk assessment of complex infrastructure SoS requires an understanding of why and how risk propagates through the system and results in significant adverse consequences. These significant adverse consequences usually manifest themselves in the form of system failures or accidents. In this case, the question of why complex SoS fails and how to detect and evaluate signs prior to system failures becomes the focus of a deductive risk assessment. Identifying specific failure modes and their failure mechanisms is a key step to extend current risk assessment tools to analyze complex infrastructure SoS.

Effective risk assessment and management of infrastructure systems depend on the understanding of their failure mechanisms, and designing mitigation strategies to manage these risks. Failure analysis tools based on reliability theory, such as sequential accident models (event-based accident models) and Epidemiological models [Qureshi, 2007] [Hollnagel, 2002] work well for accidents caused by failures of physical components or human errors in relatively simple systems. However, they are limited in their capability to explain accident causation in the more complex systems that were developed in the last half of the 20th century [Hollnagel 2004] [Lundberg *et al.* 2009]. Several theories have been proposed to understand failures and accidents in complex systems. Based on systems theory, systemic accident models endeavor to describe the characteristic performance on the level of the system as a whole, rather than on the level of specific cause-effect “mechanisms” or even epidemiological factors [Hollnagel 2004]. Two

notable systemic modeling approaches – Rasmussen's (1997) hierarchical socio-technical framework and Leveson's [2004] Systems-Theoretic Accident Model and Processes (STAMP) model – endeavor to model the dynamics of complex sociotechnical systems. However, most of these theories are qualitative in nature, thus unable to provide a rigorous system-level analysis.

Risk analysis is about the future. Pro-active risk analysis of emergent forced changes calls for a continuous process of designing a data-collection mechanism, developing metrics with which to measure changes in the system, assessing whether observed changes are sufficiently significant, and determining criteria for actions – all are requisites for effective risk modeling, assessment, and management for future emergent forced changes. With the advancements in sensing, communication, and information processing technology, automated detecting and monitoring devices generate high volume of information flow regarding past and current states of the system [Ko & Ni 2005]. This trend in technology provides a foundation for a pro-active risk assessment tool utilizing information to detect signs prior to system failure, through the observation of precursors and the changes in the relationship among different state variables in the state space of the system. Precursors are important signs prior to system failures and thus deserve our further investigations.

Precursor analysis identifies factors that increase the likelihood of and detects indicators and warnings of possible future system failures. Precursors can be used to determine whether adversarial events are either occurring or expected, and provide opportunities for decision makers to take preparedness and response actions to avoid projected adverse consequences. Analyzing precursors to system failures based on

available information and evidence is a deductive and pro-active risk assessment approach thus an integral part in risk assessment of complex infrastructure SoS.

Precursor analysis, which has been widely adopted in nuclear power, aviation, and healthcare industries, has been proven to be an effective approach to identify potential risk factors leading to accidents and other adverse consequences [National Research Council, 2004]. Although literatures across different disciplines have discussed the potential use of precursors to anticipate system failures and accidents, there lacks a formal process of precursor analysis based on systems theory and a rigid, systemic, and justifiable framework. In most of the literature, precursors are identified solely based on accident or near miss incident reports, analysis of available data, and expert knowledge. These methods for precursor analysis neither account for the internal causal relationships nor capture the essence of the system dynamics. Without a system model capable of describing the mechanism of system failure, it is very difficult to identify and evaluate precursors and provide useful insights for risk management.

Another major issue in precursor analysis is the frequent discrepancies between pre- and post-accident risk assessment using precursors, where in the former case the effectiveness of using precursors are quite limited. This perception may be the result of hindsight bias [Fischhoff, 1975; Hawkins and Hastie, 1990] that is, after an accident, individuals often believe that the event should have been considered highly likely, if not inevitable, by those who observed the precursors prior to the accident. We believe that the following three major facts contribute to the existence of hindsight bias:

i. Precursors are usually weak signals to forecast system failure, and in many cases the likelihood of system failure before and after the observation of a

specific precursor is not objectively quantified. Response actions are not justified based on such vague indicators.

ii. There are one-to-many relationships between a precursor and various failure modes of the system. The observation of a specific precursor may increase the likelihood of multiple failure modes simultaneously thus not necessarily increasing the situational awareness of the decision maker.

iii. Uncertainties in precursor detection and prediction are not well accounted in a non-model-based and informal analysis.

Addressing these problems in precursor analysis would require a systemic and quantitative approach to identify essential system building components, understand complex system behaviors and failure mechanisms, quantify uncertainties in detection and prediction, and integrate information from multiple precursors.

In summary, complex infrastructure SoS possess some specific characteristics that distinguish them from conventional systems. New theories and methodologies are required when extending current theories and methodologies of risk analysis to the complex infrastructure SoS.

Section 1.3 provides an overview of the proposed approaches to these issues in the risk analysis of complex infrastructure SoS.

1.3. Proposed Approaches

Risk analysts of complex SoS must understand and quantify the interdependencies among the subsystems, be able to foresee the emergent forced changes from within or from outside the system, and evaluate their impacts to the system.

1.3.1 Modeling Complex Infrastructure Systems of Systems

The proposed approaches first address the modeling of complex infrastructure SoS. Modeling the entire complex SoS, including all its components, relationships, functions and behaviors is a daunting task involving a continuous process of learning, discovery, modification, and validation. The complexity of SoS is characterized by the highly interconnected and interdependent subsystems, which introduce significant modeling challenges for the systems analysts. Decomposition is a common approach to modeling large-scale systems. When a large-scale system is decomposed into its subsystems and sub-subsystems, the interconnectedness and interdependencies among these subsystems must be preserved. In modeling the intricate complex interdependencies among the myriad components and subsystems that constitute SoS, it is a challenging task to quantify their causal intra- and interconnected and interdependent relationships [Haimes, 2007, 2008, 2012]. In the extrinsic modeling approach, these interconnections and interdependencies have been modeled as input-output relationships among subsystems.

State variables play a central role in modeling systems and their interdependencies. Shared state variables are defined as the state variables that are common to two or more subsystems. In this dissertation, the "intrinsic" modeling approach [Haimes, 2012] is applied to the SoS where subsystem interdependencies are represented by shared state variables and coupled decision variables in a state space model. Methods to decompose and coordinate the subsystems when they share state variables are discussed.

A model should be as simple as possible but as complex as required. A large-scale physical infrastructure SoS usually consists of multiple subsystems, sub-subsystems, units, and numerous components. The finer the granularity of the model the more

complex it will be, with the increase in the number of variables and equations, the time and cost to develop the model, and the difficulty in finding the solutions. The level and granularity of the model need to be determined properly for a model to answer specific questions. For example, a model of bridge infrastructure systems built from detailed physical models for each bridge element may not be very suitable for a decision maker focusing on strategic policies. However, without understanding the behaviors of each component of the system, a higher-level model cannot be built from a bottom-up approach.

Meta-modeling connotes a framework that builds on systems-based theory and methodology that enables modelers to relate, coordinate, and integrate sub-models of multiple models of subsystems of SoS for the purpose of better understanding and modeling the SoS as a whole. Meta-models are usually constructed using a data-driven, bottom-up approach. The exact, inner working of the system is not assumed to be known or even understood, solely the input-output behaviors is important to the upper level models [Queipo *et al*, 2005]. System identification [Ljung, 2010] is one of the major approaches to develop meta-models.

In this dissertation, two levels of meta-models are used to model different aspects of complex SoS. At a lower level, meta-models are used to extract the functional components of a subsystem from their constitutive physical components, such that the behavior of a subsystem can be modeled as the result of interconnected and interdependent functions instead of physical components. This meta-modeling approach enables a modeler to model and analyze the functions of a complicated subsystem without understanding how these functions are realized by various physical components.

This abstraction of functions is useful for a failure model to capture different failure modes of the system and their conditions. At a higher level, meta-models are used to model the behaviors of subsystems based on a decomposition method. Subsystems having shared states can be converted into subsystems with independent states and inputs from other related subsystems. It enables a coordinator at SoS level to coordinate multiple subsystems without knowing the internal states, structures, and working mechanism of each subsystem. This level of meta-modeling has important realistic implication because of the operational and managerial independence property of the organizations in SoS.

Among various systems engineering theories and methodologies, this proposed research in particular builds on the centrality of the states of the system in modeling and in risk analysis. State variables serve as the key to define, analyze, and detect emergent force changes to the system. As the state variables of a complex SoS are distributed across different subsystems, this approach explores the intrinsic interdependencies and interactions between and among subsystems through the understanding of all shared and unshared state variables.

1.3.2 Systemic Risks in Complex Systems of Systems

A fundamental question to be answered for any deductive risk assessment of complex infrastructure SoS is why these systems fail. A clear understanding of the failure mechanisms of complex SoS provides a theoretical base for the identification and detection of precursors to system failure. Because the structures and purposes of complex infrastructure SoS as well as their definitions of system failure vary significantly across different types of systems, literature on failure mechanism tends to be very specific from

case to case and there lacks a general theory on the failure mechanism of complex systems. However, the literature on systemic accident models shed some light on how complex systems fail, as system failures usually manifest themselves as or lead to accidents. Systems-Theoretic Accident Model and Processes (STAMP) [Leveson 2004] treats accidents as a dynamic control problem instead of a reliability problem and claims that accidents result from lack of enforcement of safety constraints in system design and operations, and (at least) one of the goals in system design and operation should ensure the behavior of the components and systems as a whole to ensure safety constraints. This is consistent with the theory that constraints and objectives are exchangeable in multi-objective decisionmaking problems [Chankong and Haimes 1983]. Thus, the enforcement of safety constraints should be treated as one of the most important objectives of the safe operation of any physical infrastructure system. Hierarchical Socio-technical Framework [Rasmussen 1997] describes the causation of accidents as the migration of systems and organizations toward states of high risk under cost and productivity pressures in an aggressive, competitive environment and normal variation in organizational behavior across the boundary of safety regulations.

Based on the theory of systemic accident models [Leveson 2004] [Rasmussen 1997], this dissertation aims to identify internal systemic risks specific to complex SoS through: 1) identification of system safety constraints through a control theory perspective; and 2) exploration of unique characteristics of complex SoS and discovery of how the identified safety constraints can be violated due to these characteristics. Complex SoS are characterized by: 1) nonlinear system dynamics; 2) multiple stakeholders and decision makers; 3) multiple goals and objectives; 4) multiple interconnected and interdependent

subsystems; 5) unknown interactions between subsystems; and 6) emergent and adaptive system behaviors. An analysis of a multi-objective sequential decision process of a nonlinear system reveals some inherent systemic risks within the subsystems of SoS, as one of the potential failure modes of the SoS, when the decision maker's preference on system safety against other objectives is reduced.

1.3.3 Risk Assessment Using Precursors

The ultimate purpose and efficacy of risk assessment are to envision, foresee, and predict emergent forced changes based on the capability of the human imagination, the availability of evidence, and the predictions of modeling tools. From a broader perspective, the risk assessment process, and to a limited extent the risk management process, are supported by envisioning, discovering, and tracking emergent forced changes.

Among various approaches to risk assessment of system failures, this dissertation proposes a pro-active, dynamic anticipatory analysis tool based on precursors, as an extension to the static, passive scenario-based risk assessment approach. Precursors are signals prior to state change or system failure. If precursors to system failures are detected and evaluated in an objective and credible way, it can be reasonably assumed that decision makers will gain more situational awareness about which risk is impending and how much time is available to response, thus lead to more efficient allocation of risk management resources. The proposed precursor analysis tool evaluates the likelihood of different competing failure modes through observing and monitoring the changes in the states of the system, so that decision makers can make more timely and informed

decisions in response to emergent forced changes to the system and prevent severe consequences in the future.

1.4. Goal and Objectives

The goal of this dissertation is to develop theories and methodologies to model and analyze risks to complex infrastructure SoS, identify and detect emergent forced changes through precursor analysis based on the knowledge of system structure, dynamics, and failure mechanism, and provide insights on the solutions to the issues stated in section 1.2. The rest of the chapters are organized in the following order:

Chapter 2 reviews current state-of-the-art paradigms on risk analysis of complex infrastructure SoS and different programs of precursor analysis.

Chapter 3 serves as a theoretical foundation to analyze the failure of complex infrastructure SoS caused by subsystem interdependencies through shared states. An intrinsic modeling approach is introduced for modeling, decomposing, and quantifying inter-subsystem interdependency, and followed by discussions of subsystem coordination which aims to reduce the unexpected impacts from interdependent subsystems.

Chapter 4 explores a new perspective to understand a unique failure mechanism of complex SoS. Using the decision maker's preference on the safety objective as a slowly changing parameter of a dynamic multi-objective decision process, the analysis shows that the system becomes more susceptible to small perturbations caused by inter-subsystem interdependencies when the preference is below a certain level. Given a known system model, it provides a quantitative way to identify the safe operation boundary of the system and its resilience to external perturbations caused by interdependencies through shared states.

Chapter 5 discusses how precursors to system failure can be identified, prioritized, detected, and evaluated against their likelihood and urgency to cause system failure. A practical yet systemic way to identify precursors to system failure is discussed, in which a meta-modeling approach is used to identify system functional components, control structure, and safety constraints and Hierarchical Holographic Modeling (HHM) is used to develop and organize major failure modes and failure scenarios. Each safety constraint is further decomposed and quantified by its corresponding essential state variables and their relationships. Multiple detected precursors then can be integrated to track the likelihood of competing failure modes and improve situation awareness of the decision maker.

Chapter 6 performs a case study on a bridge infrastructure SoS where it serves two purposes. First, it explains the theories and methodologies presented in Chapters 3 – 5 using a real-world example. Second, the bridge infrastructure SoS is studied and examined to validate the assumptions and conditions used in the theories discussed in the early chapters, and it applies the precursor based risk assessment tool with bridge data to demonstrate the applicability and effectiveness of the proposed approach in a real problem setting.

Chapter 7 summarizes the findings and provides insights to design better precursor monitoring and warning systems. Some future research questions and potential directions are discussed.

This dissertation provides useful insights and guidelines to the identified issues in the risk assessment of complex infrastructure SoS through the above procedures. Each of these tasks will be discussed in detail in the following chapters.

2. Literature Review

This chapter provides a comprehensive review of the state-of-the-art in risk assessment and management of complex infrastructure systems of systems. It is organized in the following order: Section 2.1 introduces the concept and properties of complex SoS; section 2.2 discusses the causes of complex system failure; section 3.3 reviews in detail different types of accident models that can be used for risk analysis of complex engineering systems; and section 2.4 outlines current practice in precursor and early warning system design.

2.1. Complex System, Systems of Systems, and Systems Engineering Approaches

Complex systems science is a new approach to science that studies how relationships between parts give rise to the collective behaviors of a system and how the system interacts and forms relationships with its environment through variation, interaction, and selection. The hallmarks of complex systems are interaction, adaptation, self-organization and emergence, and the common characteristic of all complex systems is that they display

organization without any external organizing principle being applied [Ottino, 2004]. Figure 2-1 shows the major factors contributing to complexity.

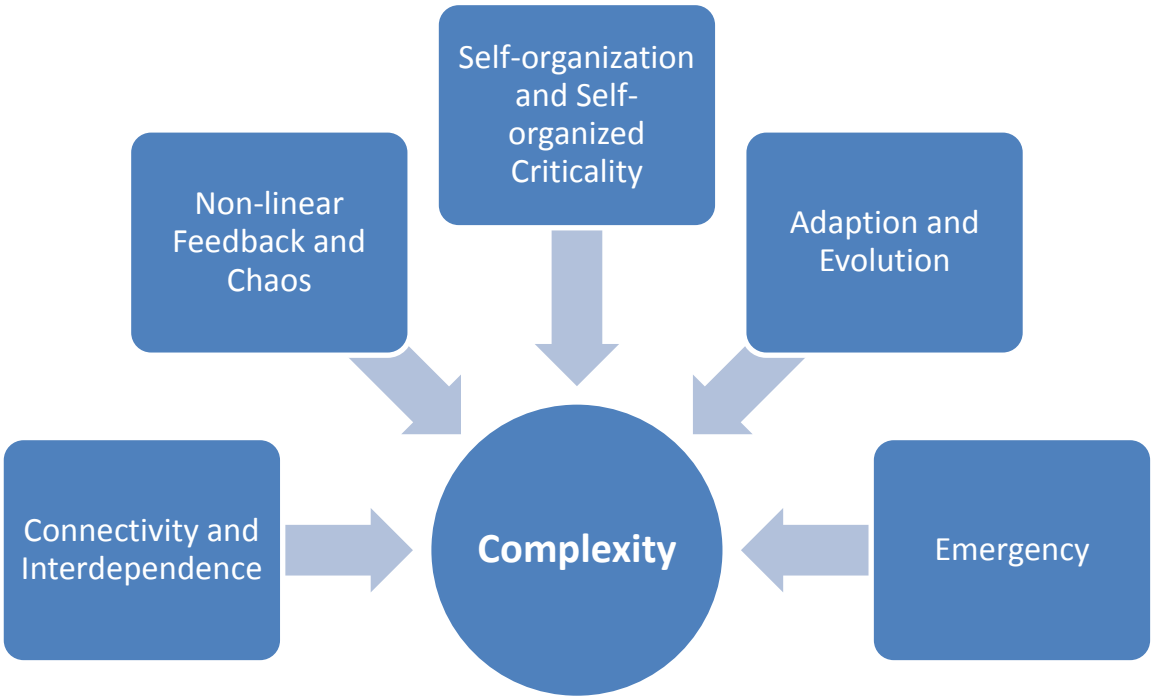

FIGURE 2-1. MAJOR FACTORS CONTRIBUTING TO COMPLEXITY

The mathematical techniques used in complex system studies include nonlinear dynamics (differential and difference equations and time series analysis), graph and network theory, agent-based modeling and simulation (ABMS) among many others [Ottino, 2003] [Barabási, 1999], as depicted in Figure 2-2. Shalizi [2006] also reviews the main methods and techniques of complex systems science, which include tools for analyzing data, constructing and evaluating models, and measuring complexity. Chang and Harrington [2005] provide a comprehensive description of agent-based models of

organizations. Amaral and Ottino [2004] describe network theory and its importance in augmenting the framework for the quantitative study of complex systems. Lloyd and Lloyd [2003] present a general method for modeling complex systems in terms of flows of information. Page [1999] discusses detailed, robust computational models such as SWARM Platforms. For the challenges from complex systems engineering, Johnson [2006] provides a comprehensive review of emergent properties and how they affect the engineering of complex systems. Bar-Yam [2003] reviews the lessons learned from problems with systems engineering over the past couple of decades and an evolutionary paradigm for complex systems engineering. Highly Optimized Tolerance (HOT) has also been developed to generate power law behaviors in complex systems which explain some characteristics of complex systems. Some common features of HOT include 1) high efficiency, performance, and robustness to designed-for uncertainties; and 2) hypersensitivity to design flaws and unanticipated perturbations [Carlson & Doyle, 1999, 2000].

The subject of large- and multi-scale complex systems has been on the agenda of researchers for over half a century [Wiener 1948, Bertalanffy 1968, Sage 1977, 1992, 1995, Blauberg et al. 1977, Haimes 1977, Sage & Rousse 1999]. Large-scale complex infrastructure systems are commonly composed of interconnected and intra- and interdependent subsystems, which in their essence constitute a system of systems; each is characterized by a hierarchy of interacting and networked components with multiple functions, operations, efficiencies and costs [Haimes, 1977, 1981, 2008, 2012]. Sage and Biemer [2007] argue that no universally accepted definition of a SoS is available at this

time. Sage and Cuppan [2001] build on the following five properties of SoS suggested by Maier [1998]:

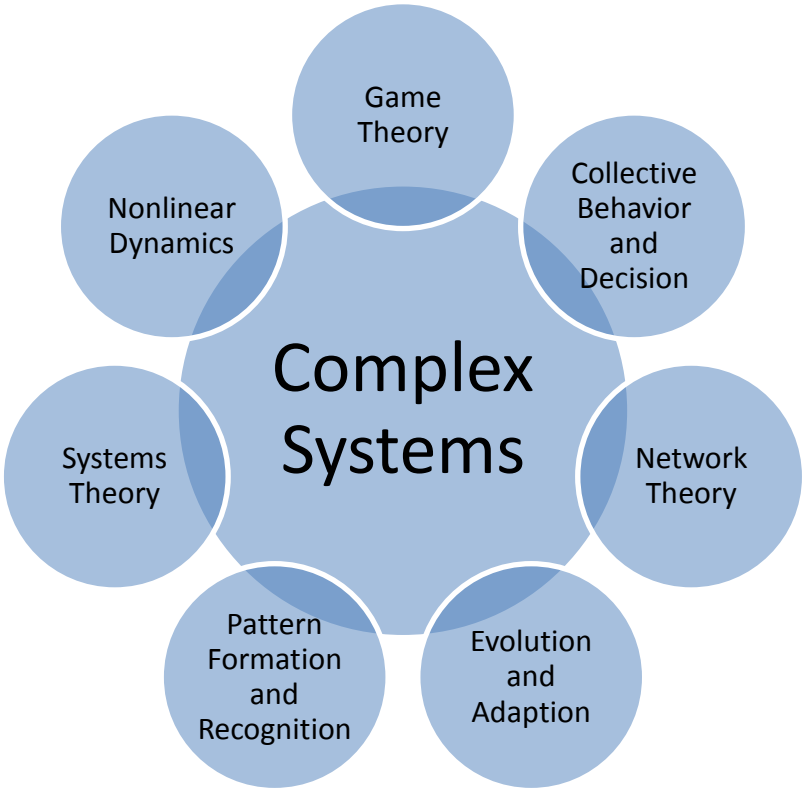

FIGURE 2-2. COMMONLY USED RESEARCH APPROACHES TO COMPLEX SYSTEMS

- Operational Independence of the Individual Systems. A system of systems is composed of systems that are independent and useful in their own right.
- Managerial Independence of the Systems. The component systems not only can operate independently, they generally do operate independently to achieve an intended purpose.
- Geographic Distribution. Geographic dispersion of component systems is often large. Often, these systems can readily exchange only information and knowledge with one another, and not substantial quantities of physical mass or energy.

- Emergent Behavior. The system of systems performs functions and carries out purposes that do not reside in any component system.
- Evolutionary Development. A system of systems is never fully formed or complete. Development of these systems is evolutionary over time and with structure, function and purpose added, removed, and modified as experience with the system grows and evolves over time.

Thus recognition of SoS is widespread in the recent decade in modeling for disciplines including cyber, finance, infrastructure, healthcare, military, environment, and many others.

Systems of systems are composed of various physical and engineered elements, users, decision makers, and the environment in which they operate. DeLaurentis [2008], Lewe et al. [2004], Parker [2010], and Aktan & Faust [2003] suggest that SoS problems require a new modeling paradigm that would account for the multiplicity of stakeholders, objectives, interdependencies and emergent outcomes. DeLaurentis & Callaway [2004] discuss the need to focus the modeling effort on the system of systems interdependencies, and they suggest that the "evaluation of an individual entity at its own level is of less importance than how it affects the higher level organization for which it is a member." Similarly, Thissen & Herder [2009] claim that the "efforts to increase understanding at the overall system of systems level are much needed, in view of the fact that the key performance indicators of complex infrastructure systems are in the end determined by the interplay of most, if not all the component systems." A system of systems modeling approach has been called for by civil engineering researchers and practitioners, who suggest that integrated modeling of large-scale infrastructure SoS, inclusive of their

engineered, human, and natural elements has been unsuccessful thus far [Aktan & Faust 2003].

While much of the academic writing has focused on defining what systems of systems are, less work has been done to develop modeling approaches for studying and managing the risks of these systems. Among which, Sage [2003] presents an overview and partial taxonomy of the diversity of categorizations that may be used to describe various conflict and risk situations that involve a system of systems. Often, these situations emerge from behavior of many independent agents who attempt to achieve both individual objectives and objectives of a larger organizational unit. Sage partitioned risk and conflict management into two interrelated components: risk and conflict program planning and risk and conflict abatement. He suggested that in risk and conflict program planning, we forecast and assess the potential for risk and conflicts. This involves formulation, assessment, and interpretation steps. In risk and conflict abatement, we implement the selected abatement tactics such that we are able to monitor the situation such that we can detect an impending risk and conflict situation, diagnose the cause of the situation, and correct it through selection of an appropriate risk and conflict abatement alternative.

Despite the documented efforts to improve the modeling of complex and large-scale systems, implementation efforts of such modeling methodologies remain scarce. The lack of quantitative modeling of complex SoS leads to a limited awareness of the types of analytical risk analysis tools that could be used for these systems.

2.2. System Accidents Caused by Interdependencies

Perrow [1984] in his book "Normal Accidents" provided a structural analysis of risky systems and first introduced the concept of system accidents in complex engineering systems. Investigations of several accidents of complex systems such as Three Mile Island Accident show that the causes of complex system failure usually include multiple component failure and their (unexpected) interactions. He distinguishes two general types of failure modes in complex systems: component failure accidents and system accidents. Component failure accidents involve one or more component failures (part, unit, or subsystem) that are linked in an anticipated sequence, while system accidents involve the unanticipated interaction of multiple failures. A system accident must have multiple failures, and they are likely to be in reasonably independent units or subsystems. It is not the source of the accident that distinguishes the two types, since both start with component failures; it is the presence or not of multiple failures that interact in unanticipated ways.

Contemporary literatures in risk analysis of complex systems lack analytical tools to address this issue of these unknowns. Some conceptual and qualitative approaches to forecast system failure through monitoring the states or stability of the system have been discussed, for example, Fisher [2011] in his book "Crashes, Crises, and Calamities" indicated several common early warning signs prior to an imminent catastrophes of complex systems, such as the existence of runaway processes including positive feedback loops, chain reactions, and the domino effect.

2.3. Accident Models for Complex Engineering Systems

Accident models provide a conceptualization of the characteristics of the accident, which typically show the relation between causes and effects. They explain why accidents occur, and are used as techniques for: risk assessment during system development, and post hoc accident analysis to study the causes of the occurrence of an accident.

One of the earliest accident causation models is the Domino theory proposed by Heinrich in the 1940's [Ferry 1988], which describes an accident as a chain of discrete events which occur in a particular temporal order. This theory belongs to the class of sequential accident models or event-based accident models, which underlie most accident models such as Failure Modes and Effects Analysis, Fault Tree Analysis, Event Tree Analysis, and Cause-Consequence Analysis [Leveson 1991]. These models work well for losses caused by failures of physical components or human errors in relatively simple systems. However, they are limited in their capability to explain accident causation in the more complex systems that were developed in the last half of the 20th century [Hollnagel 2004].

In the 1980s, a new class of epidemiological accident models endeavored to explain accident causation in complex systems. Epidemiological models regard events leading to accidents as analogous to the spreading of a disease, i.e., as the outcome of a combination of factors, some manifest and some latent, that happen to exist together in space and time. Reason's [1990, 1997] Swiss cheese model of defenses is a major contribution to this class of models, and has greatly influenced the understanding of accidents by highlighting the relationship between latent and immediate causes of accidents.

Sequential and epidemiological accident models are inadequate to capture the dynamics and nonlinear interactions between system components in complex socio-technical systems. New accident models, based on systems theory, classified as systemic accident models, endeavor to describe the characteristic performance on the level of the system as a whole, rather than on the level of specific cause-effect "mechanisms" or even epidemiological factors [Hollnagel 2004]. A major difference between systemic accident models and sequential/epidemiological accident models is that systemic accident models describe an accident process as a complex and interconnected network of events while the latter describes it as a simple cause-effect chain of events. Two notable systemic modeling approaches – Rasmussen's [1997] hierarchical socio-technical framework and Leveson's [2004] STAMP model – endeavor to model the dynamics of complex socio-technical systems.

2.4. Precursor and Early Warning System

Early warning is the provision of timely and effective information, through identified institutions, that allows individuals exposed to hazard to take action to avoid or reduce their risk and prepare for effective response [UNEP 2012]. The basic idea behind early warning is that the earlier and more accurately we are able to predict short and long-term potential risks associated with natural and human-induced hazards, the more likely we will be able to manage and mitigate disasters' impact on society, economies, and environment.

The early warning system is the integration of four main elements [UNISDR, 2005]:

- Risk Knowledge: Risk assessment provides essential information to set priorities for mitigation and prevention strategies and designing early warning systems.

- Monitoring and Predicting: Systems with monitoring and predicting capabilities provide timely estimates of the potential risk faced by communities, economies and the environment.
- Disseminating Information: Communication systems are needed for delivering warning messages to the potentially affected locations to alert local and regional governmental agencies. The messages need to be reliable, synthetic and simple to be understood by authorities and public.
- Response: Coordination, good governance and appropriate action plans are a key point in effective early warning. Likewise, public awareness and education are critical aspects of disaster mitigation.

The main goal of early warning systems is to take action to protect or reduce loss of life or to mitigate damage and economic loss, before the disaster occurs. Monitoring and predicting provides the input information for the early warning process that needs to be disseminated. It is essential to note that "predictions are not useful, however, unless they are translated into a warning and action plan the public can understand and unless the information reaches the public in a timely manner" [Glantz, 2003].

Paté-Cornell [1986] presented a method that allows probabilistic evaluation and optimization of warning systems, and comparison of their performance and cost-effectiveness with those of other means of risk management. Lakats and Paté-Cornell [2004] describe an analytical framework focused on organization performance, based on decision analysis and probability, to design and optimize such a warning system from a management perspective.

The National Research Council workshop [NRC, 2004] definition of an accident precursor is any event or group of events that must occur for an accident to occur in a given scenario. Many other definitions exist. For example, a precursor is an event that has some, but not all, of the ingredients of a more undesirable situation [Corcoran, 2003a, b]. However in practice, precursors are more commonly referred to previously undiscovered conditions, which if occurred, will greatly increase the likelihood of an undesirable scenario. Conditions that are necessary but contribute much less to the occurrence of an undesirable scenario are usually not regarded as precursors.

Most precursor analysis approaches fall into two general categories of precursor analysis framework. In healthcare and aviation industry, precursors leading to accidents are learned through the observation and reporting of near misses, close calls, or even accidents. This is because when the system itself is too complex, it is usually characterized by complex human-machine interaction, such that no model is able to provide a comprehensive analysis of all failure scenarios. This precursor analysis process uses a descriptive approach to understand the causes of accidents and is depicted in Figure 2-3. Figure 2-3. Descriptive precursor analysisCurrent precursor analysis methods belonging to this category include: Aviation Safety Action Programs (ASAPs), Aviation Safety Reporting System (ASRS), Applied strategies for Improving Patient Safety (ASIPS), and Patient Safety Reporting System (PSRS). The drawback of this approach is that the time and cost of learning is prohibitive in some cases due to the lack of a pro-active approach to discover precursors that haven't been experienced by the organization.

In other industries such as nuclear power industry, because the consequences of missing potential precursors to accidents are so high, a reliability model-based pro-active

approach that systematically explores precursors in the system is used. This prescriptive approach is shown in Figure 2-4.

Although learning from near misses is still an integrated part of this process, the focus is put more on updating and improve the system model so that more related precursors can be exploited, instead of merely identifying individual precursors. The Accident Sequence Precursor (ASP) Program in the nuclear industry belongs to this category of precursor analysis. The drawback of this approach is that unknown or unanticipated risk-significant precursors may not be identified and reported due to the lack of knowledge about the complex behaviors of the system.

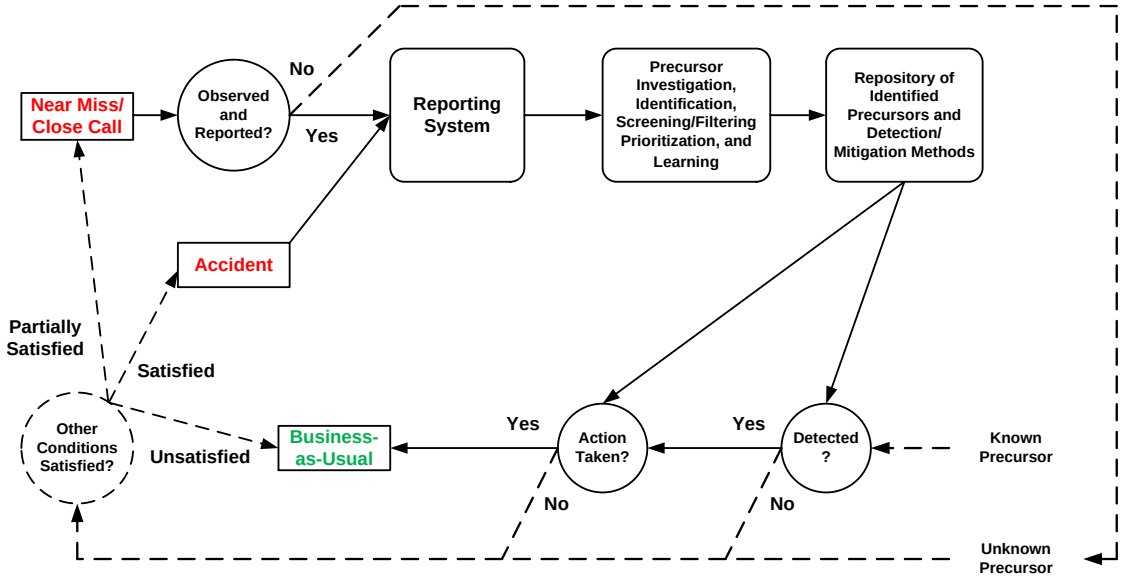

FIGURE 2-3. DESCRIPTIVE PRECURSOR ANALYSIS

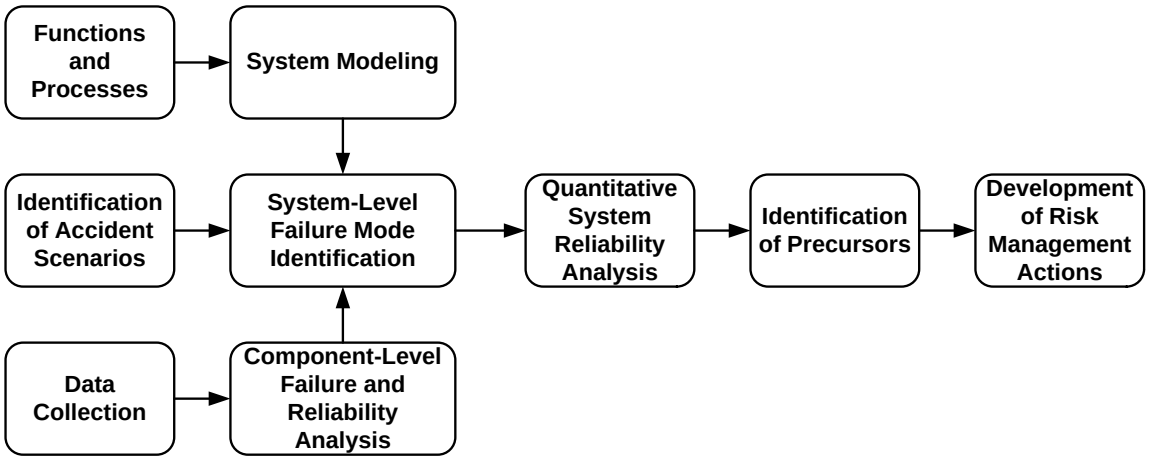

FIGURE 2-4. PRESCRIPTIVE PRECURSOR ANALYSIS

3. Decomposition and Coordination of Systems with Shared States

This chapter discusses methods to decompose subsystems with shared state variables and coordinate individual subsystems to achieve the overall objective of SoS through a two-level hierarchical control structure. The decomposition method plays a fundamental role in understanding systemic risks caused by subsystem interdependencies, which will be discussed in Chapter 4. Section 3.1 provides an overview of the needs for decomposition and coordination in modeling and managing SoS. The decomposition method is discussed in section 3.2. Section 3.3 extends the two-level non-feasible method to a state-space model for subsystem coordination.

3.1. Background

Complex SoS are usually of large scale. No single model is capable of representing the multiple perspectives of a complex SoS, thus a realistic model of SoS must consist of multiple subsystem models and their relationships. Any modeling approach failing to consider subsystem interdependencies would fail to uncover intricate complex interactions and causal relationships among the myriad components and subsystems that

constitute SoS. Modeling the entire complex SoS, including all its components, relationships, functions and behaviors requires a continuous process of learning, discovery, modification, and validation. It is a challenging task to discover and quantify all the causal intra- and interconnected and interdependent relationships among subsystems that constitute SoS.

The Phantom System Model (PSM)-based intrinsic modeling approach [Haimes, 2007, 2012] posits that some specific commonalities, interdependencies, interconnectedness, or other relationships must exist between and among any two subsystems within any SoS. The essence of each subsystem can be represented by a finite number of essential state variables. For a properly defined system of systems, any interconnected subsystem will have at least one (typically more) essential state variable(s) and objective(s) shared with at least one other subsystem. Shared state variables are defined as state variables common to two or more subsystems' models, thus they play a central role in modeling systems and their interdependencies.

With subsystems sharing state variables, both a centralized and a decentralized approach can be used to analyze and manage the SoS. A centralized approach combines these subsystems as an integrated system and treats the shared state variables as internal state variables. This approach is usually efficient and achieves better system performance, where all needed information is readily available to the central controller. However, issues arise in managing complex infrastructure SoS using a centralized approach due to the operational and managerial independence and geographic distribution [Maier 1998] characteristics of SoS. First, it is likely that no central modeler who has the knowledge about all the subsystems, or a central controller who has the authority to set

goals for and control of the operations of all subsystems. Second, the complexity of the system model and solutions increase with the scale of the overall system. Furthermore, it is not the practice of SoS organizations to share their privileged information with others, which is a key requirement for the centralized control. Thus, decomposition of SoS and a decentralized approach to subsystem coordination and control are more realistic. The decentralized approach treats each subsystem independently at a lower level of the hierarchy and coordinates them at a higher level, and only limited information is exchanged between subsystems and the coordinator.

3.2. Decomposition of Systems with Shared State Variables

A more detailed categorization of shared state variables in a state-space model is discussed first. Linear state space models are one of the most commonly used mathematical models to represent physical systems. Despite their simplicity, they can reasonably approximate the behaviors of many real systems. To simplify our discussion, we will use linear state space models to demonstrate the approach.

A linear dynamic (deterministic) state space model has the form

EQUATION 3-1

$$S(k+1) = AS(k) + BU(k)$$

$$Y(k) = CS(k) + DU(k)$$

$$k = 0, \ldots T-1$$

where S is the vector of state variables, U is the vector of control/input variables, and Y is the vector of output variables. This model describes the system behavior for time period from 0 to $T-1$ and k is the index of time. Matrices $A, B, C,$ and D are system

parameters. In a simple case where the system described in Equation 3-1 consists of two subsystems, it is obvious that matrix A determines whether two subsystems share state variables, matrix B and D determine whether two subsystems share decision variables, and matrix C determines whether the output of the two subsystem are coupled through shared state variables. Here we limit our discussion to different forms of matrix A and its relation to shared state variables.

In a special case where matrix A has the form of or can be transformed to a diagonal matrix with n sub-matrices as the diagonal elements, such as $A = \begin{bmatrix} A_{11} & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_{nn} \end{bmatrix}$, the overall system can be decomposed into n independent subsystems, without sharing any state variables.

For each subsystem, we have $S_i(k+1) = A_{ii}S_i(k) + B_{ii}U_i(k)$, for $i = 1, \dots, n$. (assuming matrix B is also diagonal.)

However, the transformation of matrix A into a diagonal matrix is not guaranteed in most cases. Here we discuss four different cases shown in Figure 3-1.

Case 1 $A = \begin{bmatrix} A_{11} & 0 & 0 & A_{1n} \\ 0 & A_{22} & 0 & A_{2n} \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_{nn} \end{bmatrix}$

Case 2 $A = \begin{bmatrix} A_{11} & 0 & 0 & 0 \\ 0 & A_{22} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ A_{n1} & A_{n2} & 0 & A_{nn} \end{bmatrix}$

Case 3 $A = \begin{bmatrix} A_{11} & 0 & 0 & 0 \\ 0 & A_{22} & 0 & A_{2n} \\ 0 & 0 & \ddots & 0 \\ A_{n1} & 0 & 0 & A_{nn} \end{bmatrix}$

Case 4 $A = \begin{bmatrix} A_{11} & 0 & 0 & A_{1n} \\ 0 & A_{22} & 0 & A_{2n} \\ 0 & 0 & \ddots & 0 \\ A_{n1} & A_{n2} & 0 & A_{nn} \end{bmatrix}$

FIGURE 3-1. FOUR DIFFERENT CASES OF MATRIX A, FROM LEFT TO RIGHT

In case 1, both state variables S_1 and S_2 depend on state variables S_n but S_n does not depend on S_1 and S_2. In this case, there is no way to completely separate subsystem 1 and 2, and state variables S_n are shared by subsystem 1 and 2 as a source of information. However, there is no interaction between subsystem 1 and 2 although they share state variables. This means that the change of states in subsystem 1 has no impacts on the states in subsystem 2.

In case 2, state variables S_n depend on state variables S_1 and S_2 but S_1 and S_2 do not depend on S_n. Similar to case 1, state variables S_n are shared by subsystem 1 and 2 as a drain of information, and there is no interaction between subsystem 1 and 2.

In case 3, state variables S_2 depend on state variables S_n and S_n depend on S_1, and state variables S_n are shared. There is one-way interaction from subsystem 1 to 2, which means that the change of states in subsystem 1 has impacts on the states in subsystem 2, but not in the reversed direction.

In case 4, both state variables S_1 and S_2 depend on state variables S_n, and S_n also depend on S_1 and S_2. There is two-way interaction between subsystem 1 and 2, which means that the change of states in either subsystem has impacts on the other subsystem.

To summarize, the type of shared state variables is determined by the structure of matrix A. Sharing is only a necessary condition for subsystem interaction through state variables. As the two-way interaction case is the most general case, we will use it for all the subsequent discussions.

We demonstrate the decomposition and coordination of two separated subsystems sharing one state variable using a simple linear state-space system example. Consider a linear system with three state variables that can be described by Equation 3-2

EQUATION 3-2

$$\begin{bmatrix} s_1(k+1) \\ s_2(k+1) \\ s_3(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} s_1(k) \\ s_2(k) \\ s_3(k) \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ 0 & b_{22} \\ b_{31} & b_{32} \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix}$$

$$\begin{bmatrix} y_1(k) \\ y_2(k) \end{bmatrix} = \begin{bmatrix} c_{11} & 0 & c_{13} \\ 0 & c_{22} & c_{23} \end{bmatrix} \begin{bmatrix} s_1(k) \\ s_2(k) \\ s_3(k) \end{bmatrix}$$

According to the structure of matrix A, it is feasible to decompose it into two separate subsystems where state variable s_1 and s_2 belong to each individual subsystem respectively and that the state variable s_3 is shared between the two subsystems. It is assumed that:

1. The input u_1 does not have any impact on the state s_2; and the input u_2 does not have any impact on the state s_1
2. The output y_1 does not depend on the state s_2; and the output y_2 does not depend on the state s_1
3. y_i is not directly dependent on u_i (coefficient matrix $D = 0$)

The above assumptions guarantee that the interdependencies among the two subsystems depend solely on the shared state variable s_3. Based on Equation 3-2, a system diagram is constructed and illustrated in Figure 3-2.

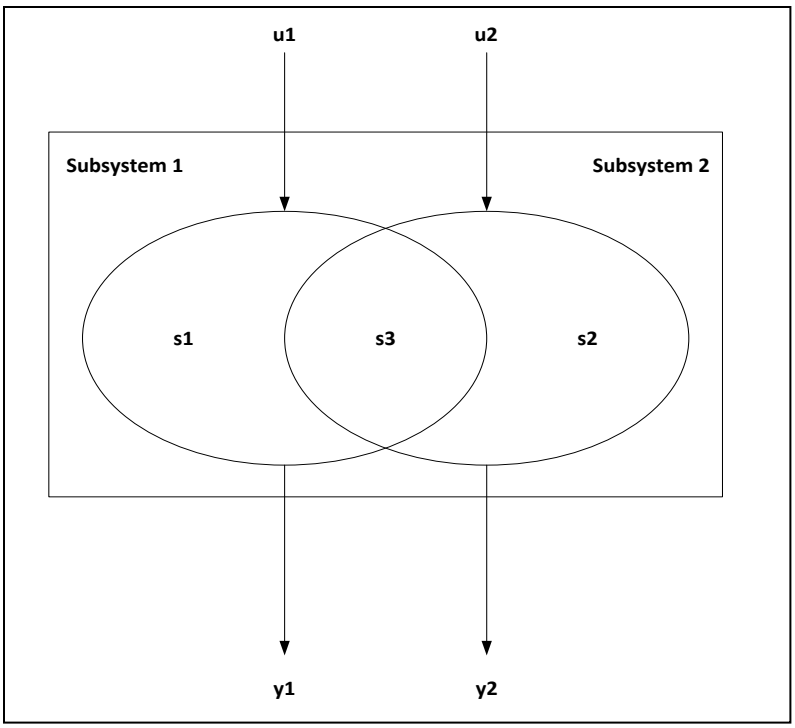

FIGURE 3-2. TWO SUBSYSTEMS SHARING ONE STATE VARIABLE

A decomposition method that transforms subsystems with shared state variables in SoS into subsystems connected through inputs and outputs will enable techniques appropriate for the coordination of input-output systems to be used with the state-space model.

In the case of two subsystems sharing one state variable, the shared state variable belongs to and contributes to the dynamics of both systems. During decomposition, the shared state variable belongs to and thus remains in both subsystems to perform their function. By analyzing the shared state variable s_3, it can be shown that:

EQUATION 3-3

$$\begin{aligned} s_3(k+1) &= a_{31}s_1(k) + a_{32}s_2(k) + a_{33}s_3(k) + b_{31}u_1(k) + b_{32}u_2(k) \\ &= a_{33}s_3(k) + [a_{31}s_1(k) + b_{31}u_1(k)] + [a_{32}s_2(k) + b_{32}u_2(k)] \\ &= a_{33}s_3(k) + z_1(k) + z_2(k) \end{aligned}$$

where $z_1(k) = a_{31}s_1(k) + b_{31}u_1(k)$ and $z_2(k) = a_{32}s_2(k) + b_{32}u_2(k)$

From Equation 3-3, the value of s_3 at $t = k + 1$ depends on three factors:

1. $a_{33}s_3(k)$, is its value at $t = k$
2. $z_1(k)$, contains all necessary information from subsystem 1 at $t = k$ to update s_3
3. $z_2(k)$, contains all necessary information from subsystem 2 at $t = k$ to update s_3

A decomposition scheme based on this property is shown in Figure 3-3. The shared state variable s_3 remains in both subsystems as s_{13} for subsystem 1 and s_{23} for subsystem 2.

For subsystem 1, $s_{13}(k+1) = a_{33}s_{13}(k) + a_{31}s_1(k) + b_{31}u_1(k) + z_2(k)$, and for subsystem 2, $s_{23}(k+1) = a_{33}s_{23}(k) + a_{32}s_2(k) + b_{32}u_2(k) + z_1(k)$.

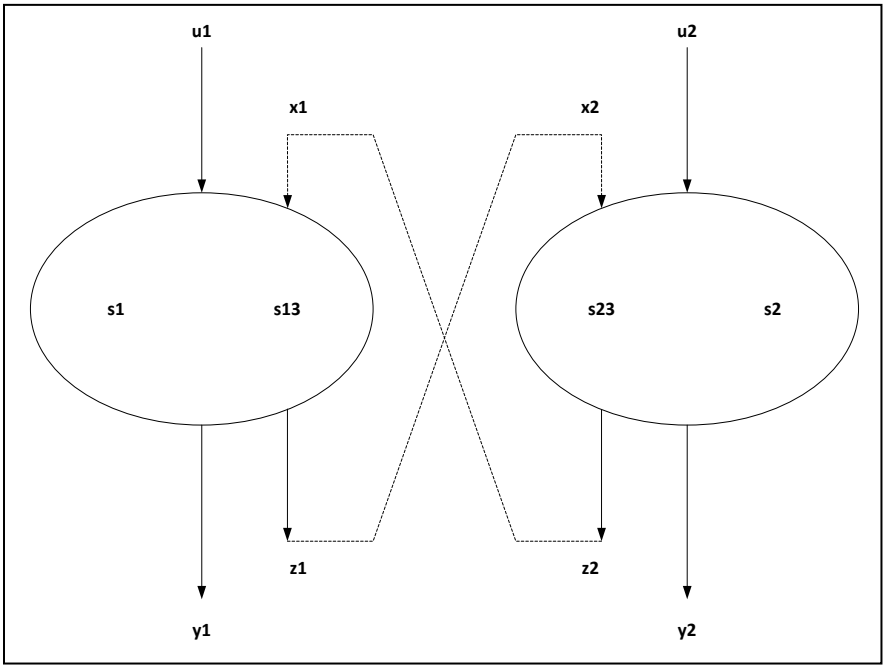

FIGURE 3-3. DECOMPOSITION OF SYSTEMS SHARING ONE STATE VARIABLE

One extra input x and one extra output z are added to each subsystem to incorporate the information needed to update the value of the shared state variable from the other subsystem.

The system models for the decomposed subsystems are derived as:

For subsystem 1

EQUATION 3-4

$$\begin{bmatrix} s_1(k+1) \\ s_{13}(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} \begin{bmatrix} s_1(k) \\ s_{13}(k) \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ b_{31} & 1 \end{bmatrix} \begin{bmatrix} u_1(k) \\ x_1(k) \end{bmatrix}$$

$$\begin{bmatrix} y_1(k) \\ z_1(k) \end{bmatrix} = \begin{bmatrix} c_{11} & c_{13} \\ a_{31} & 0 \end{bmatrix} \begin{bmatrix} s_1(k) \\ s_{13}(k) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ b_{31} & 0 \end{bmatrix} \begin{bmatrix} u_1(k) \\ x_1(k) \end{bmatrix}$$

For subsystem 2

EQUATION 3-5

$$\begin{bmatrix} s_2(k+1) \\ s_{23}(k+1) \end{bmatrix} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} s_2(k) \\ s_{23}(k) \end{bmatrix} + \begin{bmatrix} b_{22} & 0 \\ b_{32} & 1 \end{bmatrix} \begin{bmatrix} u_2(k) \\ x_2(k) \end{bmatrix}$$

$$\begin{bmatrix} y_2(k) \\ z_2(k) \end{bmatrix} = \begin{bmatrix} c_{22} & c_{23} \\ a_{32} & 0 \end{bmatrix} \begin{bmatrix} s_2(k) \\ s_{23}(k) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ b_{32} & 0 \end{bmatrix} \begin{bmatrix} u_2(k) \\ x_2(k) \end{bmatrix}$$

Subject to $x_1(k) = z_2(k)$ and $x_2(k) = z_1(k)$

This decomposition method can be extended to nonlinear systems as long as the transition equation of the shared state variable has the form

EQUATION 3-6

$$s_3(k+1) = g_3\big(s_3(k)\big) + g_1\big(s_1(k), u_1(k)\big) + g_2\big(s_2(k), u_2(k)\big)$$

where $g_1(\cdot), g_2(\cdot)$ and $g_3(\cdot)$ are arbitrary real functions.

3.3. Coordination of Subsystems with Shared State Variables

The interdependencies caused by shared states introduce significant challenges for the decision maker to understand the impacts of their decisions on other connected subsystems, and the impacts of other's decision on their subsystems. Consider the question: How do decision makers of a subsystem achieve their goals when some of its state variables are shared with other subsystems. Clearly the value of the shared state variables cannot be totally controlled by any one subsystem. Furthermore, with some states being partially uncontrollable, what would be the strategy of the decision makers?

On the other hand, an important issue faced by a coordinator of the SoS is how to allocate limited risk management resources to each subsystem over a decision horizon so that the overall performance of the SoS can be optimized, given his understanding of the interdependencies among the subsystems that constitute SoS. A decentralized coordination method is needed for a coordinator to achieve the objectives of the integrated SoS through a properly designed incentive, subsidy or regulation structure, such that decision makers of each subsystem can coordinate with others while achieving their individual objectives.

Fundamental to the decentralization concept is the decentralization principle: *A set of subsystems is optimally controllable in a decentralized manner if and only if there are no variables common to two or more subsystems*. The presence of common or shared variables is equivalent to the existence of constraints between the optimization problems, requiring an integrated or simultaneous solution of two or more sub-problems. The decentralized control problem with coupled decisions has been formulated and discussed in [Lasdon and Schoeffler, 1966]. Their decentralized approach decomposes the overall

system into a set of subsystems, connected through input-output relationships. The solution of the optimized decision problem is obtained by severing the links among the subsystems. Members of the input vector x to each subsystem are regard as independent variables, while they are determined by the other subsystems. Then for each subsystem the output vector y depends only on its input variables u and x. A two-level hierarchical control structure is used. At the first-level each subsystem aims to achieve its own objectives by varying its input and decision variables; and the second-level controller is assigned the job of coordinating the first-level subsystems so that overall system goal is achieved. The decomposition method approaches the decentralized control problem though transforming the interdependency caused by shared states into input-output relations.

An example problem of the coordination between two subsystems in SoS over a decision horizon is illustrated. We consider here the optimization of a single objective discrete-time dynamic state-space system using a multilevel non-feasible method. The linear state space model described in Equation 3-2 and Figure 3-2 is used as an example. The multilevel non-feasible method for the decomposed systems has two dimensions of coordination – between different time stages and between the two subsystems. Assuming the planning horizon to be $[0, T-1]$, the multilevel method decomposes the dynamic problem into T independent static subproblems by cutting the input-output relations among different stages. In the state space representation, the information exchanged between different stages is actually the state variables, as shown in Figure 3-4.

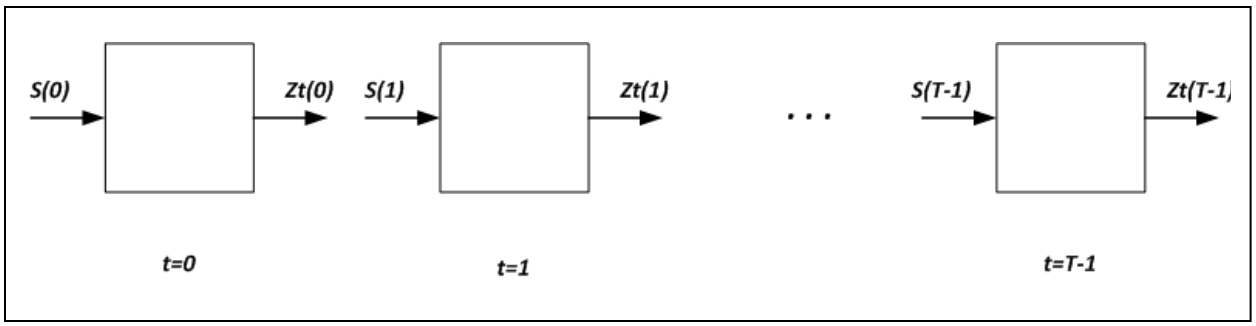

FIGURE 3-4. CUTTING STATE SPACE BETWEEN CONSECUTIVE TIME STAGES FOR DECENTRALIZED CONTROL

We define for each subsystem $i, i = 1,2$ at each stage $k, k = 0, \dots, T-1$ the following variables:

Independent manipulated variables:

- $S_i(k)$, the state vector of subsystem i at time stage k, as vector "inputs" to subsystem i;
- $U_i(k)$, the vector of control variables (decisions) to subsystem i at time stage k;
- $X_i(k)$, the vector of input variables from subsystem j to subsystem i at time stage k

Output/controlled variables:

- $Zt_i(k)$, the state vector of subsystem i at time stage $k+1$, as vector "outputs" from subsystem i at time k to subsystem i at time $k+1$;
- $Zs_i(k)$, the vector of output variables from subsystem i to subsystem j at time stage k;
- $Y_i(k)$, the vector of objective/output variables from subsystem i at time stage k;

The relationships between these variables are summarized here:

EQUATION 3-7

$$Zt_i(k) = A_i S_i(k) + B_i \begin{bmatrix} U_i(k) \\ X_i(k) \end{bmatrix}$$

$$Zs_i(k) = zs_i(k) = a_{3i} s_i(k) + b_{3i} u_i(k)$$

$$Y_i(k) = C_i S_i(k) \text{ for linear system, or}$$

$$Y_i(k) = \left(S_i(k)\right)^T C_i S_i(k) \text{ for quadratic system}$$

(We only consider the case where coefficient matrix $D = 0$.)

The constraints between subsystems are summarized here:

- $Zt_i(k) = S_i(k+1)$ is the constraints for subsystem i between consecutive time stages k and $k+1$ for $k = 0, \dots, T-2$;
- $Zs_i(k) = X_j(k), i, j = 1,2; i \neq j$ are the constraints between subsystem i, j for $k = 0, \dots, T-1$

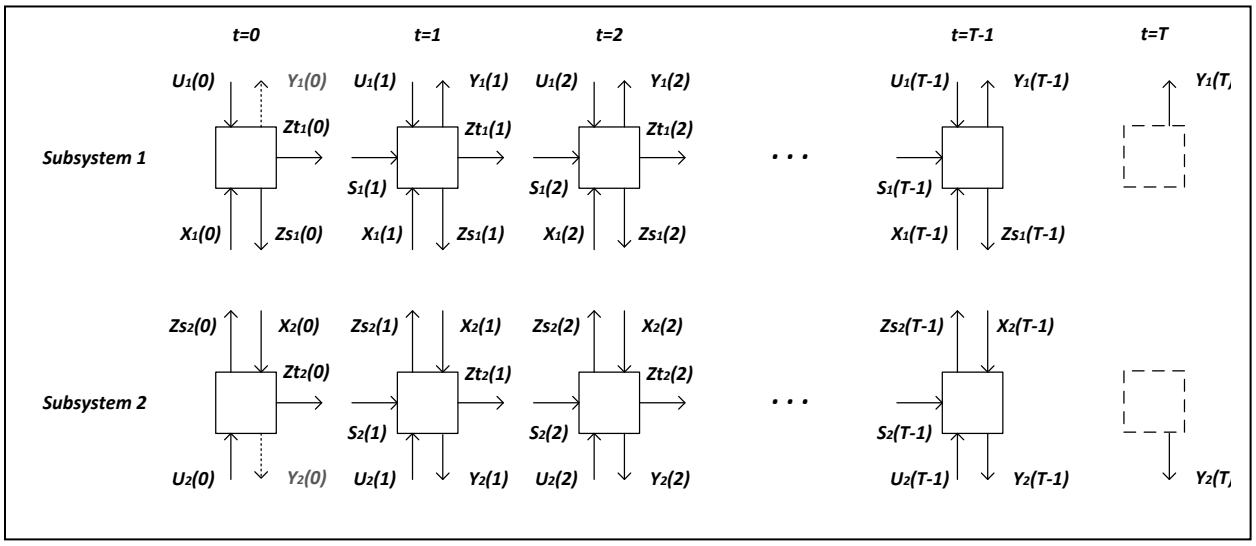

FIGURE 3-5. DECOMPOSITION FROM BOTH TIME AND SUBSYSTEMS DOMAIN, WITH INPUTS AND OUTPUTS OF EACH SUBSYSTEM

Thus we decompose the integrated problem with 2 subsystems and T time stages into $2 \times$ T subproblems, as shown in Figure 3-5.

We assume that the performance of the overall system is a function of the performance of individual subsystem and we choose an additive performance function of the form:

EQUATION 3-8

$$F = \sum_{k=0}^{T-1} \big(y_1(k+1) + y_2(k+1)\big)$$

to be minimized over time period T, by the selection of manipulated/control variables u_i. This overall objective function is separable to each subsystem and each time stage.

We formulate the problem as:

EQUATION 3-9

$$\min_u F = \sum_{i=1}^{2} \sum_{k=0}^{T-1} y_i(k+1)$$

$$s.t. \quad Zt_i(k) = S_i(k+1), \text{for } i = 1,2 \text{ and } k = 1, \dots, T-1$$

$$Zs_i(k) = X_j(k), \text{for } i, j = 1,2, i \neq j; \text{and } k = 0, \dots, T-1$$

To incorporate the constraints, a Lagrangian is formed as:

EQUATION 3-10

$$L = \sum_{i=1}^{2} \sum_{k=0}^{T-1} y_i(k+1) + \sum_{i=1}^{2} \sum_{k=0}^{T-2} \lambda_{ik}[S_i(k+1) - Zt_i(k)]$$

$$+ \sum_{\substack{i,j=1,2 \\ i \neq j}} \sum_{k=0}^{T-1} \mu_{ik}\big[X_j(k) - Zs_i(k)\big]$$

where λ_{ik} and μ_{ik} are Lagrangian multipliers. By reorganizing the Lagrangian function, we decompose it into $2 \times T$ sub-objectives. Define $\tilde{y}_{ik}$ as the adjusted objective function of subsystem i at time stage k:

EQUATION 3-11

$$L = \sum_{i=1}^{2} \sum_{k=0}^{T-1} \tilde{y}_i(k+1)$$

where

EQUATION 3-12

$$\tilde{y}_i(1) = y_i(1) - \lambda_{i0} Zt_i(0) - \mu_{i0} Zs_i(0) + \mu_{j0} X_i(0)$$

$$\text{for } k = 0;$$

$$\tilde{y}_i(k+1) = y_i(k+1) - \lambda_{ik} Zt_i(k) + \lambda_{i(k-1)} S_i(k) - \mu_{ik} Zs_i(k) + \mu_{jk} X_i(k)$$

$$\text{for } k = 1, \dots, T-2$$

and

$$\tilde{y}_i(T) = y_i(T) + \lambda_{i(T-2)} S_i(T-1) - \mu_{i(T-1)} Zs_i(T-1) + \mu_{j(T-1)} X_i(T-1)$$

$$\text{for } k = T-1;$$

During the course of computation, the constrains that the cut variables be equal are not satisfied in general, and we define the error $Et_i(k) = Zt_i(k) - S_i(k+1)$, and $Es_i(k) = Zs_i(k) - X_j(k)$. A two-level non-feasible approach can be used to solve the integrated problem in a completely decentralized manner, where at the lower level the ik^{th} subproblem is to find $S_i(k)$ and $X_i(k)$ such that $\tilde{y}_i(k+1)$ is minimized for fixed λ_{ik} and μ_{ik}, and at the higher level the goal of the coordination process is to find λ_{ik}^* and μ_{ik}^* such that $Et_i(k) = 0$ and $Es_i(k) = 0$ for all i and k. A recursive algorithm through the

coordination of the second-level controller is used to obtain optimal values of λ_{ik} and μ_{ik}, and the optimal solution must be such that the constraints are met. One key problem is how the second-level controller adjusts the parameters by some algorithm such that the differences will be reduced. There are a number of gradient minimizers such as the steepest descent, which will converge to λ_{ik}^{*} and μ_{ik}^{*}. The simplest of those has the form of $\frac{d\lambda_{ik}}{dt} = Et_i(k)$ and $\frac{d\mu_{ik}}{dt} = Es_i(k)$.

Notice that each $\tilde{y}_i(k+1)$ is an unconstrained function only of the variables associated with subsystem i at time stage k, for fixed multipliers λ_{ik} and μ_{ik}, and these variables are independent variables. Thus, if L is to be minimized, each $\tilde{y}_i(k+1)$ must be independently minimized for given λ_{ik} and μ_{ik}. Searching for optimal values of the Lagrangian multipliers λ_{ik} and μ_{ik} is left to the second-level controller. For each subsystem, given certain values of the Lagrangian multipliers, the decision maker tries to minimize a new unconstrained objective function $\tilde{y}_i(k+1)$, which is the sum of the original objective function $y_i(k+1)$ and a adjusted factor: $-\lambda_{ik}Zt_i(k) + \lambda_{i(k-1)}S_i(k) - \mu_{ik}Zs_i(k) + \mu_{jk}X_i(k)$, where $S_i(k)$ and $X_i(k)$ are the inputs from connected subsystems and $Zt_i(k)$ and $Zs_i(k)$ are the outputs to connected subsystems. Minimizing this adjusted factor means that we need to minimize the inputs to the shared state variable and maximize the outputs to the shared state variable from the individual subsystem's perspective. This can be interpreted as maximizing the subsystem's controllability on the shared state variable. If the decision maker has no control over the shared state variable, there is no guarantee that the objective function can be optimized. In such a case, gaining controllability of the shared state variables is an important objective to be considered in the decision making strategy. In other words, when there are shared state variables among

subsystems, the decision makers of each subsystem now optimize an adjusted objective function, which incorporates the controllability of the shared state variable as the second objective. The subsystem decision maker's original single objective decision problem now becomes a two-objective decision problem, and the Lagrangian multipliers serve as weighting factors between the subsystem's original objective and the controllability of the share state variable.

In a real system, the physical form of the Lagrangian multipliers includes incentive, subsidy, policy and regulation, and many other forms, all of which have an influence on the decision maker's preference between the two objectives. The coordinator's role is to find the optimal values of the multipliers such that the optimal decision of subsystem 1 will also help to optimize subsystem 2 through shared states, and vice versa. In such an optimal situation, the values of the Lagrangian multipliers are determined such that the optimal decision for one subsystem generates exactly the required output z to optimize the other subsystem, and together, they optimize the overall objective of the SoS.

The decomposition and coordination methods will be applied in Chapter 4 to understand the stability of complex systems under external perturbations.

4. Systemic Risks in Complex Infrastructure Systems of Systems

System failures, such as power outages, cyber incidents, and collapse of bridges, are among the most severe risk scenarios for infrastructure systems. These failures have to be understood in advance, monitored in real time and if possible forecasted based on available information so that adverse consequences resulting from these failures can be effectively managed. The term "systemic risk" refers to system failures resulted from component interactions rather than component failures. These component interactions are determined by functions, interdependencies, system dynamics, topology and structure, and tradeoffs among competing objectives. This chapter builds on the theories of modeling, decomposition, and coordination of Systems of Systems (SoS) in Chapter 3 and develops a framework to explore systemic risks inherent in complex infrastructure SoS to understand the uniqueness of sources of failures of complex SoS to enable decision makers to mitigate and manage these risks. Among various theories on failure mechanisms of complex infrastructure SoS, this Chapter focuses on a specific failure mode in a nonlinear dynamic multi-objective sequential decision process. We discuss a model-based quantitative analysis of systemic risks in complex infrastructure SoS and

demonstrate that the sources of risk may come from: (i) the decision maker's inappropriate preference between multiple objectives; (ii) unknown coupling (interdependencies) between interconnected subsystems. The purpose of this chapter is to demonstrate the unique source of risks in complex SoS, which doesn't necessarily exist in other systems. Specifically, it demonstrates that complex SoS can fail even if all its components are functioning as they are designed. Thus, reliability theory alone is not sufficient to explain the failures in complex SoS. This result provides a justification for the development of control-based system meta-model to explain complex SoS failure and the identification of precursors in Chapter 5.

Sections in this chapter are organized as follows. Section 4.1 introduces some special characteristics of complex infrastructure SoS as potential sources of systemic risks. Section 4.2 models a dynamic multi-objective sequential decision making process and explores a unique failure model in a standalone subsystem in the SoS. Section 4.3 discusses subsystem stability when it shares state variables with other subsystems, and section 4.4 provides some discussions on how to mitigate risks from subsystem interdependencies.

4.1. Characteristics of Complex Infrastructure Systems of Systems

Complex systems possess some unique features that distinguish them from conventional systems. These features include but are not limited to: nonlinear dynamics, decision based on local information, local connectivity and interaction, strong and multiple feedback loops, long-term memory, cascading failures, and adaptive and emergent behaviors.

Complex infrastructure SoS in particular, is a subset of complex systems where human decisions play an important role in its behaviors. The following set of characteristics is common to many infrastructure SoS:

- Multiple stakeholders and decision makers
- Multiple goals and objectives
- Multiple interconnected and interdependent subsystems through shared states and decisions
- Unknown interactions between subsystems
- Nonlinear system dynamics
- Emergence and adaption

The characteristics of SoS may be useful to unveil some internal risk to the system. One approach to understanding the failure mechanism of complex infrastructure SoS starts from capturing unique system structures and behaviors that may cause system failures. Specifically, we focus more on how the interactions among different system functional components become a source of risk to system failure, rather than the physical reliability of each system component. The next section discusses in detail how a simple model with the above characteristics is developed to analysis systemic risks to complex infrastructure SoS.

4.2. Risks in a Nonlinear Dynamic Multi-objective Sequential Decision Making Process

The decomposition method developed in Chapter 3 provides an opportunity to focus on the understanding of each individual subsystem's behaviors and dynamics. To understand

the systemic risks of complex infrastructure SoS, we first developed a simple mathematical model for an individual subsystem. Although the model is simple, it does contain many important characteristics of complex SoS summarized in section 4.1. We demonstrate through this example that an optimal control strategy doesn't necessarily guarantee system safety. This illustrative model has two state variables, two objective functions, and one decision maker. We assume one of the subsystem's objectives is related to the performance of the subsystem and the other objective is related to the safety of the subsystem. These two objectives are usually non-commensurable and competing with each other such that tradeoffs between objectives must be considered. The decision maker makes a set of sequential decisions based on the observed states of the subsystem over a time horizon to optimize both objectives. A set of system constraints determine the safety operation boundary of the subsystem, and system failure is defined as the situation when the system constraints are no longer satisfied. This type of system structure is a common representation of many real-world systems and has been discussed in detail in Rasmussen's Hierarchical Socio-technical Framework [Rasmussen, 1997], which is used to explain system failure in complex systems. According to Rasmussen's theory:

> *"Any socio-technical system is required to operate between the bounds of workspace defined by economic, functional and safety constraints. Decisionmaking and human activities are dynamic processes continually adapting to the human's environmental conditions and perturbations, such as market competition, economic and political pressures, and legislation. Systems and organizations migrate toward a state close to the boundary of safety regulations either under the influence of pressure toward cost-effectiveness in an aggressive, competitive environment or to buffer the variations in the economic and functional objectives resulting from external perturbations. It is important to identify the dynamic forces that may cause the socio-technical system to migrate towards or cross these boundaries."*

We hypothesize that the decision maker's (increasing) preference on system performance objective against safety objective constitutes a driving force which pushes the system to migrate towards the boundary of safety operation. We extend Rasmussen's theory and test our hypothesis using a model-based quantitative approach.

4.2.1. Model Formulation

Recall the notation used in Chapter 3, and let $s_1(k)$ and $s_3(k)$ be two state variables of the subsystem at time k, where $s_3(k)$ is a shared state variable with other subsystems. Let y_{11} and y_{12} be two objective functions of the subsystem respectively, where they are the functions of both $s_1(k)$ and $s_3(k)$, $y_{11} = y_{11}(s_1(k), s_3(k), k)$ and $y_{12} = y_{12}(s_1(k), s_3(k), k)$. Let $u(k)$ be the decision made by the decision maker at time k.

A second-order state transition function is used to represent the nonlinear dynamics of the subsystem, and a system model is shown in Equation 4-1 (For this section, we temporarily ignore the interdependencies between subsystems, which is the topic of the next section.

EQUATION 4-1

$$\begin{bmatrix} s_1(k+1) \\ s_3(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} s_1^2(k) \\ s_3^2(k) \\ s_1(k) \\ s_3(k) \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(k)$$

$$k = 0, \dots, T-1$$

To simplify the discussion, second order interaction terms (e.g., $s_1(k)s_3(k)$,) are omitted in the model. Although the model is simple, it does contain many important characteristics of complex SoS summarized previously. We will demonstrate later that even this simplified model reveals some interesting findings. Parameters a_{ij} and b_i are

system parameters for the state variables and control respectively. Their values can be obtained through parameter estimation (system identification) based on available databases, or through expert elicitations.

Assume there are two objective functions of the system, $y_{11}(k)$ and $y_{12}(k)$, representing system safety and performance respectively. In many cases, such as quality control, the decision maker aims to maintain the states of the system within a certain boundary and to minimize the deviation from a nominal value. The safety objective can also take this form, where either a higher or lower shift in the level of state would cause system failure. For simplicity, the two objectives assume the following form:

EQUATION 4-2

$$\begin{bmatrix} y_{11}(k) \\ y_{12}(k) \end{bmatrix} = \begin{bmatrix} (s_1(k) - c_1)^2 \\ (s_3(k) - c_3)^2 \end{bmatrix}$$

$$k = 0, \dots, T-1$$

where c_1 and c_3 are control objectives of the state variables, representing system safety and performance respectively. Due to uncertainties from within or outside of the subsystem, the decision maker does not plan for a long decision horizon to optimize the overall system performance. Instead, the system operates more as a closed-loop feedback system, where the decision maker makes decisions at each time step to compensate for possible errors and adjust the system's states. The problem is formulated as:

EQUATION 4-3

$$\min_{u(k)} \{y_{11}(k+1), y_{12}(k+1)\}$$

for each $k \in \{0, \dots, T-1\}$.

Among various approaches to multi-objective decision making problem, the weighting method [Chankong and Haimes, 1983] is the simplest one, because the objective functions are convex. Let θ be the weight for these two objectives, $\theta \in (0,1)$, the above problem becomes a single-objective decision making problem with θ as a system parameter:

EQUATION 4-4

$$\min_{u(k)}\{\theta y_{11}(k+1) + (1-\theta)y_{12}(k+1)\}$$

for each k $\in \{0, \dots, T-1\}$.

To find the optimal value of $u(k)$, the objective function is rewritten as:

EQUATION 4-5

$$\begin{aligned} y(k+1) &= \theta y_{11}(k+1) + (1-\theta)y_{12}(k+1) \\ &= \theta[s_1(k+1) - c_1]^2 + (1-\theta)[s_3(k+1) - c_3]^2 \end{aligned}$$

$$= \theta[a_{11}s_1^2(k) + a_{12}s_3^2(k) + a_{13}s_1(k) + a_{14}s_3(k) + b_1u(k) - c_1]^2$$

$$+(1-\theta)[a_{21}s_1^2(k) + a_{22}s_3^2(k) + a_{23}s_1(k) + a_{24}s_3(k) + b_2u(k) - c_3]^2$$

By setting the derivative of y with respect to u to zero, we get:

EQUATION 4-6

$$\begin{aligned} \frac{dy}{du} &= 2\theta b_1[a_{11}s_1^2(k) + a_{12}s_3^2(k) + a_{13}s_1(k) + a_{14}s_3(k) + b_1u(k) - c_1] \\ &\quad + 2(1-\theta)b_2[a_{21}s_1^2(k) + a_{22}s_3^2(k) + a_{23}s_1(k) + a_{24}s_3(k) + b_2u(k) - c_3] \\ &= 0 \end{aligned}$$

By rearranging the above equation, we get:

EQUATION 4-7

$$u^*(k) = \frac{\theta b_1[c_1 - a_{11}s_1^2(k) - a_{12}s_3^2(k) - a_{13}s_1(k) - a_{14}s_3(k)]}{\theta b_1^2 + (1-\theta)b_2^2} + \frac{(1-\theta)b_2[c_3 - a_{21}s_1^2(k) - a_{22}s_3^2(k) - a_{23}s_1(k) - a_{24}s_3(k)]}{\theta b_1^2 + (1-\theta)b_2^2}$$

Let $= \frac{1}{\theta b_1^2 + (1-\theta)b_2^2}$, and substitute $u^*(k)$ into the state transition matrix in Equation 4-1, we get:

EQUATION 4-8

$$s_1(k+1) = a_{11}s_1^2(k) + a_{12}s_3^2(k) + a_{13}s_1(k) + a_{14}s_3(k) + b_1u^*(k)$$

$$= a_{11}s_1^2(k) + a_{12}s_3^2(k) + a_{13}s_1(k) + a_{14}s_3(k)$$

$$+\rho\theta b_1^2[c_1 - a_{11}s_1^2(k) - a_{12}s_3^2(k) - a_{13}s_1(k) - a_{14}s_3(k)]$$

$$+\rho(1-\theta)b_1b_2[c_3 - a_{21}s_1^2(k) - a_{22}s_3^2(k) - a_{23}s_1(k) - a_{24}s_3(k)]$$

By rearranging the above equation, we get:

EQUATION 4-9

$$s_1(k+1) = [a_{11} - \rho\theta a_{11}b_1^2 - \rho(1-\theta)a_{21}b_1b_2]s_1^2(k)$$

$$+[a_{12} - \rho\theta a_{12}b_1^2 - \rho(1-\theta)a_{22}b_1b_2]s_3^2(k)$$

$$+[a_{13} - \rho\theta a_{13}b_1^2 - \rho(1-\theta)a_{23}b_1b_2]s_1(k)$$

$$+[a_{14} - \rho\theta a_{14}b_1^2 - \rho(1-\theta)a_{24}b_1b_2]s_3(k)$$

$$+\rho\theta b_1^2c_1 + \rho(1-\theta)b_1b_2c_3$$

Following the same procedure, we can get:

EQUATION 4-10

$$s_3(k+1) = a_{21}s_1^2(k) + a_{22}s_3^2(k) + a_{23}s_1(k) + a_{24}s_3(k) + b_2u^*(k)$$

$$= [a_{21} - \rho\theta a_{11}b_1b_2 - \rho(1-\theta)a_{21}b_2^2]s_1^2(k)$$

$$+[a_{22} - \rho\theta a_{12}b_1b_2 - \rho(1-\theta)a_{22}b_2^2]s_3^2(k)$$

$$+[a_{23} - \rho\theta a_{13}b_1b_2 - \rho(1-\theta)a_{23}b_2^2]s_1(k)$$

$$+[a_{24} - \rho\theta a_{14}b_1b_2 - \rho(1-\theta)a_{24}b_2^2]s_3(k)$$

$$+\rho\theta b_1b_2c_1 + \rho(1-\theta)b_2^2c_3$$

To analyze the dynamics of the state variable $s = [s_1; s_3]$, we are interested in the function

EQUATION 4-11

$$\boldsymbol{f}(\theta, \boldsymbol{s}) = \boldsymbol{s}(k+1) - \boldsymbol{s}(k) = \begin{bmatrix} s_1(k+1) - s_1(k) \\ s_3(k+1) - s_3(k) \end{bmatrix} = \beta_2 \begin{bmatrix} s_1^2(k) \\ s_3^2(k) \end{bmatrix} + \beta_1 \begin{bmatrix} s_1(k) \\ s_3(k) \end{bmatrix} + \beta_0$$

WHERE ACCORDING TO EQUATION 4-9 AND

Equation 4-10

EQUATION 4-12

$$\beta_2 = \begin{bmatrix} a_{11} - \rho\theta a_{11}b_1^2 - \rho(1-\theta)a_{21}b_1b_2 & a_{12} - \rho\theta a_{12}b_1^2 - \rho(1-\theta)a_{22}b_1b_2 \\ a_{21} - \rho\theta a_{11}b_1b_2 - \rho(1-\theta)a_{21}b_2^2 & a_{22} - \rho\theta a_{12}b_1b_2 - \rho(1-\theta)a_{22}b_2^2 \end{bmatrix}$$

EQUATION 4-13

$$\beta_1 = \begin{bmatrix} a_{13} - \rho\theta a_{13}b_1^2 - \rho(1-\theta)a_{23}b_1b_2 - 1 & a_{14} - \rho\theta a_{14}b_1^2 - \rho(1-\theta)a_{24}b_1b_2 \\ a_{23} - \rho\theta a_{13}b_1b_2 - \rho(1-\theta)a_{23}b_2^2 & a_{24} - \rho\theta a_{14}b_1b_2 - \rho(1-\theta)a_{24}b_2^2 - 1 \end{bmatrix}$$

and

EQUATION 4-14

$$\beta_0 = \begin{bmatrix} \rho\theta b_1^2 c_1 + \rho(1-\theta) b_1 b_2 c_3 \\ \rho\theta b_1 b_2 c_1 + \rho(1-\theta) b_2^2 c_3 \end{bmatrix}$$

where parameters $\beta_2, \beta_1, \beta_0$ are functions of θ.

When $\boldsymbol{f}(\theta, \boldsymbol{s}) = 0$, the system state $\boldsymbol{s}$ is in a steady state. Otherwise, it is in a transient state and may or may not reach the steady state depending on its initial values.

EQUATION 4-15

$$\boldsymbol{f}(\theta, \boldsymbol{s}) = \beta_2 \begin{bmatrix} s_1^2(k) \\ s_3^2(k) \end{bmatrix} + \beta_1 \begin{bmatrix} s_1(k) \\ s_3(k) \end{bmatrix} + \beta_0 = 0$$

The solution of Equation 4-15 for different values of θ provides information about the stationary points, stable and unstable regions of the state space. The existence of solutions depends on the values of a_{ij} and b_i, as well as the value of θ. When no solution exists, the system has no steady state and will very likely drift away from the control objectives and cause system failure. Thus the question we want to answer here is that given coefficients a_{ij} and b_i, what is the boundary of θ such that Equation 4-15 has at least one solution.

A further analysis of this specific example reveals that Equation 4-15 represents two hyperbolae in the two-dimensional state space. The problem is then equivalent to determining whether these two hyperbolae intersect or not. A solution to this problem has been proposed by [Wang, et al, 2001].

4.2.2. A Numerical Example and Analysis of Results

Consider a numerical example, let $A = \begin{bmatrix} -0.06 & 0 & 0.93 & -0.01 \\ 0 & -0.06 & 0.01 & 0.91 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$. Figure 4-1 shows the state space of the system with field lines indicating the dynamic behavior of the system when $\theta = 0.2$. The contour lines with value zero depict the steady state of each individual state variable, and the two intersections of the contour lines with value zero (point A and B in Figure 4-1) are the solution to Equation 4-15 thus the stationary points of the system. A further analysis of the field lines reveals that point A is a stable stationary point and point B is an unstable stationary point. Given an appropriate initial state of the system, the states of the system will converge to the stable stationary point A over time.

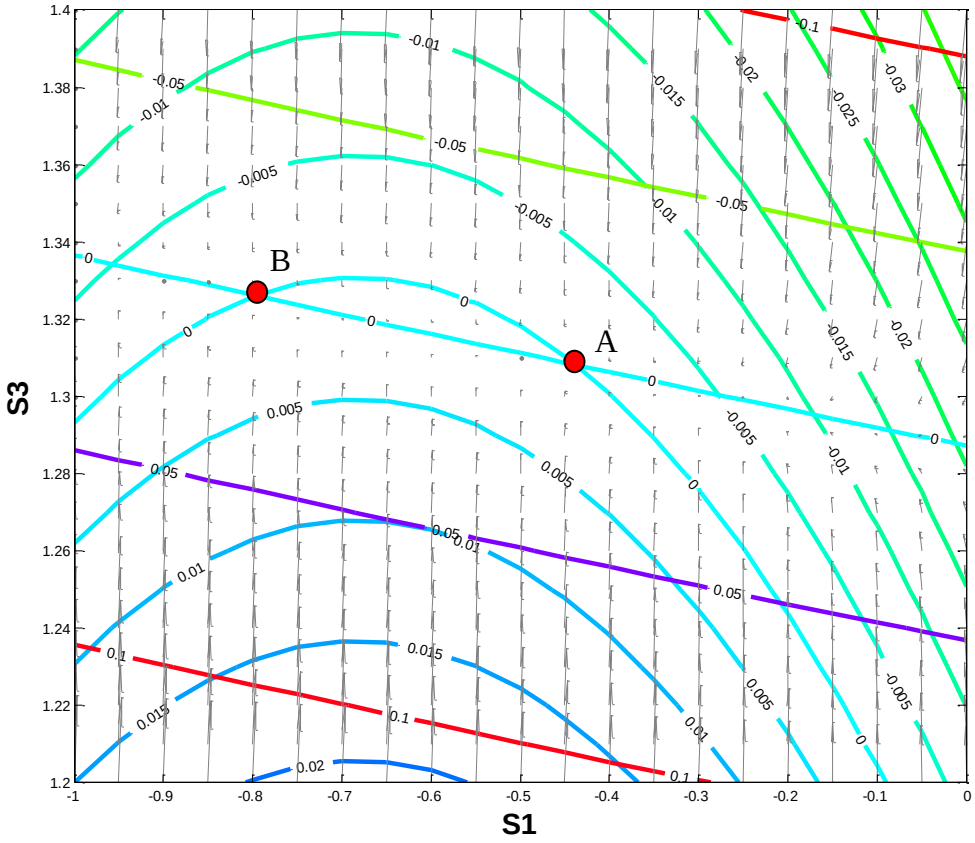

FIGURE 4-1. TWO STATIONARY POINTS IN THE STATE SPACE, $\theta = 0.2$

Now we gradually reduces the value of θ to simulate the decision maker's increased preference on objective function y_{12} over y_{11}. When $\theta \approx 0.103$, Figure 4-2 shows that the two contour lines with value zero are tangent to each other and there is only one stationary point in the state space. This means that Equation 4-15 has only one solution.

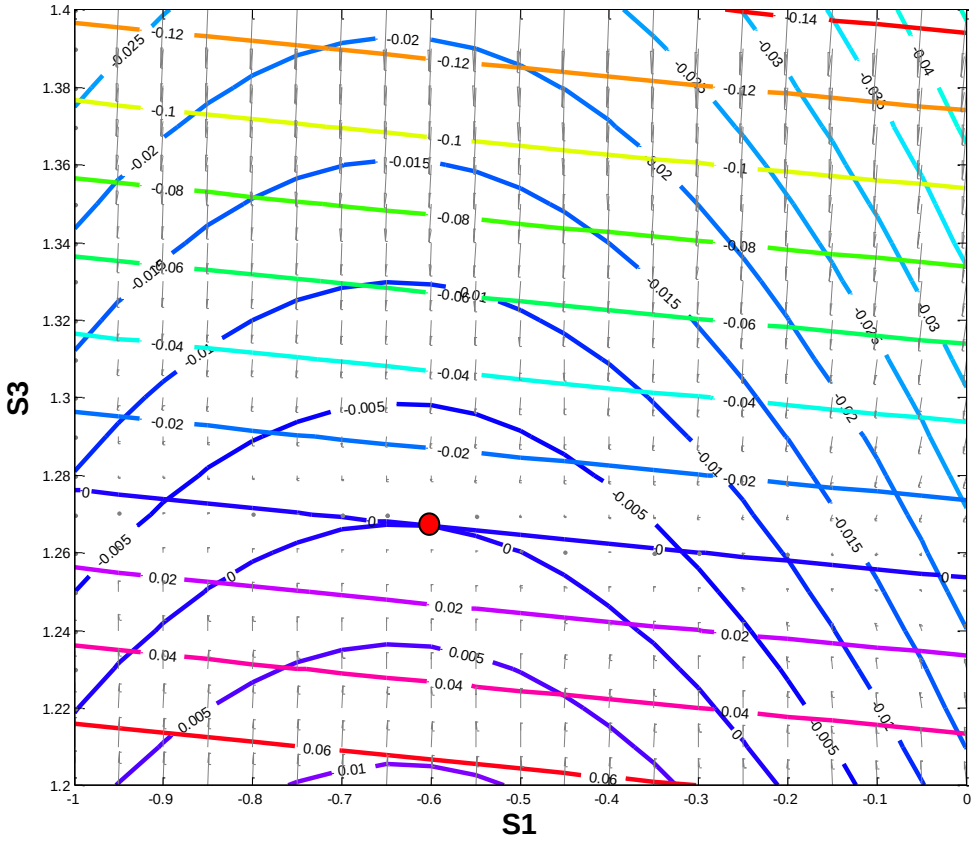

FIGURE 4-2. ONE STATIONARY POINTS IN THE STATE SPACE, $\theta = 0.103$

When θ decreases further, the two contour lines with value zero will separate from each other, thus there is no solution to Equation 4-15, as shown in Figure 4-3. This means that when the value of θ is less than a threshold which is determined by system parameters, there will be no steady state in the state space and the values of the state variables may drift (or oscillate) beyond the safety operation boundaries and cause system failures. We have to emphasize that even in this case, the system is controlled under an "optimal" control strategy $u^*(k)$, where the decision maker aims to optimize

two objective functions given a certain value of θ. As we can see, this type of optimal control strategy doesn't necessarily ensure the safe operation of the system.

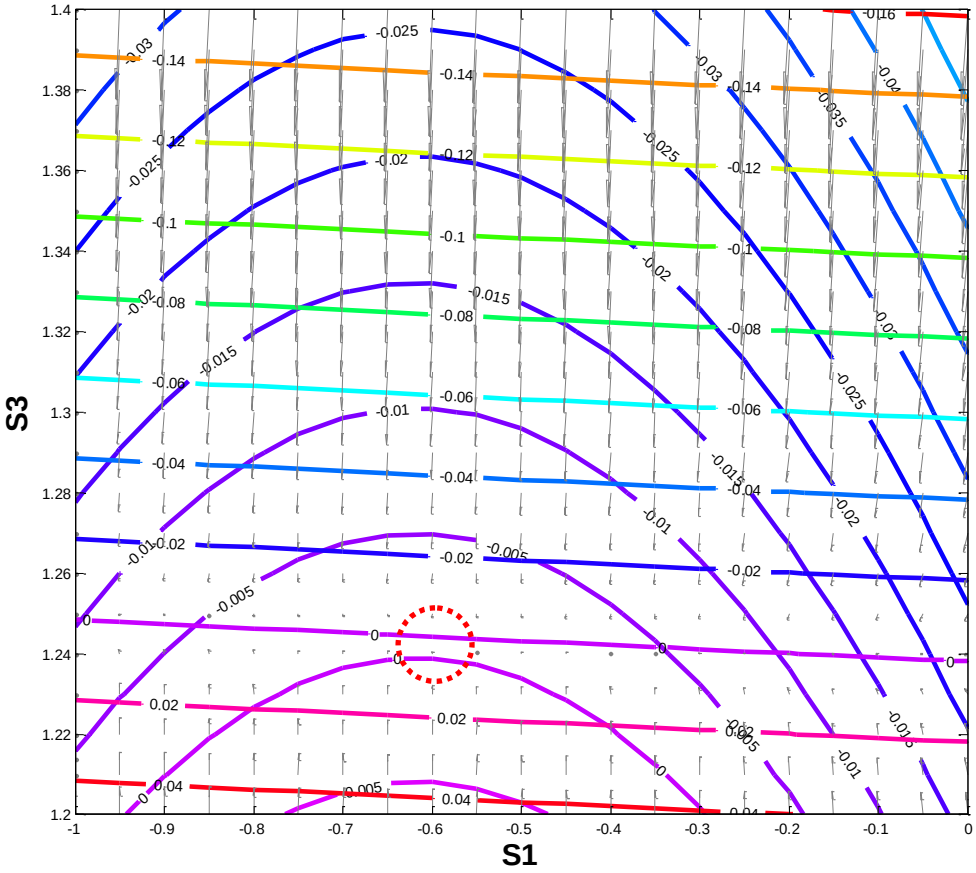

FIGURE 4-3. NO STATIONARY POINT IN THE STATE SPACE, $\theta = 0.05$

Figure 4-4 shows the trajectory of both stationary points in the state space as the value of θ decreases from 0.5 to 0.103. The blue solid line is the trajectory of stable stationary point A and the green line is the trajectory of unstable stationary point B. As θ decreases, they eventually converge to one point, and beyond a threshold of θ, there is no stationary point in the state space (a bifurcation).

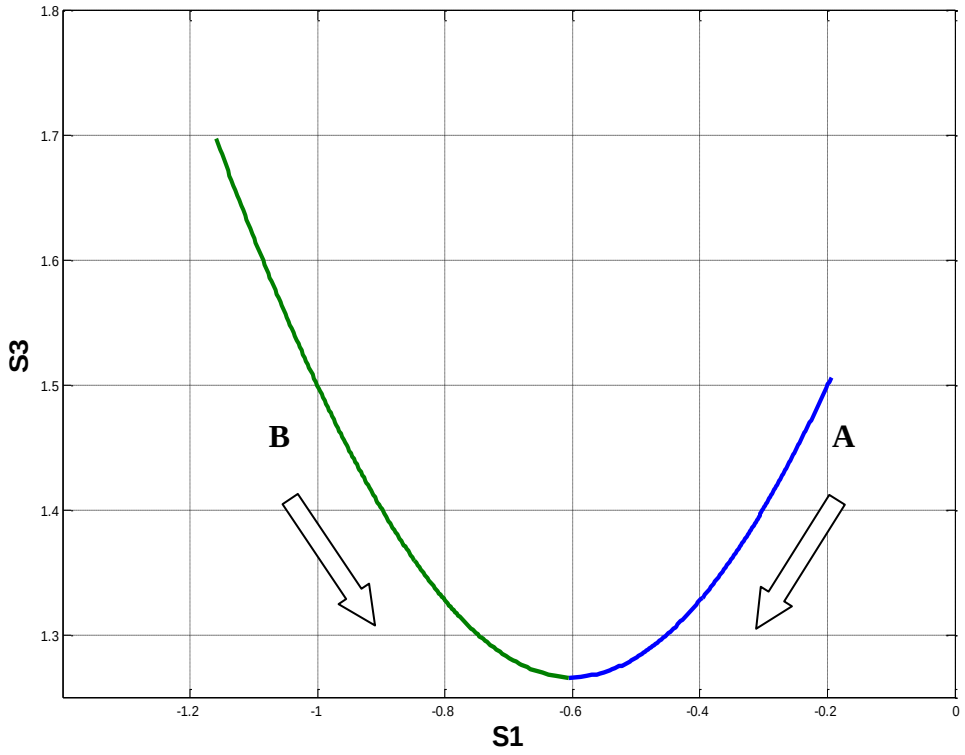

FIGURE 4-4. THE THEORETICAL TRAJECTORY OF TWO STATIONARY POINTS AS A FUNCTION OF θ

Figure 4-5 shows the trajectory of state variable s_1 (black) and s_3 (brown) as a function of θ using simulation, with initial values $s_1 = s_3 = 9.105$ and control target value $c_1 = c_3 = 1.225$. The value of θ decreases gradually from 0.5 to 0 over time. The simulation result shows that the system stays at one of the stationary point (point A) after a short transient period at the beginning of the simulation, and moves slowly as θ decreases. When θ crosses the threshold (0.103), the system still remains around the previous stationary points for a while until a sudden change in both of the state variables. The results indicate that the values of both state variables go to negative infinity. The threshold of θ is a crossover point of the system between its stable region and unstable region. If the decision maker's preference on the system performance objective over the

safety objective becomes so high such that θ moves across this crossover point, the system will not be able to return to a stable region even with a feedback mechanism. The crossover occurs only due to the decisionmaker's preferences and not due to any external forces, thus this is a systemic risk to the system. If a system failure is defined as an operation boundary for either state variable, this situation implies a system failure.

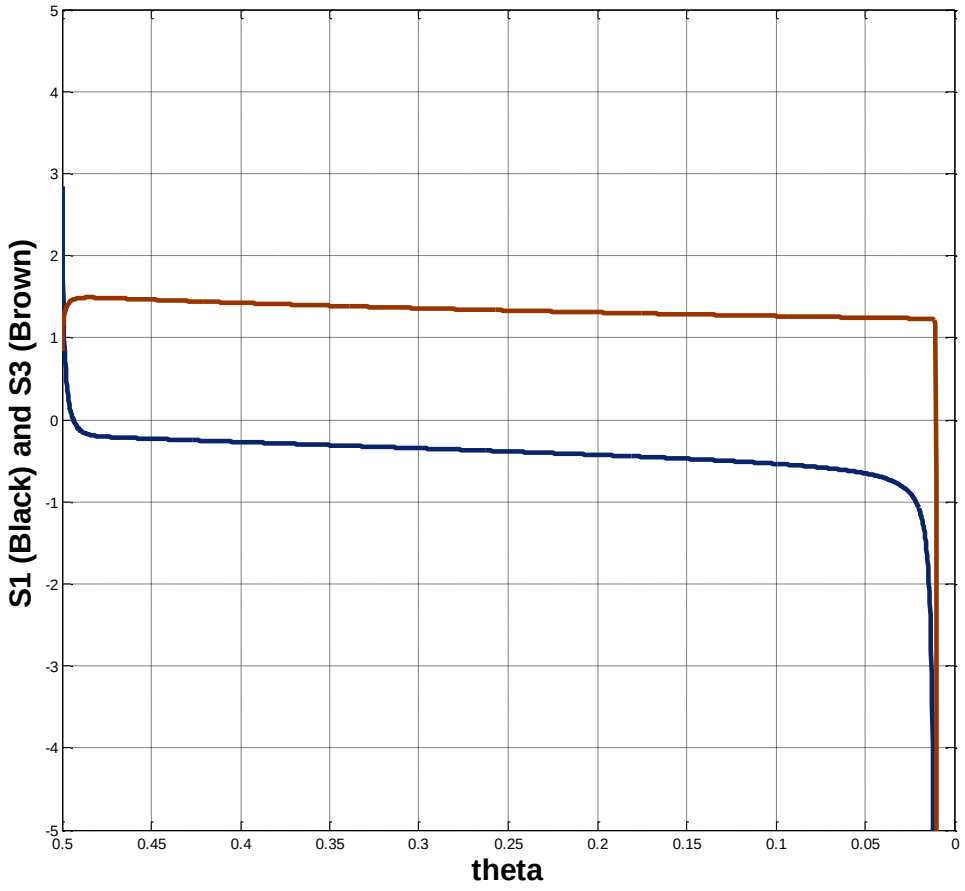

FIGURE 4-5. THE SIMULATED TRAJECTORY OF STATE VARIABLES THROUGH SIMULATION

4.2.3. Implications

Results from the above analysis deserve further discussion. In a multi-objective decision problem, all solutions on the Pareto-optimal frontier are considered feasible and optimal. Decision makers can choose whatever solution on the frontier based on their preferences. This is true if the stability of a nonlinear system is not considered. This analysis reveals that when the decision maker has a very high preference for one objective over another and thus chooses the solutions on the extreme sides of the Pareto-optimal frontier, a nonlinear system may eventually enter a region which is not stable. The decision maker's slowly changing preference, either intentionally or spontaneously, may be the result of pursuing better system performance or reacting to aggressive competitive market conditions. This dissertation doesn't intend to identify the reasons why decision maker's preference changes, but we argue that the changing preference does have impacts on system stability and safety.

The system behavior is also similar to a bifurcation structure as they are both dynamic and nonlinear with slowly changing parameters. The threshold of θ is a boundary between a stable and unstable region in the state space, and is analogous to a bifurcation point or critical point in a bifurcation system. Different types of bifurcation structure in complex systems have been explained through the principle of universality [Gros and Markovic, 2012]. In this dissertation, we believe that the decision maker's preference structure in a multi-objective decision process might be another potential mechanism for the system to have such structures.

This analysis also provides insights into identifying precursors and designing warning systems for complex infrastructure SoS. Traditional control paradigms monitor the trends

and patterns in the controlled process, and use hypothesis testing to identify anomalies in the process. This approach is reasonable for linear systems as the changes in the state space of linear systems are usually predictable through its trend. However, in a nonlinear system, state changes are usually abrupt and trend-based technologies are quite limited in predicting these changes. The simulation results call for other metrics to be used as signals to predict abrupt state changes in a nonlinear system.

The parameter θ, which is literally the weight a decision maker places on objectives, is a measure of the decision maker's subjective preference between objectives. It is a latent variable which is difficult to measure or quantify even by the decision maker himself. However, as we can see from the example, when the value of θ is very low, the system becomes unstable and introduces significant risks of system failure. The ability to interpret a decision maker's preference from the observed states of the system along the sequential decision making process is essential to predict potential system failures. A method is needed to detect the change in a decision maker's preference structure such that abrupt state changes and system failures in a nonlinear multi-objective system can be predicted.

Finally, the mathematical model developed in this section may be used as a good approximation to real-world systems where the decision frequency is high (in terms of seconds or hours) and computers automatically make decisions based on some predefined strategies, algorithms, or decision rules. Such systems include high frequency trading system, power grid, and some of the Supervisory Control and Data Acquisition systems (SCADAs). The model might not be a good approximation for a system where the decision period is long (in terms of months or years) and the human is the major decision

maker. This is because humans do not usually mechanically follow a simple decision rule, and when new information such as abnormal signals is received and there is enough time to investigate the cause, in most cases system failure can be prevented. In addition, system topology, parameters, and decision rules may change under different situations, and one single model cannot capture this adaptive behavior of the system, especially in some extreme cases. Although the limitation of the model should be understood before applying it to a specific system, we are not claiming that low-frequency human decision systems are free from this type of systemic risks. Without a well designed process to identify and discover the emergent forced changes to the system, especially the precursors to system failures, the risk from a multi-objective decision making process cannot be ignored.

4.3. Risks Caused by Interdependencies through Shared States and Decisions

The previous section discusses the systemic risks within a single subsystem. Complex infrastructure SoS are usually of large scale, and consist of many interconnected and interdependent subsystems. These interdependencies usually take the form of shared states and decisions. When a state variable is common to two or more subsystems, it introduces couplings between these subsystems and any decision of one subsystem may propagate through the coupling and cause intended or unintended consequences to other interdependent subsystems. Understanding this source of systemic risk from subsystem interdependencies in complex infrastructure SoS is necessary for a comprehensive risk assessment and management of these systems.

In section 4.2, the subsystem operates in a stable region as long as the value of θ doesn't exceed a certain threshold. In practice, the interdependencies among subsystems

are often difficult to identify and quantify when subsystems are not coordinated or they do not share any information, and decision makers experience unexpected changes in the shared state. When one of the state variables is shared by other subsystems, the impacts from other subsystems on the shared state variables can be usually treated as perturbations that force system states to deviate from the stable region. If perturbations are introduced to the state variables, depending on the magnitude of that perturbation and the margin around the stable region in the state space, the system may or may not return back to the steady state. If the perturbations are strong enough, the system will probably become unstable and experience sudden state changes. In other words, interdependencies between subsystems further reduce the safety margin of θ due to the perturbations introduced to the subsystem. The ability of a subsystem to withstand this perturbation depends on the size of the stable region around the system operation point. The system's stability under external perturbations induced by the shared states is an important issue in understanding systemic risks in complex infrastructure SoS. In some extreme cases, a small perturbation may force the system into an unstable state. Risk management of systemic risks in complex infrastructure SoS requires an understanding of the system's stability under external perturbations.

According to the decomposition method discussed in section 3.2, a subsystem sharing a state variable with another subsystem is equivalent to a standalone subsystem with an extra input adding to its shared state variable. The decisions made in another subsystem will have certain impacts on the shared state variable through this extra input. In this case, the extra input can be treated as a source of perturbation to the system, with unpredictable magnitude.

When subsystems are not coordinated or they do not share any information, a decision maker of a subsystem has no information on the shared state variable, and Equation 4-1 becomes

EQUATION 4-16

$$\begin{bmatrix} s_1(k+1) \\ s_3(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} s_1^2(k) \\ s_3^2(k) \\ s_1(k) \\ s_3(k) \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(k) + \begin{bmatrix} \Delta(k) \\ 0 \end{bmatrix}$$

where $\Delta(k)$ represents the perturbation caused by the interdependent subsystems, and can be treated as a random variable. Assume that Δ has a much lower frequency than the decision period, the state of the subsystem will remain at the stable stationary point most of the time for a given level of θ until the perturbation forces the state away from the stationary point by a distance of $\Delta(k)$. A common fourth-order Runge–Kutta method [Press et al, 1992] can be used to determine whether the new state is in the unstable region of the state space. The fourth-order Runge–Kutta method numerically integrates ordinary differential equations by using a trial step at the midpoint of an interval to cancel out lower-order error terms. This method is reasonably simple and robust and can be approached using numerical methods combined with an intelligent adaptive step-size routine.

The distance between a stable stationary point and the boundary of a stable region is called the *safety margin* of the subsystem. If the perturbation is strong enough to push the states beyond the boundary of stable region, the system will not return back to the stable stationary point and eventually cause system failure.

Using the Runge–Kutta method, we can calculate the trajectory of system states under different levels of perturbation Δ. Figure 4-6 shows 10 trajectoris with Δ ranging from 0.05 to 0.5 with step of 0.05, with $\theta = 0.2$. The black trajectories are those returned to the stable stationary point, and the blue trajectories are those not returned to the stable stationary point. In this case, when $\Delta > 0.35$, the system will leave the stable region, and thus the boundary of the stable region is around 0.35.

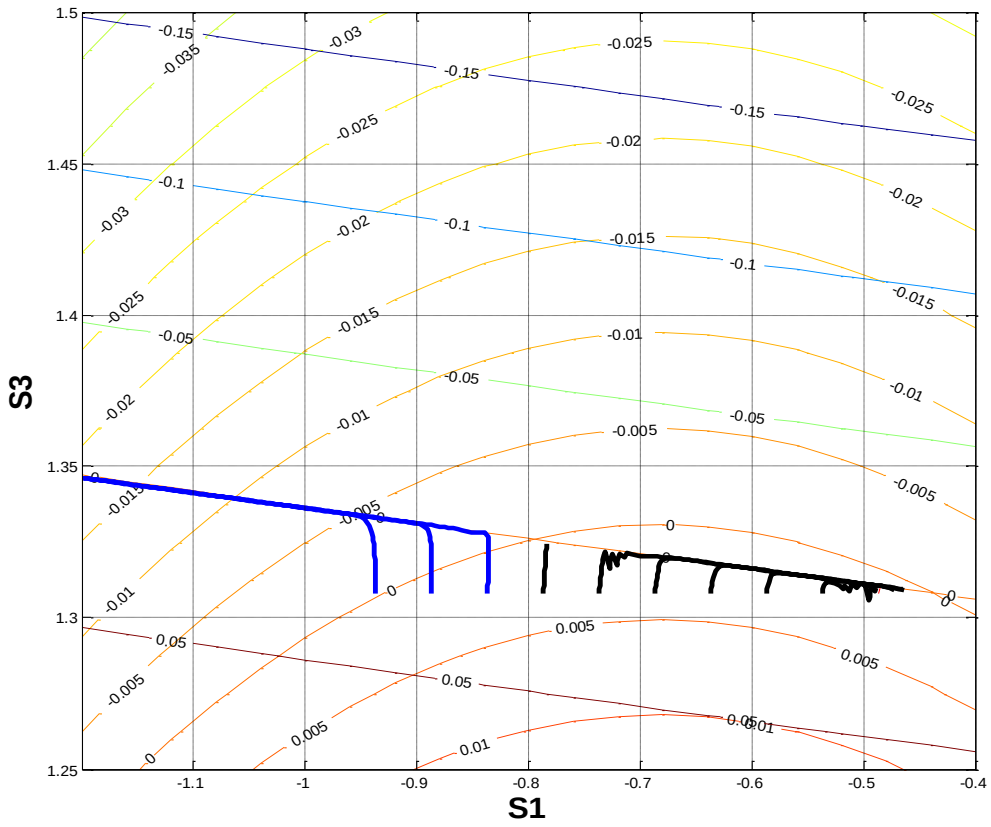

FIGURE 4-6. THE THEORETICAL TRAJECTORIES OF SYSTEM STATES UNDER DIFFERENT LEVELS OF PERTURBATION

Figure 4-7 illustrates the simulated trajectory of the system in the state space given three different levels of perturbations when $\theta = 0.2$. In subplot 1, the perturbation level is 0.2 and the system returns to the steady state value. In subplot 2, the perturbation

pushes the system into a trajectory that slowly departs from steady state and eventually leaves the stable region. In subplot 3, the perturbation is strong enough to push the system into the unstable region immediately.

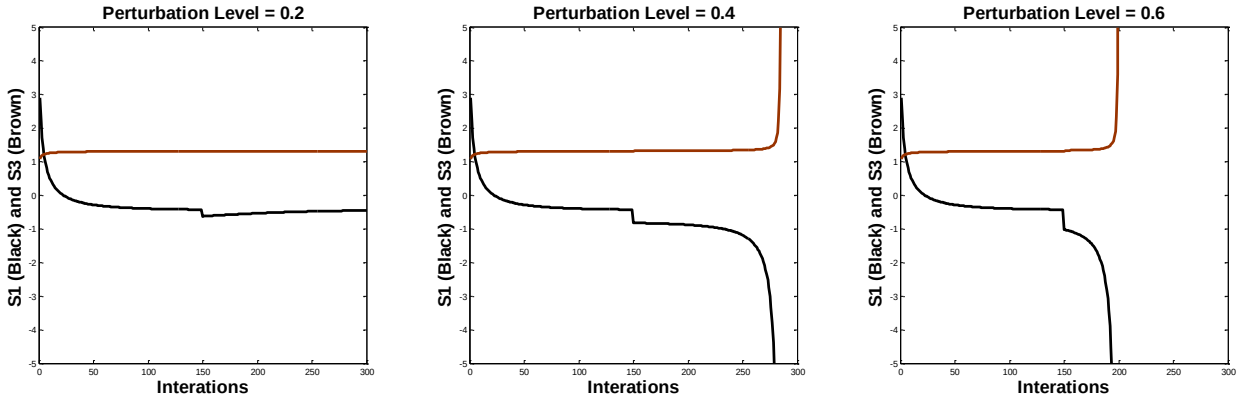

FIGURE 4-7. SIMULATED TRAJECTORIES OF SYSTEM STATES UNDER DIFFERENT LEVELS OF PERTURBATIONS

In addition to the Runge–Kutta method, the safety margin in Figure 4-1 can also be approximately estimated from the two roots of Equation 4-15. As we can see from Figure 4-1, the field line are nearly vertical until reaching the curve $s_3(k+1) - s_3(k) = 0$, at which point the field lines become horizontal. As a result, regardless of the magnitude of the perturbation on the x-axis, the state of the system will move quickly alone the vertical direction until it reaches this curve, and slowly follows the curve. If $\Delta > |AB|_x$, the state will move to the left and be repelled from the safety region; if $\Delta < |AB|_x$, the state will move to the right and reach the stable stationary point A. From Figure 4-1 we can see that the stable region is approximately the horizontal distance from point A to the unstable stationary point B. When Equation 4-15 has two roots,

EQUATION 4-17

$$safety\ margin \approx |AB|_x = |root1_x - root2_x|$$

Equation 4-17 indicates that the safety margin is also a function of θ, as shown in Figure 4-8. When θ decreases the safety margin of the system also decreases, which reduces the ability of the system to withstand unexpected impacts from other interdependent subsystems. For the subsystem to withstand the perturbation, the safety margin of the system must be greater than the potential perturbations caused by shared states.

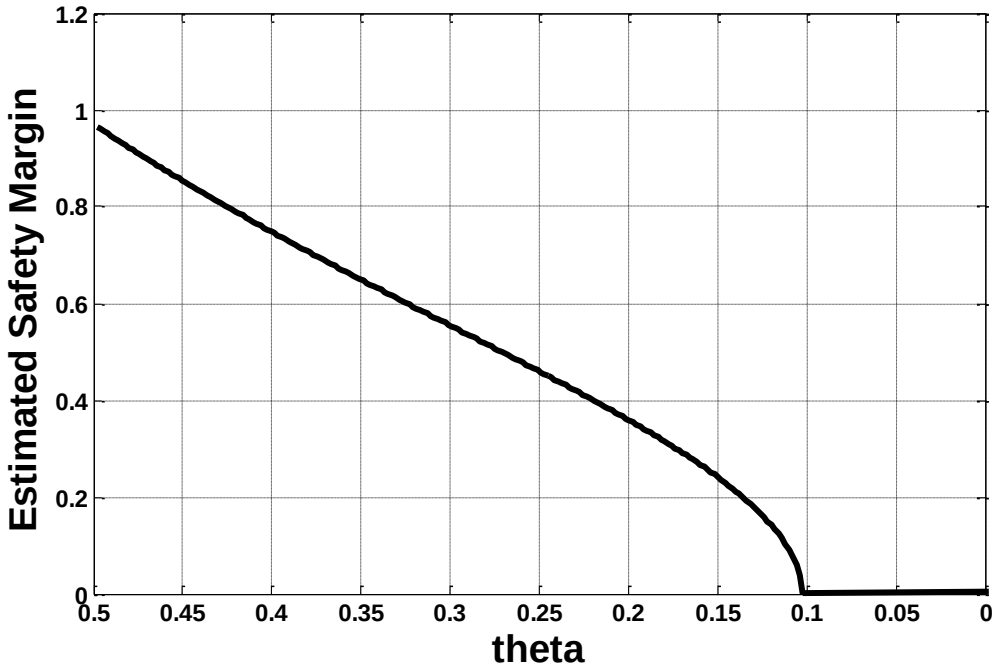

FIGURE 4-8. SYSTEM SAFETY MARGIN AS A FUNCTION OF θ

4.4. Mitigating risks caused by interdependencies through coordination

The risks caused by subsystem interdependencies can be mitigated in two ways. The decision maker may operate the system around a region in the state space where the safety margin is sufficiently large to withstand most unexpected perturbations. However, in most cases this solution is at the expense of degrading system performance. An alternative approach is through system coordination, by which the impacts from other

interdependent subsystems can be anticipated and controlled while still meeting the objectives of the subsystem.

The coordination method introduced in section 3.3 ensures that the optimal decision for one subsystem generates exactly the required output x to optimize the other subsystem, and vice versa. In other words, the two decision makers are coordinating with each other to optimize their individual objectives, and the "perturbations" through shared states now become part of the inputs necessary to optimize the system. Of course, this assumes that the system model is known to the higher-level controller, which requires information sharing among different subsystems within SoS. Information sharing between subsystems plays an important role in better modeling subsystem interdependencies and to anticipate perturbations caused by interdependencies. Knowing the impacts on shared states from other connected subsystems reduces uncertainties in the shared states, which enables the decision maker to select better options to cope with these impacts. The role of "optimization" in the decision process also deserves further investigation. We believe that the system will be much more stable if a "satisfying" decision rule [Simon, 1979] is used instead of an "optimizing" decision rule.

5. Precursor Analysis for Complex Infrastructure Systems of Systems

Chapter 4 discusses one specific failure mechanism of complex infrastructure SoS. However, many other failure mechanisms in complex infrastructure SoS may exist, and the timely detection of signs of emergent forced changes leading to system failures provides opportunities for decision makers to take preventative courses of actions (COAs) to reduce the potential adverse consequences. This chapter develops a systemic precursor analysis framework for complex infrastructure SoS, in which a pro-active, dynamic anticipatory analysis tool is designed to identify, prioritize, and evaluate different sources of emergent forced changes which have the potential to cause system failure. Section 1 provides an overview of the proposed approach and sets goals and tasks for the precursor analysis framework. Section 2 extends the theory in chapter 4 and develops a systemic way to identify various failure mechanisms of complex infrastructure SoS. Section 3 discusses a precursor filtering and prioritization process which enables the

design of an efficient precursor monitoring system. Section 4 discusses how to reduce hindsight bias and improve decision making through evaluating multiple precursors.

5.1. Introduction

This chapter introduces a systemic precursor analysis framework for complex infrastructure SoS, with the goal to (i) explore different failure mechanisms of complex infrastructure SoS; (ii) design an efficient monitoring system; and (iii) reduce hindsight bias in using precursors. First, to identify precursors to system failure, it is necessary to understand how the system fails and the causal relationships among system components so that all major failure modes and failure paths can be explored. A system meta-model describing the functional relationships among its components is needed to perform this analysis. However, the number of resulting precursors to system failure might be large for a complex infrastructure SoS such that it is not practically feasible for a monitoring system to keep track on all of them. We demonstrate that even though precursors have less chance of leading to system failures, all identified precursors must be evaluated and prioritized in terms of its likelihood to cause system failure and time-to-failure in order to design a pro-active and efficient monitoring system. Finally, in risk management of complex infrastructure SoS, the decision maker needs to decide which course of action to be taken if the probability of system failure becomes higher given some observed precursors. As there are multiple ways a system can fail, the successful risk management actions depend on an understanding of what the real causes are and what failure mode is more likely to be expected. When the monitoring system detects a precursor, the likelihood of each system failure mode needs to be quantified, evaluated, and compared, with the uncertainties in the detection and prediction process accounted. To achieve the

above goals, this chapter (i) develops a generic quantitative meta-model of the complex infrastructure SoS based on control theory; (ii) identifies, filters and prioritizes precursors to system failure based on their likelihood and urgency to cause system failure; and (iii) improves situation awareness of the decision maker through evaluating and comparing the likelihood of all identified failure modes of the system.

In risk analysis, there are generally two types of precursors of a system: precursors to the initiating events and precursors to system failures. Initiating events to the system may arise either externally or internally. Precursors to external initiating events cannot be observed using information within the system. For example, natural disasters such as earthquakes pose a great threat to the reliability of bridges. However, precursors to earthquakes cannot be observed by monitoring the state of the bridge system. An understanding of the mechanisms of earthquake and information from the surrounding geological or seismical system must be used to identify these precursors, which is beyond our discussion of bridge infrastructure systems. Precursors to internal initiating events can be observed by monitoring the states of that system. However, depending on the type of initiating event, different models are needed to identify its precursors and detecting methods. In a deductive risk analysis, an initiating event is important only if it can cause significant consequences to the system, such as system failure. In addition to that, section 5.2 demonstrates that a meta-model can be used to explain system failure and identify most of its precursors. This dissertation focuses on the identification, prioritization, and evaluation of precursors to system failures.

First, a set of terms that will be used in the following discussion are introduced. *System failure* is the state or condition of not meeting a desirable or intended objective of

the system, and may be viewed as the opposite of system normal operation. A *failure mode* is the way in which a system fails functionally. Often a system has multiple different failure modes [Langford, 1995] [SAE, 1967]. We adopt a more general definition of a *precursor* as a combination of events, conditions, system states, and the complete identified possible sequence that significantly increase the probability of system failure within a specific future time domain. A precursor can be observed and detected based on supporting evidence, which may come from multiple sources, including but not limited to inspections and observations, reports, expert judgment, intelligence, and physical sensor measurements. Observing a precursor indicates an increase of the likelihood of multiple failure modes simultaneously, which may explain partially the discrepancies between pre- and post-accident risk assessment using precursors. The goal of precursor analysis is to quantify the likelihood and urgency of each failure mode of the system given the observation of precursors.

Commonly used precursor identification and monitoring methods fall into three general categories: precursors based on trend prediction, statistical correlation, and causal relations. In a trend-based approach, some safety-related states of the system are controlled and the system fails when one or some of these states exceed a predefined threshold. These states are continuously measured or estimated and their trends are used to predict whether the threshold will be reached or not in the future. An example of this type of system can be found in [Paté-Cornell, 1986] where the density of smoke particles is used as a precursor for fire in the building. While this approach is easy to implement, the trend-based prediction is not very reliable and the selection of threshold has to balance both false alarms and false positives. In addition to that, this approach only

identifies the “sharp-end” factors [Reason, 1990] leading to system failure and thus fails to provide long-term prediction and results in limited response time. Approaches based on statistical correlation identify statistical measures, which are significant prior to system failure. Examples include using bus voltage frequency data as a precursor to blackouts in the power grid [Hines *et al*, 2011]. This approach is useful when the causal factors of system failure are difficult to understand. However, a training data set including the data prior to actual system failure is needed to enable model learning, thus signals related to unknown failure modes might not be identified. In addition to that, the relatively high false alarm rate is an issue when the correlation is not strong. Finally, approaches based on causal relationships try to understand the failure mechanism of the system and identify all causal factors contributing to system failure. These causal relationships can be based on either a reliability model or a control model of the system. This model-based approach systematically explores potential failure modes of the system and is useful in identifying precursors, which have a long-term effect on the safety of the system. However, understanding the causal relationships in a complex system are often daunting.

The precursor analysis approach developed in this dissertation focuses on identifying precursors based on the causal relationship among system components leading to system failure. It uses a meta-modeling approach to modeling the safety control structure in a complex system and thus enables a systemic way to identify causal factors leading to system failure.

This dissertation proposes three key phases for a quantitative precursor analysis framework:

- System Modeling – a generic quantitative meta-model to explain why complex infrastructure SoS fail
- Precursor Identification, Filtering and Prioritization – a pre-screening process to identify the most important precursors for further monitoring
- Precursor Detection and Evaluation – a detection and evaluation process to combine information from multiple precursors to evaluate the likelihood of each system failure mode

This overall process is shown in Figure 5-1.

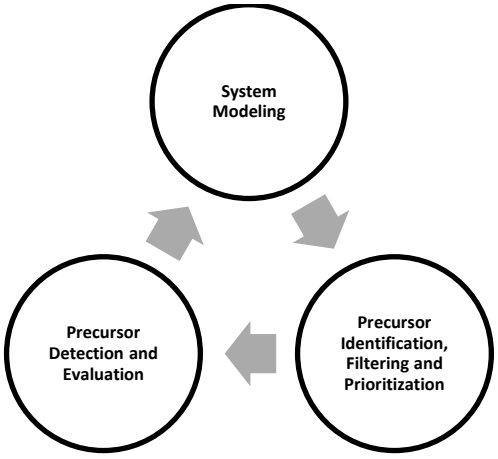

FIGURE 5-1. THE PROPOSED PRECURSOR ANALYSIS FRAMEWORK

As Figure 5-1 shows, this overall framework has to be an iterative process due to the dynamic and adaptive characteristics of the complex infrastructure SoS. Each time we go through this process, our knowledge about the behavior of the complex infrastructure SoS increases, which enables us to develop a better system model, identify more important precursors, and improve the accuracy of the detection and prediction process. These three phases will be discussed in details in the following sections.

Two major sources of uncertainty are involved in this precursor analysis framework: (i) projection uncertainty: the uncertainties in the projection of the likelihood of system failure for a precursor scenario; and (ii) detection uncertainty: the uncertainties in the detection of precursors given observed evidence. In the first case, a precursor is identified and we are interested in knowing the likelihood of system failure in a future time given that a precursor exists. The ability to project the failure likelihood into the future is constrained by the randomness in system dynamics and state observations, as well as our lack of knowledge about the system and oversimplification of the models. In the second case, the likelihood of the existence of a precursor in the first place needs to be estimated by the evidence – usually the imperfect information available at that time. These two sources of uncertainties are quantified respectively and combined in the final evaluation phase.

Both sources of uncertainty contain aleatory uncertainty and epistemic uncertainty. Aleatory uncertainty represents the natural variation in the projection and detection results and is usually quantified by statistical measures. Epistemic uncertainty represents our ignorance about the system. It can be reduced through an iterative learn-as-you-go process, but its quantification is much more difficult. In this dissertation, we restrict our focus mainly on the impacts from aleatory uncertainties.

5.2. Meta-Modeling the Failure of Complex SoS through a System Control Perspective

Complex infrastructure SoS vary widely from one domain to another, and ad hoc system models have been developed for different type of systems to explain their failures, with many of them focusing on component failure or human errors. However, new kinds of accidents have emerged in complex and tightly coupled systems [Perrow, 2011]，and the

role of humans in systems is changing [Hollnagel and Woods, 1983]. Failures in complex infrastructure SoS can no longer be explained by the chains of events from component failures or human error [Hollnagel, 2004]. A new type of system model, Systemic Accident Model, explains system failure from a system's perspective. The STAMP approach [Leveson, 2004] describes system failure resulting from the violation of safety constraints of the system, and it argues that regardless of the type of systems under study, it is possible to identify control structures which enforce those safety constraints in the system. However, STAMP remains a qualitative analysis tool for independent systems and we can extend this idea to a quantitative meta-model of the complex infrastructure SoS.

5.2.1. Identifying Failure and Failure Modes

Before developing a system model, we need to define the boundary of the system and what system failure is. The boundary of the system usually includes physical and management boundaries. The physical boundary defines the physical components of the system, and the management boundary defines the various stakeholders (owner, manager, users) and their behaviors (decisions).

The definition of system failures is more subtle in complex infrastructure SoS as there are generally two types of system failures: physical (hard) failure and functional (soft) failure. The definition of physical failure is straightforward, but the definition of functional failure tends to be ambiguous, which usually depends on the context of the problem. Physical failures in complex infrastructure SoS do occur, especially during natural disasters. Using a bridge infrastructure system as an example, physical failure usually means broken major bridge elements or even the collapse of bridge. However,

understanding the more complex mechanism of functional failures is the focus of this dissertation. One of the major functional failures is bridge deficiency which is a form of failure we are investigating. Bridge deficiency usually doesn't cause bridge collapse directly, but they reduce the safety margin of the bridge and accelerate the bridge deterioration process.

Identifying different failure modes is a first step in approximately understanding how a system fails. This process is usually based on knowledge, experience, and expert judgment. The purposes of identifying failure modes are to (i) develop perspectives and organize categories for further precursor identification; and (ii) select the most efficient risk management actions according to the way the system fails.

The next step in the analysis is to identify system constraints for each failure mode. This task is achieved by expressing a failure mode as relationships among state variables and other building blocks of the system, such as a set of quantitative equality or inequality constraints among state variables. If a failure mode needs to be expressed by more than one constraint, a decomposition method can be used to describe the Boolean logic relationships among these constraints. The logical relationship among failure, failure modes, and system constraints can be represented in Figure 5-2. To clarify, Figure 5-2, although resembling a fault tree, it is not used to physically decompose the system itself as in reliability analysis, but to describe the logical relationships among a set of system constraints.

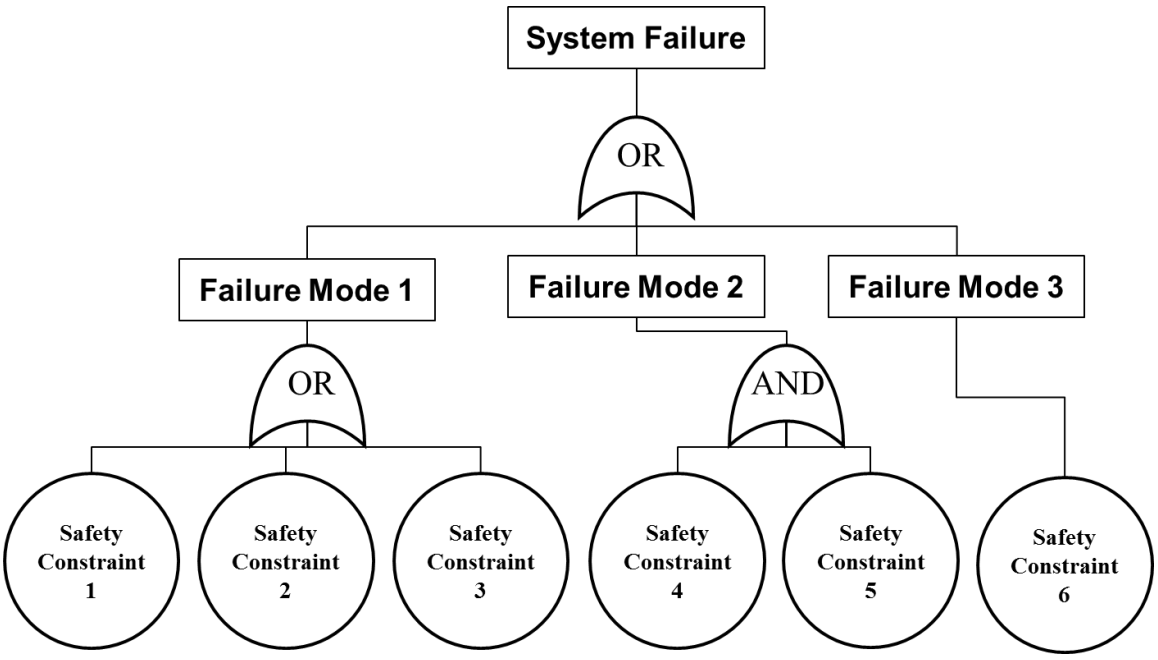

FIGURE 5-2. A LOGICAL REPRESENTATION OF THE RELATION AMONG FAILURE, FAILURE MODES, AND SYSTEM CONSTRAINTS

5.2.2. Meta-Modeling System Control Structure

In the previous section we showed how to identify system constraints for each system failure mode. These system constraints determine the safe operation boundaries of the system, and any well-designed engineering systems, regardless of their structures and purposes, should contain a control mechanism to enforce these constraints. In most cases, a feedback control system with single or multiple control loops is used for the following reasons:

- The process to be controlled is complicated and the control model is only an approximation of the actual process.
- Random noises and perturbations cannot be ignored.
- There are uncontrolled inputs or couplings between the subsystems.
- New information about system states can be observed or estimated over time.

There are five major functional components in a typical feedback control system; they are the controlled process, process model, actuator, sensor, and controller. The process model is a mental or analytical model of the controlled process, and it describes the dynamics of the system. The actuator provides controls to the process, and the sensor provides real time information about the state of the system. The controller aims to achieve some control objectives based on the process model. Each functional component may consist of hardware, software, human/organization, and procedures. Identifying these functional components enables us to understand how the system constraints are enforced in the system and to develop precursors for each failure mode. A general control structure of an engineering system is shown in Figure 5-3.

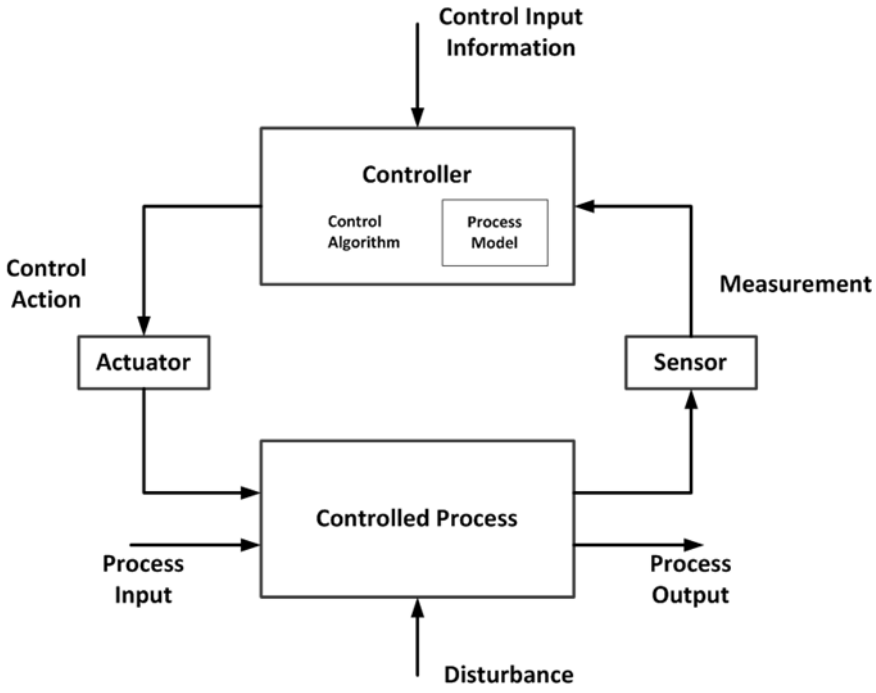

FIGURE 5-3. A GENERAL CONTROL STRUCTURE OF AN ENGINEERING SYSTEM

Physical infrastructure SoS usually contain multiple interconnected and intra- and interdependent subsystems with multiple functions, operations, and stakeholders. When modeling infrastructure SoS based on this control structure, multiple control loops with respective decision makers may be developed for each subsystem, and shared state variables are identified to represent the interdependencies among subsystems. Consider a bridge infrastructure SoS with two subsystems: the maintenance subsystem and the traffic engineering subsystem. Figure 5-4 describes the control structure of these two subsystems, with the deck conditional rating as the shared state variable between the two subsystems. More details about this bridge model are discussed in Chapter 6.

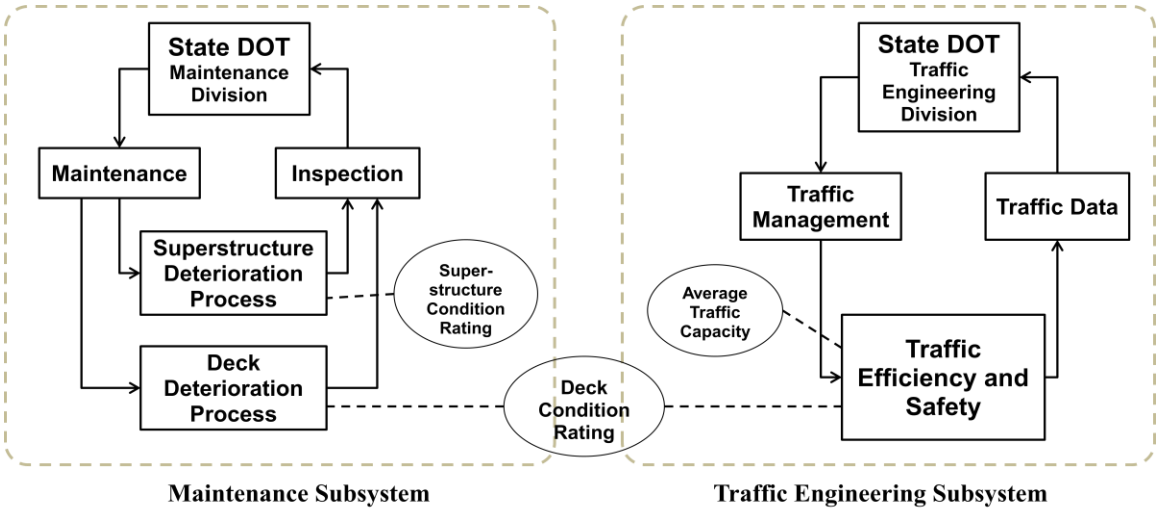

FIGURE 5-4. CONTROL STRUCTURE OF A BRIDGE SOS WITH TWO SUBSYSTEMS

The identified control structure in Figure 5-4 is a qualitative system model that can be used to identify failure precursor scenarios, which is the topic of section 5.3. However, evaluating and predicting the likelihood of system failure requires developing a quantitative system model so that analytical analysis or numerical simulation can be performed. In bridge systems, various models have been developed to model the behaviors of each individual functional component.

For example, superstructure and deck deterioration models include physicochemical models [Thoft-Christensen, 2000], Markovian models [Chase & Gáspár, 2000], and statistical regression models [Chase *et al*. 1999]. The effectiveness of bridge maintenance can be modeled using a capital budgeting model [Al-subhi et al. 1990] or empirical model [Testa & Yanev 2002]. The effectiveness of bridge inspection can be modeled using reliability models [Phares *et al*. 2000] and models provided by the inspection equipment manufacturers. The bridge maintenance decision process can be modeled as a probabilistic lifetime-oriented multiobjective optimization problems [Neves *et al*. 2006], Markovian models [Scherer & Glagola 1994], and dynamic optimization model [Jiang & Sinha 1989]. In practice, bridge maintenance decisions are often made based on computerized maintenance management systems such as Pontis [Thompson *et al*. 1998]. Bridge load capacity can be modeled using simple beam models or finite element models [Mabsout *et al*. 1997]. Traffic and live load can be modeled statistically using Weight-In-Motion data [Nowak, 1993]. The traffic capacity reduction around work zones (bridge maintenance project) can be estimated using multiple regression models [Kim *et al*. 2001] and neuro-fuzzy logic model [Adeli and Jiang, 2003]. And finally, the impact of traffic engineering can be estimated by time-series and neural network models [Smith & Demetsky, 1997], cell transmission models [Daganzo, 1994] and agent-based models [Burmeister *et al*. 1997] [Balmer *et al*. 2008].

The meta-modeling approach can take advantage of these existing sub-models of components without starting from scratch. The major task of meta-modeling is not developing sub-models for individual components, but identifying the interface between connected component sub-models in order to coordinate and integrate multiple sub-

models in a system for the purpose of better understanding and modeling the system as a whole. The concurrent use of a multi-disciplinary set of tools and models to explore the interactions between the various functional components of the system help us to examine the operation of the bridge infrastructure from various perspectives and can help us gain insights into robust bridge infrastructure management and emergent risks better than any single-model approach can [Andrijcic, Chase, Guo, and Hwang 2012].

In cases where existing models are not available, sub-models can be developed through eliciting expert evidence, performing experiments, or taking advantage of existing databases. The process of using existing databases to estimate model parameters is called system identification, which is a more objective and efficient way to model the system behaviors. System identification has been studied over the past decades and various techniques are available for different applications [Ljung, 2010]. Although it is not the purpose of this dissertation to discuss specific techniques in system identification, we will try to demonstrate in a working paper [Haimes, Andrijcic, and Guo] that being able to observe the shared states improves the results of system identification for both connected subsystems.

This system control meta-model has several advantages in analyzing the cause of system failures. It provides guidance to the analyst on what information to collect and how to integrate them in a systemic and logic way. In other methods such as Bayes Networks, the selection of the information source is arbitrary which might miss important system dynamics. This model is also flexible in combining different types of information, including sensor measurements in an automated monitoring system and subjective expert judgments. We acknowledge that some conditions of a real world engineering system are

difficult or even impossible to be quantified, for example, the effectiveness of bridge maintenance projects on improving the actual condition of bridge elements. However, this model is still able to provide a basis for estimating the likelihood of system failure based on system objectives, functions, and their causal relationships. The purpose of this model is not to predict the probability of future system failures. Instead, it creates the basis through which to compare the impacts of different precursors under the same environment, and to enable a precursor filtering and prioritization process which will be discussed in section 5.3.

5.3. Precursor Identification, Filtering and Prioritization

A complex infrastructure SoS may fail due to different reasons. Risk management resources should be allocated for different failure modes according to their likelihood and urgency. The goal of precursor identification, filtering and prioritization is to generate a manageable set of most important precursors in a systemic and justifiable way for further monitoring. The system functional components and the control structure identified in section 5.2.2 provide an insight and a systematic approach to identify different precursor scenarios of likely system failure. However, the number of resulting precursor scenarios from this process is usually very large. Hierarchical Holographic Modeling (HHM) [Kaplan, Haimes, and Garrick, 2001], [Haimes, 2009] and Risk Filtering and Ranking Method (RFRM) [Haimes, 2009] are two powerful tools to organize and prioritize the identified precursor scenarios. The output of this phase is usually a manageable set of precursors which can be effectively monitored by a monitoring system.

5.3.1. Precursor Identification

Each functional component in a generic feedback control system may malfunction in certain ways, which may lead to the failure of the entire system. Thus, a possible defect or deviation in a functional component or control structure can be seen as a precursor to system failure. Common types of defects for each functional component are summarized by Leveson (2004) as:

- Controller
    - Control input of external information is wrong or missing
    - Inadequate control algorithm
    - Flaws in creation, process changes, incorrect modification for adaptation
    - Missing or wrong communication with another controller
    - Conflicting control actions
- Process model
    - Inconsistent, incomplete, or incorrect process model is used
- Actuator
    - Inappropriate, ineffective, or missing control action
    - Inadequate operation
    - Delayed operation
- Sensor
    - Measurement inaccuracies
    - Inadequate or missing feedback
    - Feedback delays

    - Inadequate operation
    - Incorrect or no information provided
    - Feedback delays
- Controlled process
    - Component failures
    - Process changes over time
    - Unidentified or out-of-range disturbance
    - Process input missing or wrong
    - Process output contributes to system hazard

Based on the results in Chapter 4, we consider the following additional causes:

- Controller
    - Competing control/decision objectives with inappropriate tradeoffs
    - Multiple controllers and their unknown interactions
- Controlled process
    - Multiple processes with unknown interdependencies

The common failure mechanisms above can be used as a checklist for each functional component identified in the system control structure in section 5.2.2. The question to be asked in this check process is: *Is there any evidence or available information indicating the existence of a (certain type of) defect in a functional component?* If evidence shows that certain defects may exist for a functional component, this can be considered as a precursor scenario to system failure. For example, in a bridge system, bridge inspectors serve as a sensor that detects the actual condition of bridge elements and feeds back to the decision makers. If evidence shows that due to the

limitation of visual inspection (which is a common inspection method) there is a large discrepancy or uncertainty between the true condition of the bridge elements and the reported condition, this inaccurate information may mislead the decision maker to make the suboptimal decision, delay necessary bridge maintenance, and eventually lead to the functional failure of the bridge. In that case, this evidence suggests that a precursor of insufficiency/uncertainty may exist in the visual inspection of a bridge. As we proceed with the checklist for all the functional components in the identified system control model, a comprehensive set of precursor scenarios leading to system failure can be identified.

The difference between the terms a precursor vs. a precursor scenario is whether it would significantly increase the probability of future system failure. A precursor scenario connotes any deviation from the normal operations of a functional component of the system; however, it may or may not actually significantly increase the assessed likelihood of system failure, because the feedback mechanism may work to compensate for some of the defects in the system. A projection of the system failure probability using a quantitative model will be used for the following filtering and prioritization process.

Investigations of industrial accidents reveal that it is very rare that a complex engineering system fails due to a single cause [Perrow, 1984]. In most cases, it is the combination of multiple causes with their unexpected interaction that leads to the failure of the system. So in the precursor identification process, we should also consider i) common functional components and/or common precursor scenarios under multiple failure modes; and ii) the combination of multiple precursor scenarios leading to system failure. This process requires both expertise in the specific field and out-of-the-box

thinking. The HHM structure provides an intuitive way to organize all identified precursors in a systemic way. As system failures are caused by the violation of system constraints under each failure mode as shown in Figure 5-2, it is natural to arrange these precursor scenarios under each failure mode and the functional components of the control structure, as shown in Figure 5-5. This graphical representation of precursor scenarios is useful to ensure that all aspects of the system are systematically explored and considered.

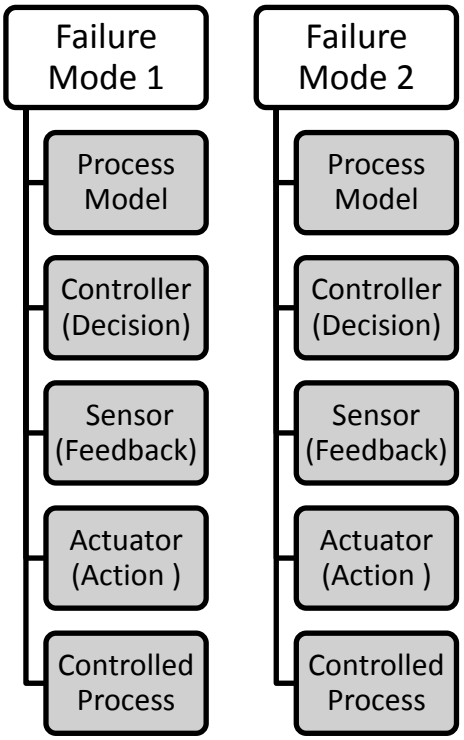

FIGURE 5-5. HHM HEAD TOPICS FOR COMPLEX INFRASTRUCTURE SOS

In the above HHM, each failure mode is listed as a head topic, and functional components of the control structure constitute the subtopic under each head topic. Then, different precursor scenarios that have the potential to lead to system failure can be organized under each functional component. The purpose of organizing precursor scenarios in a structured way is to safeguard against missing critical precursors and to

provide an added assurance that the proposed methodology captures critical precursors if and when new acquired knowledge about the system signals the emergence of new or heretofore undetected system behaviors.

HHM has been used to structure risk scenarios in probabilistic risk analysis. Although similar in appearance, the system model serving as the "backbone" of the HHM in precursor scenario identification is different. In developing risk scenarios, the system model leading to an HHM usually decomposes the system based on its multiple perspectives. In this dissertation, the system model supporting HHM constitutes the identified system functional components and the safety control structure developed in section 5.2.2.

5.3.2. Precursor Filtering and Prioritization

The process of precursor identification may generate a large number of precursors that have the potential to cause system failure. However, it is practically not feasible for a monitoring system, either automatic or human operated, to monitor and track all these precursors due to the limited resources it has. On the other hand, not all the identified precursors have the same likelihood and urgency to cause system failure. The purpose of precursor filtering and prioritization is to select a manageable subset of precursors for the monitoring system.

A model-based quantitative bi-criteria filtering and prioritization process is used based on the system control meta-model developed in section 5.2.3. The two criteria we use are the likelihood and urgency of future system failure given the precursor, which are defined as:

- Likelihood: *is the system failure probability at a specific future time given the existence of a precursor*
- Urgency: *is the time needed for a system to reach a specific failure probability given the existence of a precursor*

The likelihood criterion of a precursor is a measure of the maximum conditional probability of system failure within a specific time domain given the existence of a precursor. The urgency criterion of a precursor is a measure of the time to reach a specific failure probability given the existence of precursor. Each precursor is evaluated by these two criteria respectively and the precursors with higher likelihood and/or higher urgency receive higher priority for further monitoring. This precursor filtering and prioritization process is not limited to the above criteria. Additional criteria that matter to decision makers should be added when necessary. For example, the likelihood of dire consequences with low likelihood of failure is another important aspect to be considered if the decision maker is interested in mitigating risks from extreme events. However, an increase in the number of criteria would make the evaluation process and graphical presentation more difficult, and certain rules would have to be used to combine the results from multiple criteria.

The system control meta-model developed in section 5.2.2 describes the system behaviors under the business-as-usual scenario. The key step in evaluating each precursor scenario is to translate the difference in that precursor scenario (compared to the business-as-usual scenario) into the changes in system structure, parameters, or states, such that the system behaviors under that precursor scenario can be simulated and

compared with the business-as-usual scenario. We use an example to demonstrate this process.

The system control model in section 5.2.2 in Figure 5-4 is an ideal model without considering the practical issues of each functional component. For example, it is assumed that the bridge inspection team that functions as a sensor within the control loop is able to measure the true deterioration state of the bridge and provide unbiased and precise condition rating to the decision maker. However, the true state of the superstructure is usually not directly observable and current practice relies heavily on visual inspections due to limited inspection resources. This practice increases the bias and uncertainty in the quality of the observed bridge states and may affect the failure probability of the bridge in the long run. To evaluate the impacts of this precursor scenario on the likelihood and urgency of bridge failure, we incorporate the uncertainties in bridge inspection by adding another random variable $\omega(k)$. The mean of $\omega(k)$ represents the accuracy of the inspection results and the variance of $\omega(k)$ represents the precision of the inspection results. Based on a study by Phares et al. [2004], the mean and standard deviation of the superstructure inspection results is +0.5 and 0.8 respectively. The actual observed states incorporating inspection uncertainty becomes $s_1(k) + \omega(k)$, thus the optimal decisions are adjusted accordingly. This modification enables us to compare an inspection uncertainty scenario with the business-as-usual scenario through the evaluation of the two criteria by either analytical methods or numerical simulation using the model in Figure 5-4. A Monte Carlo simulation can be used to calculate the maximum probability of failure. Within each iteration the simulation calculates a sequence of 25 optimal decisions for a 50-year projection period (one decision every two years). At each decision period,

the values of the state variables are checked to see if system safety constraints are violated. If yes, a failure is recorded. By running the simulation for multiple times, the failure probability for each decision period can be estimated as a parameter for the binomial distribution. Then the highest failure probability among the 25 decision period is used as the maximum probability in the 50-year projection period. The following Chapter 6 will discuss in detail an example model formulation on highway bridges.

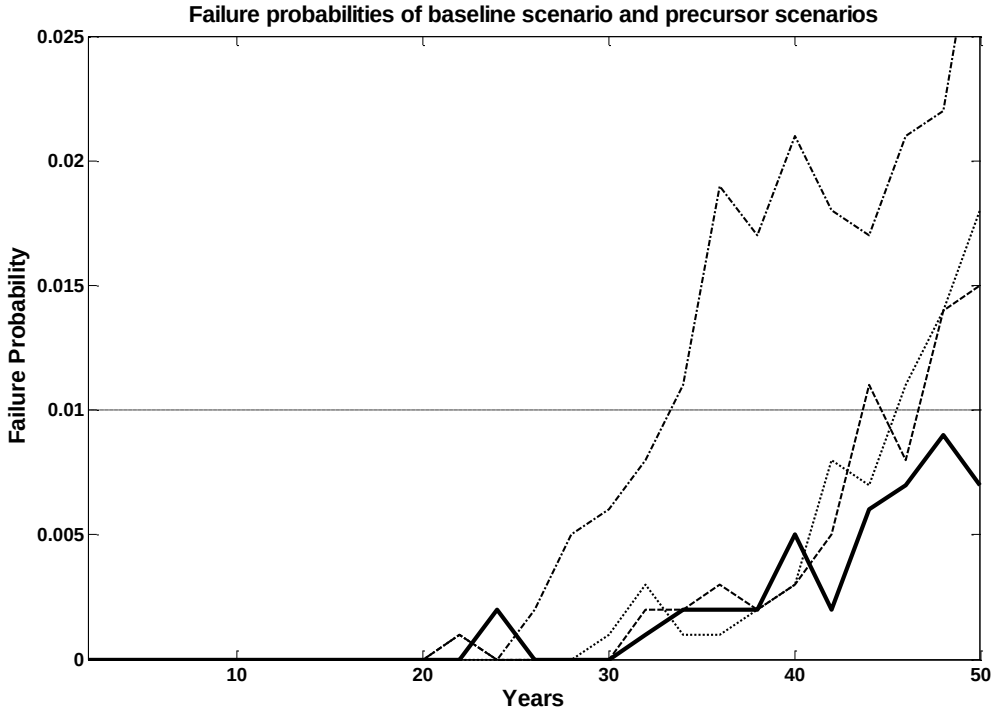

FIGURE 5-6. COMPARISON OF SIMULATION RESULTS OF THE ESTIMATED SYSTEM FAILURE PROBABILITIES BETWEEN THE BASELINE SCENARIO (SOLID LINE) AND INSPECTION ERROR SCENARIOS (THREE DOTTED LINES)

Figure 5-6 shows typical simulation results of the baseline scenario (solid line) and three precursor scenarios (dotted lines). The maximum probability of system failure within 50 years and the time to reach a failure probability of 0.01 is used to compare

these scenarios. As the figure shows, different precursor scenarios have different levels of impacts on future system failure in terms of its likelihood and urgency, and a filtering a ranking process is necessary to identify important precursors for monitoring.

Quantification of the likelihood and urgency of other identified precursors will be discussed in details in the case study in Chapter 6.

After quantifying all identified precursors against both criteria, we can plot all precursors in a two-dimensional figure as shown in Figure 5-7. The x-axis represents the time to system failure given the observation of the precursors and the y-axis represents the likelihood of system failure given the observation of the precursors. This bi-criteria figure divides all precursors into four different groups. The precursors at the top-left corner have higher likelihood and urgency so they must be monitored continuously. The precursors at the top-right corner have higher likelihood but lower urgency, and the precursors at the bottom-left corner have lower likelihood but higher urgency. These are more selective precursors to be monitored if resources are available. The precursors at the bottom-right corner have lower likelihood and lower urgency so they can be dismissed if detecting and monitoring resources are very limited.

This section concludes with two messages. First, the model and simulation used for the screening process is not a prediction tool to predict the actual failure probabilities of the system. We acknowledge that there are so many factors and uncertainties leading to system failure that cannot be included in this simple model. Quantifying failure probability is not our task for this analysis. Instead, we hope that through risk management based on this analysis, the decision makers are able to perform risk mitigation actions to actively reduce the probability of future system failure. Secondly,

this filtering and prioritization is an iterative process. Whenever new knowledge about the system or new threat to the system is discovered, the process should be conducted again to ensure that the resulting precursors are updated.

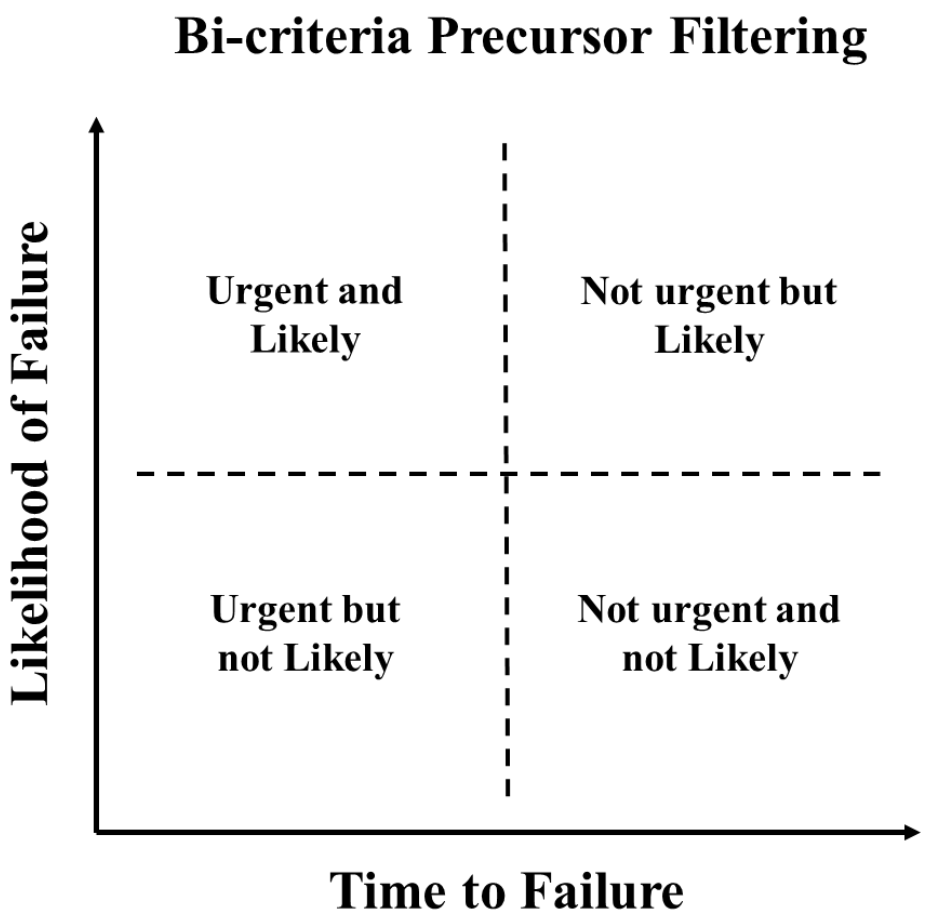

FIGURE 5-7. TWO-DIMENSIONAL FILTERING OF PRECURSORS BASED ON THE LIKELIHOOD AND TIME TO FAILURE

5.4. Precursor Detection and Evaluation

5.4.1. Precursor Detection

The capability to detect precursors – whether performed by people, an automatic process or some combination thereof – is a critical function of a successful precursor monitoring system. If identified precursors cannot be detected in a timely manner and within certain confidence levels, their contributions to risk management will be significantly limited.

The detection of precursors depends on the definition and characteristics of each precursor and there is no standard procedure for it. Some precursors have well defined specifications such that the detection is trivial. For example, the precursor of limited funding can be observed whenever the bridge maintenance funding is less than what is required.

However, in many other cases the detection of precursors requires careful analysis and design. In general, we are interested in detecting whether certain conditions or relationships exist or not among state variables and other building blocks of the system such as inputs, outputs, and decisions. An example of detecting rebar corrosion will be described in detail in section 6.4.3.

Regardless of the technologies used in precursor detection, quantifying and understanding the impacts of uncertainties in detection results is an important issue. It is not the purpose of this dissertation to develop or review state-of-the-art detection and estimation technologies; instead we believe that the process of precursor detection should be subject to uncertainty analysis, and failure to consider these impacts will compromise

the applicability of the proposed precursor analysis method. This understanding helps to apply the precursor-based risk assessment tool on a real working environment where noises and errors are inevitably embedded in the detection process.

We use the term "*evidence*" to refer to the supporting information in measuring conditions for the existence of precursors. For example, in the results of physical sensor measurement, the evidence refers to the precision and accuracy of the measurements. Measurements with high precision and accuracy indicate strong evidence that certain precursors exist; while measurements with low precision or low accuracy indicate a weak evidence for the existence of certain precursors. Evidence also refers to the signal to noise ratio in detection systems or false positive/false negative rate in the results of predictive models. In the results of statistical models and information elicited from experts, evidence usually has the form of distribution, variance, confidence intervals, upper and lower bounds, or other statistical measures. Regardless of the form, the key component in a piece of evidence is the contextual information for detecting a specific precursor and its quantified uncertainty.

In precursor detection process, we are interested in knowing the probability of system failure given an observed piece of evidence: Pr*(system failure within a specific time domain | Evidence)* and the expected time-to-failure given evidence T*(system failure with a specific probability | Evidence)*, which can be calculated using the theorem of total probability as:

EQUATION 5-1

$$\Pr(system\ failure\ within\ a\ specific\ time\ domain \mid Evidence)$$
$$= \Pr(system\ failure\ within\ a\ specific\ time\ domain \mid Existence\ of\ Precursor)$$
$$\cdot \Pr(Existence\ of\ Precursor \mid Evidence)$$
$$+ \Pr(system\ failure\ within\ a\ specific\ time\ domain \mid NonExistence\ of\ Precursor)$$
$$\cdot \Pr(NonExistence\ of\ Precursor \mid Evidence)$$

and

EQUATION 5-2

$$\text{Time}(system\ to\ reach\ a\ specific\ failure\ probability \mid Evidence)$$
$$= \text{Time}(system\ to\ reach\ a\ specific\ failure\ probability \mid Existence\ of\ Precursor)$$
$$\cdot \Pr(Existence\ of\ Precursor \mid Evidence)$$
$$+ \text{Time}(system\ to\ reach\ a\ specific\ failure\ probability \mid NonExistence\ of\ Precursor)$$
$$\cdot \Pr(NonExistence\ of\ Precursor \mid Evidence)$$

Thus, the task of the detection phase is to identify the information to be collected and to quantify the uncertainties in the detection process using detection probability Pr(Precursor | Evidence). Depending on the type of the system and the information to be collected, the tools to perform this task and the methods with which to quantify the uncertainties in the detection process vary widely. The total probability rule used in

Equation 5-1 and Equation 5-2 incorporates the uncertainties in precursor detection into the projection of system failure given the existence of precursors.

As an example, the existence of a precursor can be defined as the level that a certain state variable s is above a predefined threshold:

EQUATION 5-3

$$\Pr(Existence\ of\ Precursor) = \begin{cases} 1 & s \geq threshold \\ 0 & s < threshold \end{cases};$$

The measurements (evidence) we have show that this state variable can be modeled as a random variable with known distribution, for example, a normal distribution with mean u and variance σ^2, thus $s \sim N(u, \sigma^2)$. Then it can be calculated that

EQUATION 5-4

$$\Pr(Existence\ of\ Precursor \mid Evidence) = \Pr(s \geq threshold \mid s \sim N(u, \sigma^2))$$

and

$$\Pr(NonExistence\ of\ Precursor \mid Evidence) = \Pr(s < threshold \mid s \sim N(u, \sigma^2))$$

Although we demonstrate above how to calculate the value of failure probability after the detection of a precursor, it is not the purpose of this dissertation to quantify or predict a single value of bridge failure probability in the future. Instead, we are interested in comparing the likelihood of each failure mode after multiple precursors are observed. This process is called precursor evaluation and will be discussed in the following section.

5.4.2. Precursor Evaluation

It is commonly believed, especially in a post-accident risk assessment, that using precursors to predict potential accidents are straight forward. In such a retrospective

view, whenever a precursor is detected, it can be used to prevent the accident from happening. This phenomenon has been known as hindsight bias [Fischhoff, 1975], [Hawkins and Hastie, 1990]. However, in the traditional pre- accident risk assessment, warnings of accident solely based on a single precursor are quite unreliable and the effectiveness of using precursors for risk management is quite limited. In the previous sections we discussed how to quantify the uncertainties in precursor detection and how a model-based approach can be used to evaluate the failure probability before and after the detection of precursors. In this section, we discuss approaches to address another major contributing factor to hindsight bias, which is the one-to-many relationship between a precursor and various failure modes of the system. In a pre-accident assessment the observation of a specific precursor may increase the likelihood of multiple failure modes simultaneously (the increase in the likelihood of one failure mode doesn't necessarily reduce the likelihood of the other failure modes), while in a post-accident assessment extra information is given on which failure mode has actually occurred which doesn't objectively represent the situation faced by decision makers before system failure. Any successful risk management action depends on the understanding of what the real cause is and what failure mode is more likely.

When precursors are detected, the likelihood of each system failure mode needs to be quantified, evaluated, and compared to improve the situation awareness of the decision makers and assist them to take appropriate course of actions. Solutions to this problem depend on (i) the detection and integration of multiple, independent precursors; and (ii) providing quantitative information on the likelihood of all possible failure modes to the decision maker to improve contextual understanding and interpretation of the situation in

order to perform a more informed risk management process. This dissertation treats the precursor evaluation as a distributed detection problem. It fuses the information from multiple precursors at different time stages using a model-based approach. Inevitably all the solutions above increase the cost of the precursor analysis program so the cost-benefit analysis of such solutions may deserve further investigation. However, estimating cost of these solutions is beyond the scope of this dissertation and this analysis is not included here.

In many cases, multiple precursors or signs of system failure can be observed well before the actual failure occurs. For example, before the I-35W Mississippi River bridge collapse accident on August 1, 2007, a set of evidence has indicated the poor condition of the bridge, including [Weeks, 2007], [NTSB, 2007], [NTSB, 2008]:

- In 1990, significant corrosion in bearings was found
- In 2001, U. Minnesota civil engineering dept. report cracking in the cross girders and lack of redundancy
- In 2005, rated as "structurally deficient", in possible need of replacement (scoring 50)
- In 2006, inspection found problems of cracking and fatigue
- Inspection not performed in 2007 due to construction work
- In December 2006, a steel reinforcement project was planned for the bridge. However, the project was canceled in January 2007 in favor of periodic safety inspections.
- In internal Mn/DOT documents, bridge officials talked about the possibility of the bridge collapsing and worried that it might have to be condemned

- Prior to the collapse, there were 575,000 pounds (261,000 kg) of construction supplies and equipment on the bridge

When each precursor contains independent information (or at least partially independent information) regarding the states of the system, integrating these multiple precursors will provide us with a better understanding of future system behaviors than relying on any single precursor.

The system control meta-model developed in Figure 5-4 in section 5.2.2 can be readily used to incorporate new information whenever a precursor is detected along the timeline to fuse the information from multiple precursors and evaluate the likelihood of multiple failure modes. Instead of calculating the failure probability of the whole system that is the top event in Figure 5-2, the system control meta-model can calculate the failure probability of each failure mode, which is the intermediate event under the top event. Uncertainties in the precursor detection and projection phases are estimated and updated over time for each failure mode. When the trend of likelihood of each failure mode and the likelihood of no failure is displayed along the timeline when each precursor is detected, the decision maker may have better understanding of the current situation and possibly the root cause of the situation. Risk management actions targeted on a specific failure mode can be developed and evaluated. A hypothetical example is used to demonstrate the precursor evaluation process.

To summarize this chapter, the precursor detection and evaluation is an online process after a monitoring system is established and a set of high priority precursors are identified. This process is able to incorporate the uncertainties in the detection results and estimate the failure probability accordingly. When multiple precursors are detected, this

process is also capable to fuse the information from them and readjust the failure probability along the timeline. Besides that, the failure probability and its associated confidence interval of each failure mode can be estimated respectively such that improved situation awareness can be achieved for decision maker to identify root causes and select appropriate risk management actions.

6. A Case Study on Bridge Infrastructure Systems of Systems

6.1. Bridge Infrastructure as Systems of Systems

This chapter demonstrates the theories and methodologies developed in this dissertation with a case study on the US highway bridge infrastructure systems, with the objectives to i) explain the theories and methodologies in this dissertation through a real-world case study; ii) provide real-world evidence to support the validity of assumptions and conditions used in the theories discussed in the early chapters; and iii) demonstrate the applicability, effectiveness, and efficacious contributions of the theories and methodologies to the risk management of bridge infrastructures. The aging infrastructure issue is widely known as a complex problem transcending multiple domains including engineering, economy, social wellbeing, environment, and politics. The complexity of the highway infrastructure maintenance decision process is well recognized. It is not the purpose for the case study to solve this unsolvable complex problem with numbers; rather, the objective is to provide insights and invoke discussions from a new systems-of-

systems and risk analyses perspectives. We state the problem is unsolvable because we believe all models are simplification of the complex real world. Although part of the analysis is quantitative, the numerical results are dependent on our assumptions in the modeling process and not necessarily indicate a solution to the original problem. It is hoped that this new approach will add one more tool in the decision makers' toolbox for them to better cope with the challenges from infrastructure management.

The highway bridge infrastructure is an essential element of transportation networks. The condition of highway bridges is continuously deteriorating due to the lack of appropriate maintenance, with 26% of America's bridges are structurally deficient or functionally obsolete [ASCE, 2009a]. Many bridges receive insufficient inspection and maintenance due to limitations of funding, equipment, manpower, and available technology. The bridge infrastructure system has multiple subsystems which constitutes multi-scale infrastructure systems. This study considers bridge infrastructure as a complex engineering system of systems with broad social and economic impacts from bridge failure. Thus, assessing and quantifying the risks associated with projected traffic load, environmental factors, and other natural and human-induced emergent forced changes are critically important. This modeling framework for risks of bridge system failure allows examining the impacts of current bridge inspection and maintenance practices on the overall reliability of a bridge infrastructure SoS; enables decision makers to make more timely and informed decisions to efficiently allocate limited risk management resources; and thus, prevent future severe consequences. The developed methodology is expected to help bridge owners to efficiently prioritize and plan for

inspection, maintenance, and rehabilitation activities based on precursors, and to reduce the risk of bridge failure.

All the characteristics of a complex SoS are summarized in section 4.1 exist in the bridge system of systems, e.g., (i) multiple stakeholders such as DOT, bridge users, inspectors, and constructers; (ii) each stakeholder makes its decisions according to its goals and objectives; (iii) even within DOT, there exist different divisions with different goals. (For example, the goal of maintenance division is usually to ensure the safety and reliability of the bridge itself, while the goal of traffic engineering division is to improve driver's safety and accessibility.); (iv) divisions operate under limited total budget of DOT, and this common constraint makes them interdependent subsystems. Such interdependencies are common among subsystems; (v) decisions made by one subsystem have direct or indirect impacts on other subsystems. For example, de-icing in winter improves driver safety. However, the de-icing chemicals will accelerate the deterioration process of the rebar in reinforced concrete and reduce the load capacity of bridge over time. Due to our limitations in fully understanding these interdependencies between subsystems, those uncertainties manifest themselves as unexpected perturbations to the connected subsystems. The deterioration process is nonlinear in nature, as well as material behavior beyond the elastic limit. This is a major reason that bridges suddenly collapse and most of time unexpectedly. Finally, (vi) due to the complexity and uncertainty of the system dynamics, decision makers need to adjust and adapt their decisions based on current states of the system, in addition to a long range plan such as life cycle management. Thus, bridge infrastructure must be viewed and modeled from a complex SoS perspective.

6.2. Modeling Bridge Infrastructure Systems of Systems

A typical highway bridge consists of many physical elements, including deck, superstructure, substructure, and other auxiliary elements as shown in Figure 6-1. The Deck is the roadway portion of a bridge, including shoulders. The superstructure consists of the components that actually span the obstacle the bridge is intended to cross. The substructure consists of all parts that support the superstructure.

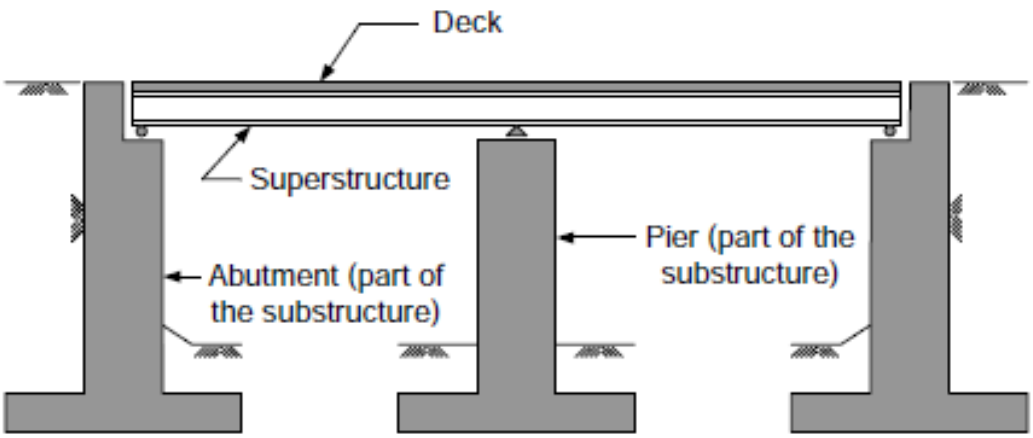

FIGURE 6-1. MAJOR BRIDGE COMPONENTS [NBIS, 2006]

Bridges are traditionally considered merely as physical systems from an engineering perspective. However, if we consider bridges as part of an overall transportation infrastructure system of systems, we must address the various functionalities, objectives, stakeholders and decision makers of the encompassing system, with the broader social and economic implications. The bridge infrastructure system typically has the following functional components:

- The physical bridge includes all bridge physical elements such as deck, beam, bearing, abutment, and pier. Research in this area mainly focuses on understanding the corrosion and deterioration process of materials under various environmental conditions.
- Bridge inspection and monitoring functions provide information on current states of the bridge to state DOT. It operates under the budget allocated by DOT, and DOT uses the information as basis for its decision-making.
- Bridge maintenance functions perform actual maintenance and repair activities on bridge. Their work quality determines the effectiveness of DOT's decision.
- Bridge management functions include bridge owners and decision makers with the goals to maintain the reliability and functionality of the bridge. State Department of Transportation (DOT) is the owner of most bridges in the states. DOT makes decisions on planning, design, operation, maintenance, repair, rehabilitation, and possibly retirement or replacement of these structures. In many cases, these decisions delay the deterioration process in the engineering subsystem. Decisions are also made to improve traffic safety and efficiency. DOT may be further decomposed into lower level subsystems as individual functional divisions.
- Bridge users include private and commercial vehicles commuting across the bridge. The increasing traffic load and changing pattern are major contributors to material fatigue, stress, wear and tear.

Regardless of the type, structure, or material from which a bridge is built, a physical bridge in a transportation network must meet two basic requirements: the reliability of the

bridge itself and the effectiveness and safety of the traffic across the bridge. In bridge engineering, the reliability of bridge is mainly determined by the condition of the superstructure while the traffic efficiency is mainly determined by the condition of the deck. In a bridge system, the deck and superstructure conditional ratings are descriptive measures of the deterioration state of deck and superstructure. Other appraisal ratings such as structural evaluation and deck geometry can be derived from these conditional ratings. Thus, the deterioration process of deck and superstructure is one of the controlled processes of the system. The average traffic capacity is a measure of the effectiveness of the bridge to move people and commodities across the bridge, thus the changes in traffic capacity also need to be controlled. The state Department of Transportation (DOT) plays the role of both controllers, with the goal to maintain certain required conditional rating of the deck and superstructure through planned maintenance activities, and to maintain a required traffic capacity through traffic engineering. The bridge deterioration models and traffic models used by DOT serve as the process model. The bridge maintenance team is the functional unit who performs maintenance activities on bridge to improve its conditional rating, thus it play the role of actuator. The bridge inspection team provides actual bridge conditions back to DOT and functions as a sensor.

A bridge usually contains multiple interconnected and intra- and interdependent subsystems with multiple functions, operations, and stakeholders. This case study considers two basic subsystems of a bridge infrastructure SoS: the maintenance subsystem and the traffic engineering subsystem. In DOT practice, superstructure and deck are managed and maintained separately. Thus, each subsystem has its own functional components and control structure. A subsystem here consists of not only the

physical elements of the bridge, but also all the organizations with their functionalities and decisions. In DOT practice, superstructure and deck are managed and maintained separately. Thus, each subsystem has its own functional components and control structure. Figure 6-2 illustrates the functional components and the structure of the two interdependent subsystems.

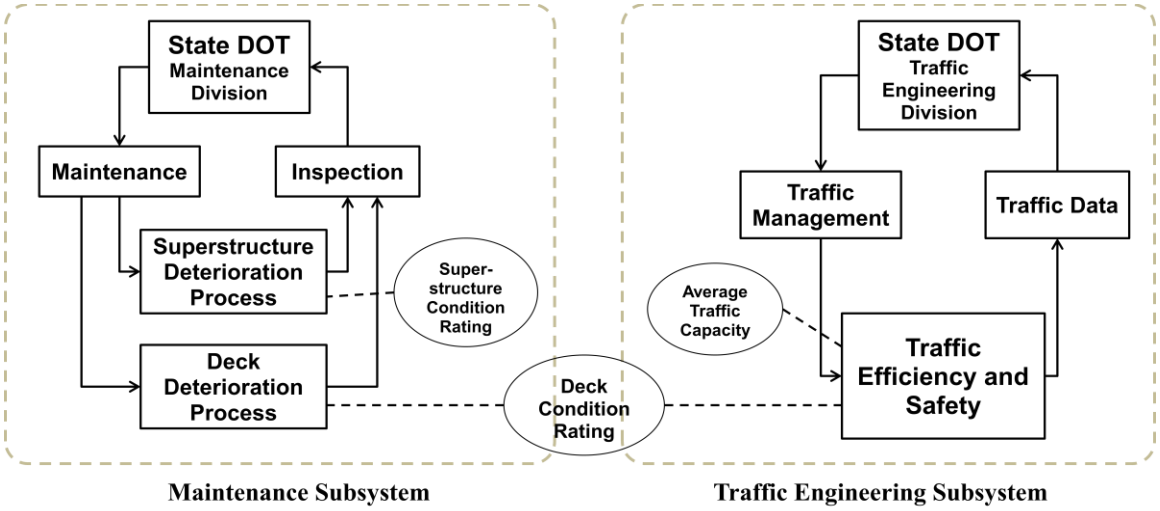

FIGURE 6-2. A BRIDGE SOS WITH MAINTENANCE AND TRAFFIC ENGINEERING SUBSYSTEMS

The decision maker of the maintenance subsystem is usually the maintenance division in state DOT, with the goal to maintain the reliability and structural integrity of both superstructure and deck at an acceptable level. The decisions they make include the maintenance spending and the type of the repair project. These objectives are achieved through performing maintenance activities on these structures, and the maintenance decisions rely on the information provided by the bridge inspection team. The decision maker of the traffic engineering subsystem is usually the traffic engineering division in state DOT, with the goal to ensure the safety and efficiency of traffic across the bridge. The decision they make range from deicing, lane allocation, work zone design, signal and

illumination placement, to speed and tonnage posting. Figure 6-2 illustrates the structure of the two interdependent subsystems. These objectives are achieved through managing traffic on bridge through different methods, and their decisions rely on the information on past, current, and projected traffic over the bridge.

However, these two subsystems are not independent, and interaction between them exists through the actual condition of the deck. Heavy traffic load accelerates the deterioration process of the deck while the maintenance activities on deck create work zones and reduce the traffic capacity of the bridge. The consequences resulting from decisions in the traffic engineering subsystem will propagate through the change in deck condition and have impacts on the superstructure. Any model of complex systems should capture these interdependencies between subsystems so that interaction between them can be understood. To achieve that, essential state variables of both subsystems should be identified, with special interest in the common state variable of both subsystems.

Three state variables are chosen to represent the essential states of the two subsystems: the condition rating of the superstructure s_s, the condition rating of the deck s_d, and the average daily traffic capacity of the bridge s_t. The maintenance subsystem has two of the state variables, the condition rating of the superstructure and the condition rating of the deck. The traffic engineering subsystem also has two of the state variables, the condition rating of the deck and the average daily traffic capacity across the bridge. The deck condition rating s_d is common to both subsystems thus a shared state variable between two subsystems. The impacts from decisions made in maintenance subsystem will propagate to the traffic engineering subsystem through the change in the condition of

the deck, and vice versa. Thus, the shared state variable is an essential factor causing interaction between subsystems, and must be captured in the modeling process.

The condition rating of the superstructure s_s and the condition rating of the deck s_d are constructed based on bridge condition rating system, which is a method of evaluating highway bridge conditions. It uses a numeric value which is indicative of bridge reliability to remain in service. The result of this method is on a scale from 0 to 9 in which 9 would represent an excellent condition bridge and 0 would represent a failed condition. The Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges [US DOT, 1995] provides instructions for the coding of condition rating for bridge elements. These instructions are summarized in Table 6-1. Each bridge element is assigned a condition rating based upon the above scale at the time of each inspection, which is usually every 2 years. Any component with a rating of 4 or less is documented in greater detail with notes and sketches. Although the condition rating is coded in integers, we will treat them as real numbers to simplify the analysis without losing its meanings. The average daily traffic capacity of the bridge $s_t$describes the traffic capacity of the bridge. It is calculated annually by dividing the yearly total traffic count by the number of days in that year.

The relationships between the bridge element's condition and time are represented by deterioration models that predict the level of a specific condition measured as a function of a bridge-element's use or wear. Several approaches have proven useful, including state-space deterioration models and statistical regression deterioration models [Chase and Gáspár, 2000]. If one assumes that bridge elements deteriorate continuously but that

the deterioration is observed and recorded periodically, then it is reasonable to use a discrete time state space approach to model the deterioration process.

TABLE 6-1. INSTRUCTIONS FOR THE CODING OF CONDITION RATING FOR BRIDGE SUPERSTRUCTURES SOURCE: [CHASE AND GÁSPÁR, 2000]

Code (1)	Description (2)
9	Excellent condition
8	Very good condition: No problems noted.
7	Good condition: Some minor problems.
6	Satisfactory condition: Structural elements show some minor deterioration.
5	Fair condition: All primary structural elements are sound but may have minor section loss, cracking, spalling or scour.
4	Poor condition: Advanced section loss, deterioration, spalling or scour.
3	Serious condition: Loss of section, deterioration, spalling, or scour has seriously affected primary structural components. Local failures are possible. Fatigue cracks in steel or shear cracks in concrete may be present.
2	Critical condition: Advanced deterioration of primary structural elements. Fatigue cracks in steel or shear cracks in concrete may be present or scour may have removed structural support. Unless closely monitored it may be necessary to close the bridge until corrective action is taken.
1	''Imminent'' failure condition: Major deterioration or section loss present in critical structural components or obvious vertical or horizontal movement affecting structural stability. Bridge is closed to traffic but corrective action may put back in light service.
0	Failed condition: Out of service. Beyond corrective action.

To simplify the discussion without losing generality, a nonlinear dynamic state-space model is used to model the above two subsystems instead of a meta-model with existing models for each functional component. The nonlinear dynamic state-space model can be seen as an integrated meta-model. In the bridge example, the meta-models for many components such as actuator and sensor are simply linear static models and the model

parameters/coefficients contain all the information about the component. These parameters are estimated from observed input-output data. So there is no need to deliberately make them separate meta-models. They can just be directly integrated into the nonlinear dynamic state-space model. However, if we are doing this analysis for a specific bridge and specific models are available for its components, then a meta-model for each component will be needed.

The parameters in the state-space model are estimated from the existing component models using the national level bridge data in the National Bridge Inventory (NBI). In this way, there is no need to discuss the details of each existing model. Unless we are trying to model a specific bridge using data from a specific bridge, the state-space model provides a more explicit view of the system without losing much of the fidelity.

We also assume that there is no direct interaction between superstructure condition rating and the average daily traffic capacity, and the only way they impact each other is through the shared state of deck condition rating. In addition, any decision in the maintenance subsystem has no direct impact on the average daily traffic capacity, and decision in the traffic engineering subsystem has no direct impact on the superstructure condition rating. The above assumptions guarantee that the interdependencies between the two subsystems solely depend on the shared state variable. Let $s_s(k)$ represent the condition rating of the superstructure at time stage k; $s_d(k)$ represent the condition rating of the deck at time stage k; and $s_t(k)$ represent the average traffic capacity at time stage k. The overall system can be described by Equation 6-1 and Equation 6-2:

EQUATION 6-1

$$\begin{bmatrix} s_s(k+1) \\ s_d(k+1) \\ s_t(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} s_s^2(k) \\ s_d^2(k) \\ s_t^2(k) \end{bmatrix} + \begin{bmatrix} a_{14} & a_{15} & 0 \\ a_{24} & a_{25} & a_{26} \\ 0 & a_{35} & a_{36} \end{bmatrix} \begin{bmatrix} s_s(k) \\ s_d(k) \\ s_t(k) \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \\ 0 & b_{32} \end{bmatrix} \begin{bmatrix} u_m(k) \\ u_t(k) \end{bmatrix}$$

EQUATION 6-2

$$\begin{bmatrix} f_s(k) \\ f_d(k) \\ f_t(k) \end{bmatrix} = \begin{bmatrix} (s_s(k) - c_s)^2 \\ (s_d(k) - c_d)^2 \\ (s_t(k) - c_t)^2 \end{bmatrix}$$

$$\text{s.t.} \quad u_1(k) \geq 0, u_2(k) \geq 0$$

$$for\ k = 0, \dots, T-1$$

for the decision makers of the maintenance subsystem, the objective is to

EQUATION 6-3

$$\min_{u_m(k)} \theta f_s(k) + (1-\theta) f_d(k)$$

and for the decision makers of the traffic engineering subsystem, the objective is to

EQUATION 6-4

$$\min_{u_t(k)} f_t(k)$$

$$for\ k = 0, \dots, T-1$$

where k represents the time interval between each decision period. As new information on the conditions of bridge is obtained every two years through bridge inspection, k is selected for a two-year period in this case. Let $u_m(k)$ represent the decision (total maintenance spending in million dollars) made for the maintenance subsystem and $u_t(k)$ represents the decision (total traffic engineering spending in million dollars) made for the traffic engineering subsystem; the functions $f_s(k)$ and $f_d(k)$ are the output (objective)

functions of the maintenance subsystem; and $f_t(k)$ is the output (objective) function of the traffic engineering subsystem; and c_s, c_d, and c_t are the control targets of each state. The decision maker of the maintenance subsystem aims to maintain appropriate superstructure and deck condition ratings of c_s and c_d by controlling the maintenance spending. The decision maker of the traffic engineering subsystem aims to maintain the traffic capacity of the bridge at c_t. Equation 6-1 describes the deterioration process of superstructure, deck, and traffic capacity. Coefficients a_{11} and a_{14} quantify the speed of natural deterioration of superstructure due to different environment factors. Coefficients a_{22} and a_{25} quantify the speed of natural deterioration of deck due to different environment factors. Coefficientsa_{12},a_{15}, a_{21} and a_{24} quantify the deterioration of superstructure and deck due to their interactions. Coefficient a_{35} and a_{36} describes the change in traffic capacity due to deck conditions and other factors, and coefficient a_{26} quantifies the impact of traffic on the deterioration of deck. Coefficients b_{11} and b_{21} quantify the effectiveness of maintenance activities in improving superstructure and deck conditions. Coefficients b_{22} and b_{32} quantify the effectiveness of traffic management activities in improving deck condition and traffic capacity. Some of these parameters such as deterioration coefficients can be estimated from bridge inspection data, but some parameters such as effectiveness of maintenance decision vary case by case and a nominal value is used based on expert estimations, see [Andrijcic *el al*. 2013].

The quadratic state space representation of the bridge model developed in Equation 6-1 to Equation 6-4 may look like an over-simplified model to even approximates the bridge maintenance and traffic engineering subsystems. The reason it is still useful in this analysis is that this model is not a “prescriptive” model to tell decision makers what to do

to optimize or improve bridge maintenance, but a “descriptive” model to describe what they actually do in the decision making process of bridge maintenance, and how the things actually work out. A “prescriptive” model can be well formulated and structured mathematically based on the theory or tool it uses, but a “descriptive” model is difficult, or even impossible to construct due to the complexity in the decision process with engineering, organizational, financial, social, and sometimes political considerations. There might not be a mathematical model which is capable to capture all the dynamic behaviors of real world decision making process, and any attempt to do so is a tradeoff between model complexity and fidelity. The model used in this case study is built based on national bridge inventory data, and it represents a “national average” bridge. The purpose of this case study is bridge infrastructure in general, not the prediction of the failure of a specific bridge, and a quadratic model does fit the data well. We also concern that an unnecessary increase in model complexity may mislead the readers to believe that we are solving a specific problem, instead of the purpose of the case study, which is to demonstrate the methodology.

When two subsystems are sharing one state variable, the shared state variable belongs to and contributes to the dynamics of both subsystems. Accordingly, when it has to be decomposed into two separate subsystems, the shared state variable, the deck condition rating in this case, should remain in both subsystems. From Equation 6-1, it can be shown that

EQUATION 6-5

$$s_d(k+1) = a_{21}s_s^2(k) + a_{22}s_d^2(k) + a_{24}s_s(k) + a_{25}s_d(k) + a_{26}s_t(k) + b_{21}u_m(k) + b_{22}u_t(k)$$
$$= a_{22}s_d^2(k) + a_{25}s_d(k) + [a_{21}s_s^2(k) + a_{24}s_s(k) + b_{21}u_m(k)] + [a_{26}s_t(k) + b_{22}u_t(k)] = a_{22}s_d^2(k) + a_{25}s_d(k) + z_1(k) + z_2(k)$$

where $z_1(k) = a_{21}s_s^2(k) + a_{24}s_s(k) + b_{21}u_m(k)$ and $z_2(k) = a_{26}s_t(k) + b_{22}u_t(k)$.

From Equation 6-1, the value of $s_d(k+1)$ depends on three factors:

- $a_{22}s_d^2(k) + a_{25}s_d(k)$, which is a function of $s_d(k)$;
- $z_1(k)$, which contains all necessary information from the maintenance subsystem;
- $z_2(k)$, which contains all necessary information from the traffic engineering subsystem.

A decomposition scheme based on this separation approach is shown in Figure 6-3. The shared state variable s_d remains in both subsystems. For the maintenance subsystem, we have:

EQUATION 6-6

$$s_d(k+1) = a_{22}s_d^2(k) + a_{25}s_d(k) + a_{21}s_s^2(k) + a_{24}s_s(k) + b_{21}u_m(k) + x_1(k)$$

and for traffic engineering subsystem,

EQUATION 6-7

$$s_d(k+1) = a_{22}s_d^2(k) + a_{25}s_d(k) + a_{26}s_t(k) + b_{22}u_t(k) + x_2(k)$$

with $x_1(k) = z_2(k)$ and $x_2(k) = z_1(k)$.

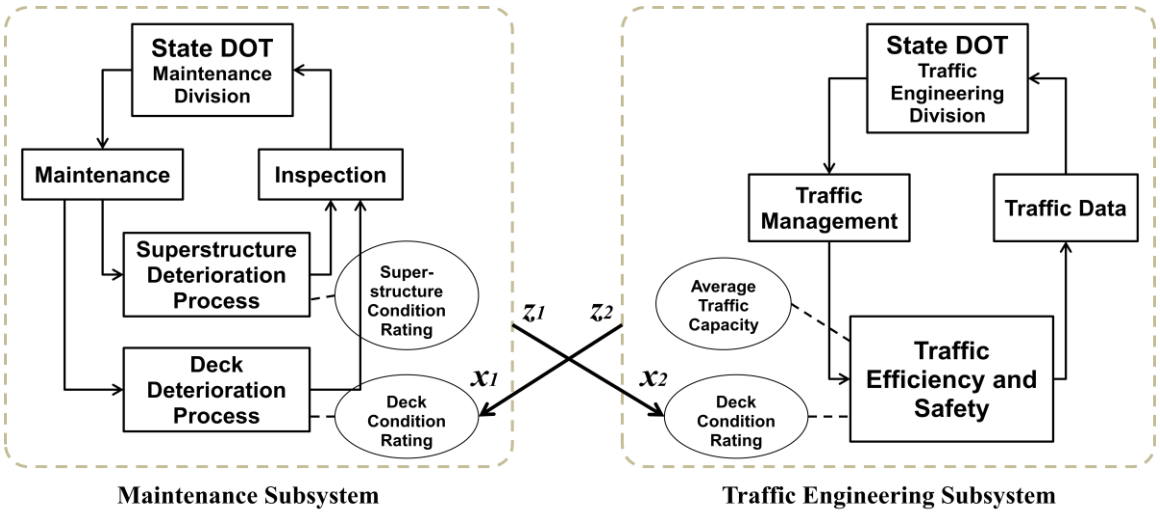

FIGURE 6-3. DECOMPOSED BRIDGE SUBSYSTEMS WITH INPUT-OUTPUT CONNECTION

The extra input x quantifies the impact from the interdependent subsystems on the shared states. When it is added to the subsystem it conveys all necessary information from the other subsystem to the shared state variable and thus the shared state variable becomes an internal state variable for both separated subsystems.

The resulting subsystem models are derived as:

For maintenance subsystem

EQUATION 6-8

$$\begin{bmatrix} s_s(k+1) \\ s_d(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_s^2(k) \\ s_d^2(k) \end{bmatrix} + \begin{bmatrix} a_{14} & a_{15} \\ a_{24} & a_{25} \end{bmatrix} \begin{bmatrix} s_s(k) \\ s_d(k) \end{bmatrix} + \begin{bmatrix} b_{11} & 0 \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} u_m(k) \\ x_1(k) \end{bmatrix}$$

$$z_1(k) = a_{21} s_s^2(k) + a_{24} s_s(k) + b_{21} u_m(k)$$

For traffic engineering subsystem

EQUATION 6-9

$$\begin{bmatrix} s_d(k+1) \\ s_t(k+1) \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_d^2(k) \\ s_t^2(k) \end{bmatrix} + \begin{bmatrix} a_{25} & a_{26} \\ a_{24} & a_{25} \end{bmatrix} \begin{bmatrix} s_d(k) \\ s_t(k) \end{bmatrix} + \begin{bmatrix} 1 & b_{22} \\ 0 & b_{32} \end{bmatrix} \begin{bmatrix} x_2(k) \\ u_t(k) \end{bmatrix}$$

$$z_2(k) = a_{26} s_t(k) + b_{22} u_t(k)$$

Subject to

$$x_1(k) = z_2(k) \text{ and } x_2(k) = z_1(k)$$

6.3. Systemic Risks in Bridge Maintenance Subsystem

Precursors to complex SoS are usually dynamic, evolving, and possibly unexpected. Bridge failure – either major structural damage or total collapse – has broad social and economic consequences. The following section discusses in detail how we use the maintenance subsystem model developed in the previous section to analysis systemic risks to the bridge maintenance subsystem.

In this section, we focus on analyzing the maintenance subsystem within the bridge SoS. The same two state variables are chosen to represent the essential states of the system: the condition rating of the superstructure and the condition rating of the deck.

The decision maker, which is state DOT in this example, usually has two objectives for the maintenance subsystem: (i) maintain the reliability and structural integrity of the bridge, which is mainly determined by the condition of the superstructure; and (ii) improve the safety and efficiency of traffic over the bridge, where the condition of the deck plays an important role. To achieve the first objective, the condition rating of the superstructure must be maintained at a certain level through maintenance activities on superstructures. To achieve the second objective, the condition rating of the deck must also be maintained at a certain level through maintenance activities on the deck. As in many cases the total maintenance budget is limited, the decision maker has to choose its preference between these two objectives to make tradeoff between bridge safety and efficiency. As federal regulations require each bridge to be inspected at least once every two years, the decision maker is able to receive information about the actual conditions of

the superstructure and deck every two years and make maintenance decisions accordingly. Although there are bridge management software tools such as BrM™ (formerly PONTIS) [AAASHTO, 2013] which assists in managing highway bridges and other structures by developing long-term maintenance plans, due to the uncertainties from various sources, decision makers make adjustment decisions every two years based on the observed states of the bridge elements and try to bring the state of the system back to the target values. Thus, the decision process is more like a closed loop feedback system and we assume that a sequential set of decisions will be made every two years through the life span of the bridge.

Based on the above discussion, a quantitative system model is constructed to model the dynamics of the bridge maintenance subsystem. We define s_s as the logit transformation of the superstructure condition rating, and s_d as the logit transformation of the deck condition rating as:

EQUATION 6-10

$$\mathrm{s} = \mathrm{logit}(condition\ rating) = \log\left(\frac{condition\ rating/9}{1 - condition\ rating/9}\right)$$

The logit transformation transforms a number within an interval of $(0,9)$ to $(-\infty, +\infty)$, which works better with a nonlinear state space model. Let $s_s(k)$ and $s_d(k)$ be two state variables of the subsystem at time k, $k = 0, \dots, T-1$. k has a unit of two years which is consistent with the inspection period. Let $u(k)$ be the maintenance spending decision for both the superstructure and the deck made by the decisionmaker at time k.

In the first step of the analysis, we ignore the interdependency between the maintenance and traffic engineering subsystem and assume that there is no impact from the traffic engineering subsystem on the shared state s_d. In the second step, the impact from the traffic engineering subsystem is modeled as perturbations on the shared state s_d. A simplified nonlinear state space model based on Equation 6-8 is used to describe the element deterioration and maintenance activities of the subsystem:

EQUATION 6-11

$$\begin{bmatrix} s_s(k+1) \\ s_d(k+1) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} s_s^2(k) \\ s_d^2(k) \\ s_s(k) \\ s_d(k) \end{bmatrix} + \begin{bmatrix} a_{15} \\ a_{25} \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(k)$$

$$k = 0, \dots, T-1$$

Equation 6-11 is the same quadratic state-space model as Equation 6-8, and the only difference is that $x_1(k)$ in Equation 6-8 doesn't explicitly appear in Equation 6-11, but as a perturbation. This model describes the system behavior for time period from 0 to $T-1$ and k is the index of time with a two-year interval. Parameters a_{ij} and b_i are system parameters for the state variables and control respectively, which have been discussed in the previous section. Their values can be obtained through parameter estimation (system identification) based on available databases, or through expert elicitations. The following values of matrix a and b are used as a reasonable approximation of a general class of bridges.

EQUATION 6-12

$$a = \begin{bmatrix} -0.06 & 0 & 0.93 & -0.01 & -0.07 \\ 0 & -0.06 & 0.01 & 0.91 & -0.095 \end{bmatrix}; b = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

Among them, the values of $a_{11}, a_{22}, a_{13}, a_{24}, b_1$ and b_2 are estimated from [Andrijcic et al., 2013], assuming that the deck deteriorates much faster than the superstructure but cost much less for maintenance. Although plenty of literature discuss the existence of superstructure-deck interaction, we found no quantitative modeling of that interaction, which determines the value of a_{12}, a_{21}, a_{14} and a_{23}. We start from a very small value to see if a weak interaction introduces systemic risks to the system.

Let f_1 and f_2 be two objective function of the subsystem, which are the deviation of the states from the desired control target. We assume that the decision maker wants to maintain the condition rating of both superstructure and deck at a certain level and minimize the deviation from it, and a shift in the lower direction in the superstructure rating will cause system failure. To simply the discussion, the two objectives take the following form:

EQUATION 6-13

$$\begin{bmatrix} f_1(k) \\ f_2(k) \end{bmatrix} = \begin{bmatrix} (s_s(k) - c_s)^2 \\ (s_d(k) - c_d)^2 \end{bmatrix}$$

$$\text{for } k = 0, \dots, T-1$$

where c_s and c_d are control objectives of the superstructure and deck condition rating respectively. For a target value of "good condition" as shown in Table 6-1, $c_s = c_d = \text{logit}(7) \approx 1.2$.

The decision problem is then formulated as

EQUATION 6-14

$$\min_{u(k)} \{f_1(k+1), f_2(k+1)\}$$

Let $\theta, \theta \in (0,1)$ be the weight between these two objectives, the above problem becomes a single objective decision problem:

EQUATION 6-15

$$\min_{u(k)}\{\theta f_1(k+1) + (1-\theta) f_2(k+1)\}$$

for each $k \in \{0, \dots, T-1\}$.

Following the derivation procedure from Equation 4-5 to Equation 4-15, we can identify the stationary points, stable and unstable regions of the maintenance subsystem state space. The steady states of the superstructure and deck under this decision rule are functions of θ and shown in Figure 6-4. It shows that when the decision maker has equivalent preference on both objectives, the deck condition rating is very close to the target value ("good condition"), while the superstructure condition rating will be lower than the target value. With the decision maker's preference on deck condition increasing, the deck condition rating converges to the target value, while the superstructure condition rating keeps decreasing gradually. However, when the value of θ continues to decrease and pass a crossover point, both states will experience a sudden decrease in the condition rating, which leads to bridge failure. The decision maker's decreasing preference on the deck over superstructure may be the result of growing traffic demand, higher user cost for delays due to poor deck condition, or higher maintenance cost for the superstructure.

A closer examination on the decision variable u – the spending on bridge maintenance that is depicted in Figure 6-5 – shows that just before the abrupt state change, the decision maker does increase the maintenance spending at an exponential rate. However, even with the increase in maintenance spending, the rapid drop of both condition ratings is not prevented.

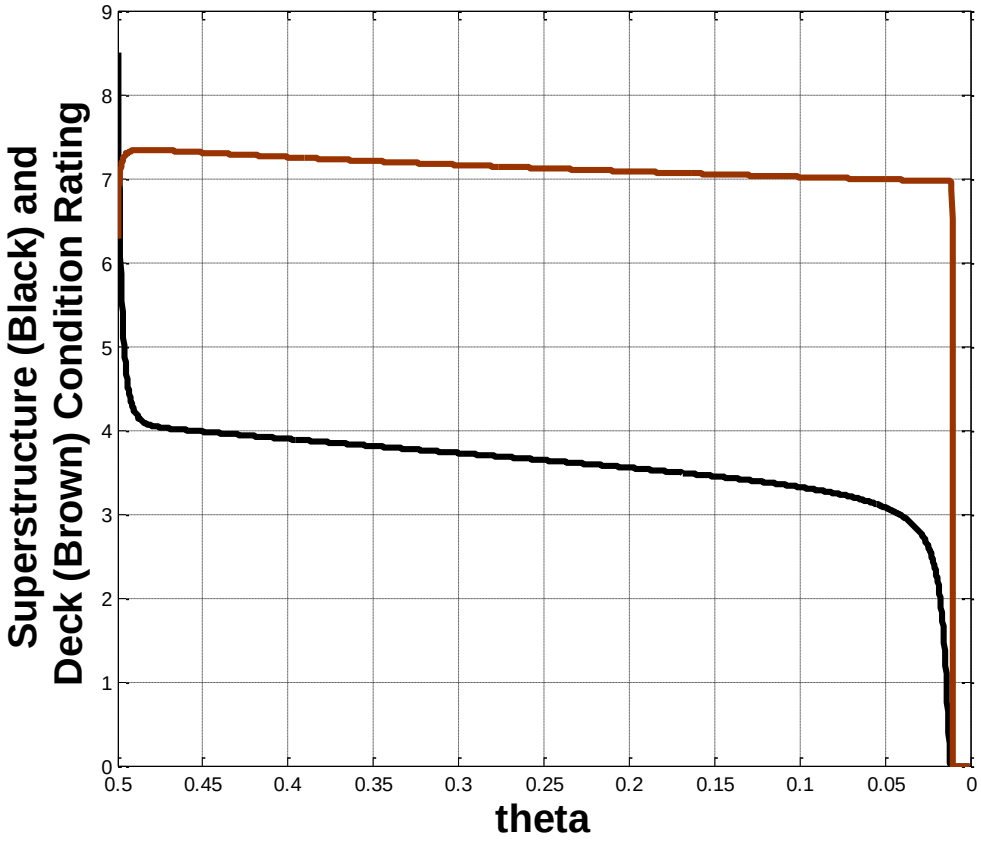

FIGURE 6-4. STEADY STATES OF SUPERSTRUCTURE AND DECK AS A FUNCTION OF θ

Analysis shows that θ must be greater than 0.1 for the maintenance subsystem to have a steady state along the decision process, not considering the perturbations introduced by subsystem interdependencies. This implies that if the decision maker's preference on bridge performance (which is based on deck condition rating) is about 9 times higher than bridge reliability (which is based on superstructure condition rating) and makes maintenance decisions solely based on these two criteria, the bridge may not be maintained in a sustainable way.

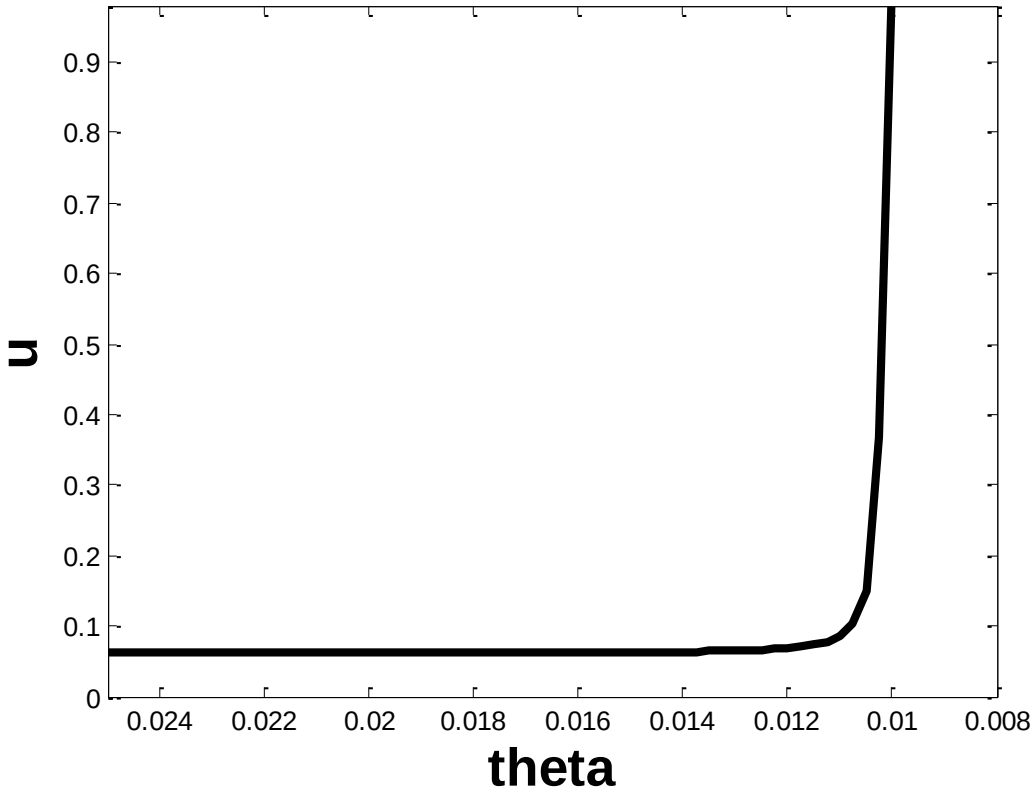

FIGURE 6-5. MAINTENANCE SPENDING AS A FUNCTION OF θ BEFORE ABRUPT STATE CHANGE

This risk modeling and assessment approach models the physical deterioration, inspection, decision, and maintenance loop as a whole system, and demonstrates that systemic risks from a multi-objective decision process do exist in a maintenance subsystem of the bridge infrastructure SoS, such that a precursor analysis process is needed to detect this emergent forced changes to the system.

6.4. Precursor Analysis for Bridge Infrastructure Systems of Systems

This section demonstrates the application of the precursor analysis framework in Chapter 5 on a bridge infrastructure SoS with two subsystems: the maintenance subsystem and traffic engineering subsystem.

6.4.1 Identify Failure and Failure Modes

In a bridge system, physical failure usually means broken critical bridge elements or even the collapse of bridge. One of the major bridge functional failures is bridge deficiency which is the form of failure we are investigating in this case study, which usually doesn't cause bridge collapse directly, but it reduces the safety margin of the bridge and accelerates the bridge deterioration processes.

In a bridge system, three failure modes for bridge deficiency are identified:

- Structurally Deficient (SD) – significant load carrying elements are found to be in poor condition due to deterioration and/or damage.
- Functionally Obsolete (FO) – the deck geometry, load carrying capacity, clearance, or approach roadway alignment no longer meet the usual criteria for the system of which it is an integral part.
- Overload (OL) – live load and dead load exceed the load capacity of bridge.

The FHWA Bridge Preservation Guide [FHWA, 2011] defines the following conditions for the first two failure modes:

- Structurally Deficient (SD)
    - Deck condition rating ≤ 4
    - Superstructure condition rating ≤ 4
    - Structural evaluation ≤ 2
    - and others
- Functionally Obsolete (FO)

- Deck geometry ≤ 3
- Structural evaluation ≤ 3
- and others

The system constraint for overload (OL) is

- Live load + Dead load < Bridge load capacity

The relationship among these constraints is shown in Figure 6-6.

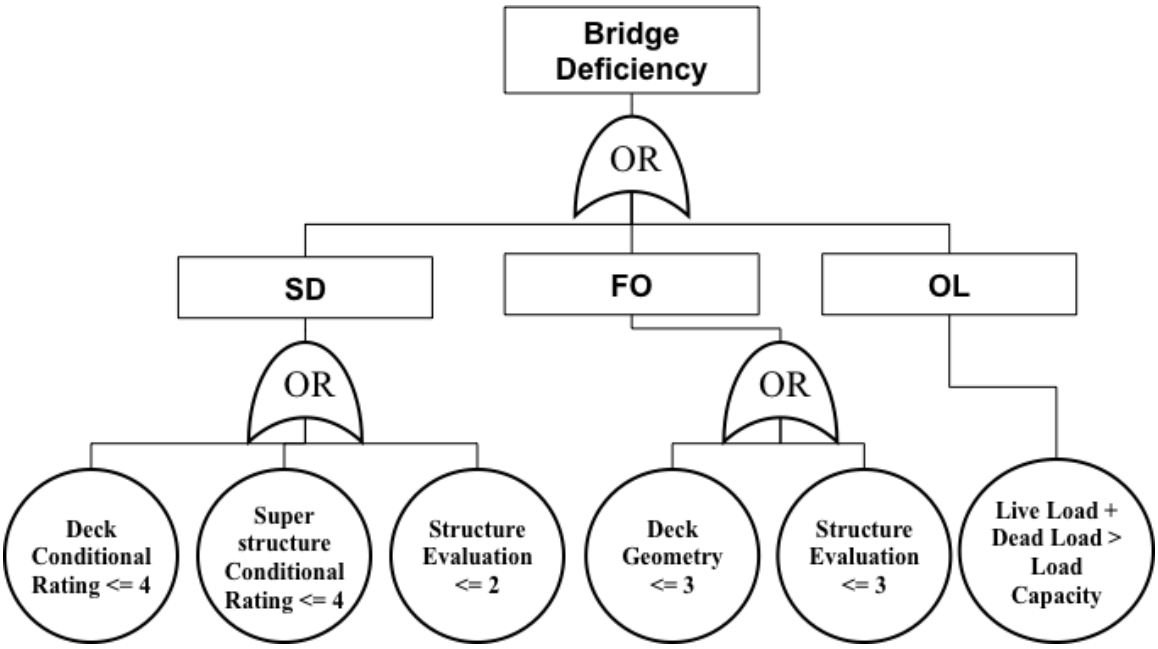

FIGURE 6-6. SYSTEM CONSTRAINTS AND RELATIONSHIPS FOR BRIDGE DEFICIENCY

6.4.2 Precursor Identification, Filtering, and Prioritization

The bridge SoS functional components and control structure identified in Section 6.2 Figure 6-2 are used for precursor scenario identification. Following the methodology developed in Section 5.3.1, potential defects in each functional component can be explored and examined and a set of precursor scenarios can be identified.

In the precursor identification process, we should not only be concerned with the relationship of single precursor scenario to single failure mode, but also the relationship of a single precursor scenario to multiple failure modes and multiple precursor scenarios to single failure mode. In order to facilitate this discovery process, an HHM is constructed and used to organize precursor scenarios under different failure modes and functions, which is shown in Figure 6-7. By exploring the HHM along the horizontal directions, it provides an intuitive way to identify common system functional components thus common precursor scenarios to multiple failure modes of the system. For example, potential defections in all aspects of the inspection function may become precursor scenarios for failure modes of both structurally deficient and functionally obsolete. By exploring the HHM along the vertical directions, it reveals all precursor scenarios and their possible combinations to a single failure mode of the system.

The likelihood criterion of a precursor scenario is a measure of the maximum conditional probability of system failure within a specific time domain given the existence of precursor scenario. The urgency criterion of a precursor scenario is a measure of the time to reach a specific failure probability given the existence of a precursor scenario. Each precursor scenario is evaluated by these two criteria respectively and the precursor scenarios with higher likelihood and/or higher urgency receive higher priority for further monitoring.

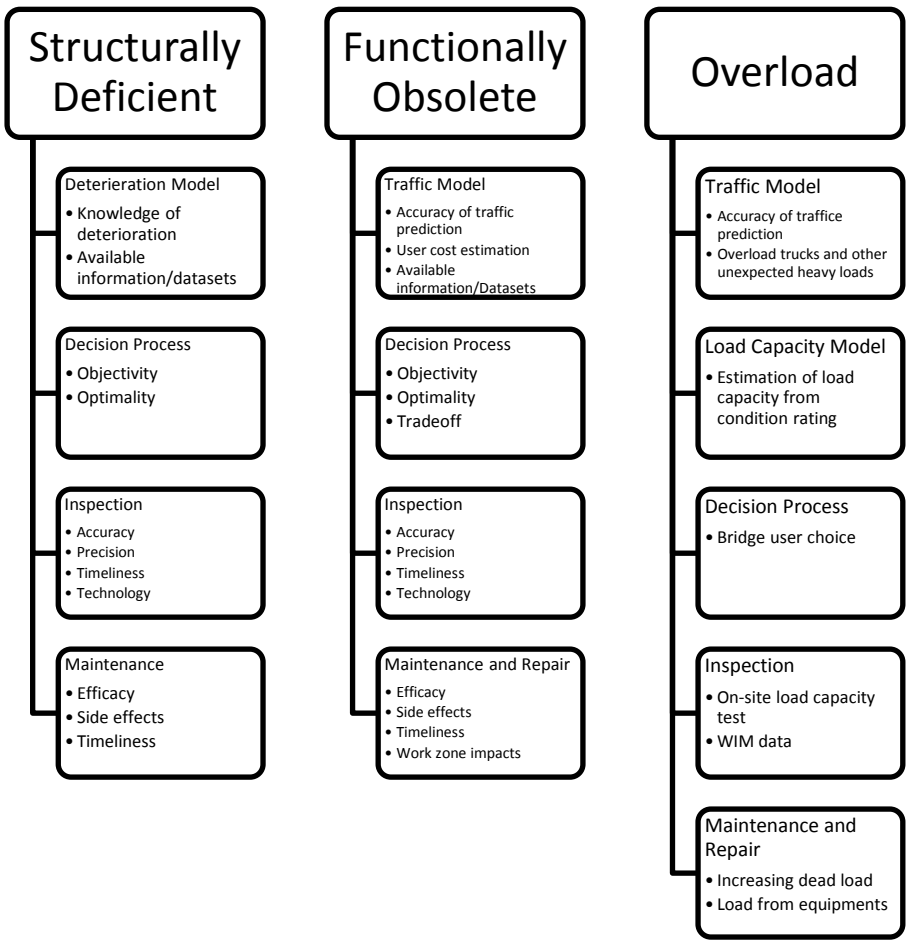

FIGURE 6-7. EXAMPLE HHM FOR THE BRIDGE SYSTEM

A subset of the identified precursors is summarized in Table 6-2.

TABLE 6-2. EXAMPLE PRECURSORS RESULTING FROM HHM FOR BRIDGE SYSTEM

Precursor Code	Failure Mode (Head Topic)	Functional Component (Subtopic)	Description
1.1.1	Structurally Deficient	Deterioration Model	The actual component deterioration process is 2% faster than the national average rate and this is captured by the process model.
1.1.2	Structurally Deficient	Deterioration Model	Maintenance decisions are made based upon a simplified process model which doesn't account for nonlinear behaviors of the system.

1.1.3	Structurally Deficient	Deterioration Model	Maintenance decisions are made based upon a simplified process model which doesn't account for component interactions during deterioration process.
1.2.1	Structurally Deficient	Maintenance Decision	Decisions are not made based on actual condition of the component (e.g., using a predefined static maintenance plan/budget).
1.2.2	Structurally Deficient	Maintenance Decision	Solutions to optimal decision strategy are not practically available. Actual decisions deviate from the optimal ones.
1.3.1	Structurally Deficient	Inspection	Inspection interval is longer than two years such that maintenance decisions are often based on outdated data.
1.3.2	Structurally Deficient	Inspection	Precision and accuracy in visual inspection results is low. Inspection error $\sim N(0.5,0.8^2)$
1.4.1	Structurally Deficient	Maintenance	Maintenance is performed too late (lagging two years).
1.4.2	Structurally Deficient	Maintenance	Maintenance doesn't restore the component condition as expected (assume 75% restoration).
1.4.3	Structurally Deficient	Maintenance	Maintenance cause unexpected consequences on other bridge components or subsystems (deicing causes increased interaction).
1.4.4	Structurally Deficient	Maintenance	Limited funding for maintenance projects (assume $10M every two years).
1.5.1	Structurally Deficient	Deterioration Process	Unknown couplings, interdependencies, and unexpected perturbations exist.
1.5.2	Structurally Deficient	Deterioration Process	Process change under different conditions and not captured by the process model (Chloride concentration level at the interface of rebar exceeds the corrosion reaction threshold).

	Functionally Obsolete		All precursor scenarios for Structurally Deficient are applicable to Functionally Obsolete. Only unique precursors to Functionally Obsolete are listed below.
2.1.1	Functionally Obsolete	Traffic Model	The model doesn't capture the increasing trend in traffic volume over the bridge (assume an additional 50% increase)
2.1.2	Functionally Obsolete	Traffic Model	Maintenance (work zones) on deck has a higher than expected impact on traffic and user cost.
2.2.1	Functionally Obsolete	Maintenance Decision	In a multi-objective decision process, the priority (preference) on some objectives is too low.

The system control model in Equation 6-1 through Equation 6-4 describes the system behaviors under the business-as-usual scenario based on its assumptions. The key step in evaluating each precursor scenario is to translate and quantify the difference in that precursor scenario (compared to the baseline scenario) into the changes in system structure, parameters, or states, such that the system behaviors under that precursor scenario can be simulated and compared with the baseline scenario. We use an example to demonstrate this process.

The system control model in Equation 6-1 to Equation 6-4 is an ideal model without considering the practical issues of each functional component. For example, it is assumed that the bridge inspection team that functions as a sensor within the control loop is able to measure the true deterioration state of the bridge and provide accurate and precise measurements of condition rating to the decision maker in state DOT. In this ideal

scenario, any rational decision maker will choose an optimal maintenance spending based on Equation 4-7 and Equation 6-1

EQUATION 6-16

$$u_m^*(k) = \rho\theta b_{11}[c_s - a_{11}s_s^2(k) - a_{12}s_d^2(k) - a_{14}s_s(k) - a_{15}s_d(k)] + \rho(1-\theta)b_{21}[c_d - a_{21}s_s^2(k) - a_{22}s_d^2(k) - a_{24}s_s(k) - a_{25}s_d(k)]$$

for $k = 1, \dots, T-1$.

However, the true state of the superstructure is usually not directly observable and current practice relies heavily on visual inspections due to limited inspection resources. This practice increases the bias and uncertainty in the quality of observed bridge states and may affect the failure probability of the bridge in the long run. To evaluate the impacts of this precursor scenario on the likelihood and urgency of bridge failure, we modified Equation 6-16 to incorporate the errors in bridge inspection by adding another random variable $\omega(k), k = 1, \dots, T$. The mean of $\omega(k)$ represents the accuracy of the inspection results and the variance of $\omega(k)$ represents the precision of the inspection results at time stage k. Based on a study by Phares et al. [2004], the mean and standard deviation of the errors in superstructure inspection results follows approximately a normal distribution of $N(0.5, 0.8^2)$. The new solution incorporating inspection uncertainty becomes

EQUATION 6-17

$$u'_m(k) = \rho\theta b_{11}\Big[c_s - a_{11}\big(s_s(k)+\omega(k)\big)^2 - a_{12}s_d^2(k) - a_{14}\big(s_s(k)+\omega(k)\big) - a_{15}s_d(k)\Big] + \rho(1-\theta)b_{21}\Big[c_d - a_{21}\big(s_s(k)+\omega(k)\big)^2 - a_{22}s_d^2(k) - a_{24}\big(s_s(k)+\omega(k)\big) - a_{25}s_d(k)\Big]$$

Equation 6-16 and Equation 6-17 enable us to compare the inspection error scenario (code 1.3.2 in Table 6-2) with the baseline scenario through the evaluation of the likelihood and urgency of each precursor scenario by either analytical method or numerical simulation. For some type of precursor scenarios and corresponding system models, there is no analytical solution available. To standardize the process, all precursor scenarios in this paper are evaluated through numerical simulation.

Figure 6-8 shows a typical simulation result of the baseline (no precursor) scenario and inspection error precursor scenarios. The maximum probability of system failure within 50 years and the time to reach a failure probability of 0.01 is used to compare these two scenarios. As the figure shows, the errors in bridge inspection have almost no impacts on the maximum failure probability over the 50 years time domain, as well as the time to reach a failure probability of 0.01. This result shows that the errors in superstructure inspection alone don't have significant impacts on system failure and this precursor scenario doesn't justify itself as a precursor thus may not deserve further monitoring, subject to other assumptions in the model formulation.

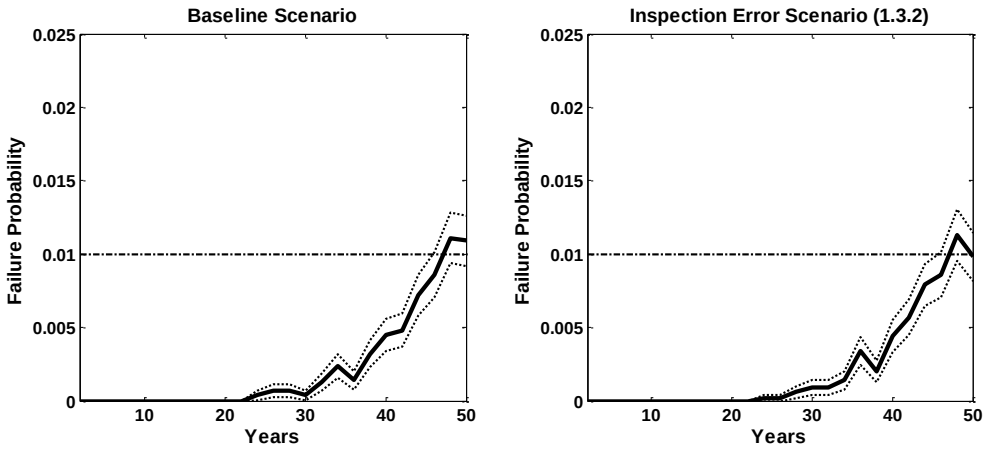

FIGURE 6-8. COMPARISON OF SIMULATION RESULTS OF THE ESTIMATED SYSTEM FAILURE PROBABILITY WITH 90% CONFIDENCE INTERVAL, BETWEEN THE BASELINE AND INSPECTION ERROR SCENARIOS

Figure 6-9 shows another simulation result where the precursor scenario of faster deterioration (code 1.1.1 in Table 6-2) is compared to the baseline scenario. This scenario simulates the impacts from a 2% increase in the parameters that determine the deterioration rate of superstructure and deck compared to the national average. This figure shows that even though the process model captures this change and adjusts the decision accordingly, a faster deterioration rate does have significant impacts on both criteria. The maximum failure probability within the 50-year period is increased by 2.5 times, and the time to reach a failure probability of 0.01 is reduced by approximate 15 years.

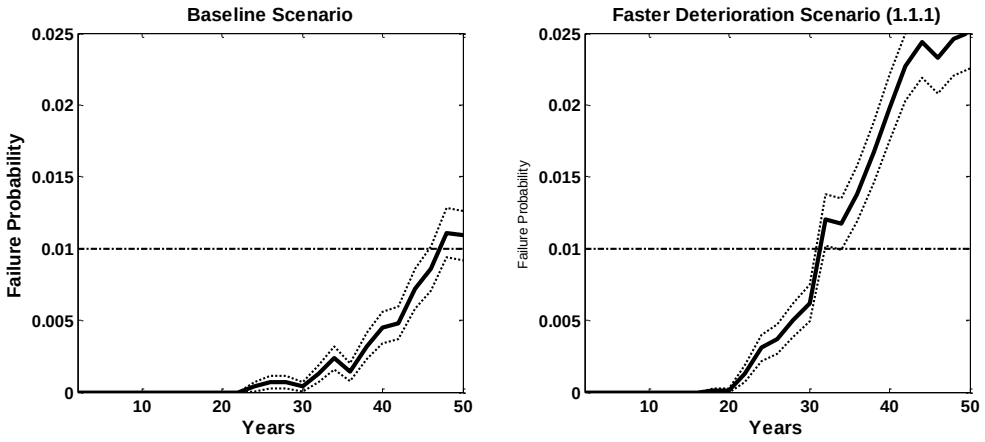

FIGURE 6-9. COMPARISON OF SIMULATION RESULTS OF THE ESTIMATED SYSTEM FAILURE PROBABILITY WITH 90% CONFIDENCE INTERVAL, BETWEEN THE BASELINE AND FASTER DETERIORATION SCENARIOS

Figure 6-10 shows the simulation result where the precursor scenario of limited funding (code 1.4.4 in Table 6-2) is compared to the baseline scenario. Assuming the maximum funding for superstructure maintenance is 10 million US Dollars in every two years, the simulation of limited funding is done by modifying Equation 6-16 as

EQUATION 6-18

$$u'_m(k) = \begin{cases} 0 & if\ u^*_m \leq 0 \\ u^*_m & if\ 0 < u^*_m < 10 \\ 10 & if\ u^*_m \geq 10 \end{cases}$$

Results show that the maximum failure probability of baseline scenario is around 0.012, and the maximum failure probability of limited funding scenario is around 0.016. The time for the bridge to reach a state with a failure probability of 0.01 in the baseline scenario takes around 47 years, and almost the same in the limited funding scenario. Thus, limited maintenance funding increases slightly the maximum failure probability but has negligible impacts on the urgency criterion.

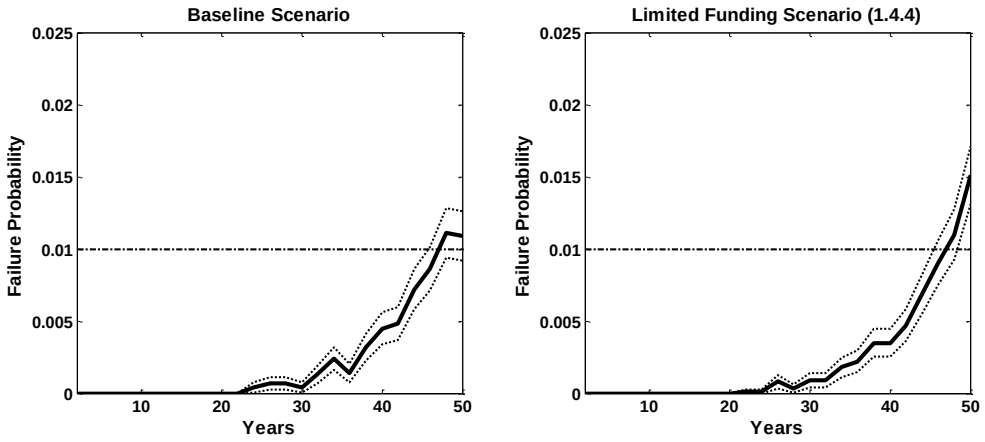

FIGURE 6-10. COMPARISON OF SIMULATION RESULTS OF THE ESTIMATED SYSTEM FAILURE PROBABILITY WITH 90% CONFIDENCE INTERVAL, BETWEEN THE BASELINE AND LIMITED FUNDING SCENARIOS

After evaluating all precursor scenarios in Table 6-2 and both criteria are recorded, we can plot all precursor scenarios in a two dimensional figure as shown in Figure 6-11. The precursor scenarios at the bottom-left corner (precursor scenario 1.2.1, 1.1.1, 1.5.1, 1.4.2, 1.5.2) have higher likelihood and higher urgency to cause system failure. They can be considered as precursors and must be monitored continuously. The precursor scenarios at the top-left corner (precursor scenario 1.4.3, 2.1.2) have lower likelihood but higher urgency. These are more selective precursors to be monitored if resources are available. The precursor scenarios at the top-right corner (precursor scenario 1.1.2, 2.2.1, 1.2.2, 1.3.1, 1.4.4, 1.4.1, 2.1.1, 1.1.3, 1.3.2) have lower likelihood and lower urgency so they can be dismissed if there is limited resource for a monitoring system. However, this filtering process needs to be revisited every time new information about the system becomes available, such that it does not miss any critical precursors to system failure.

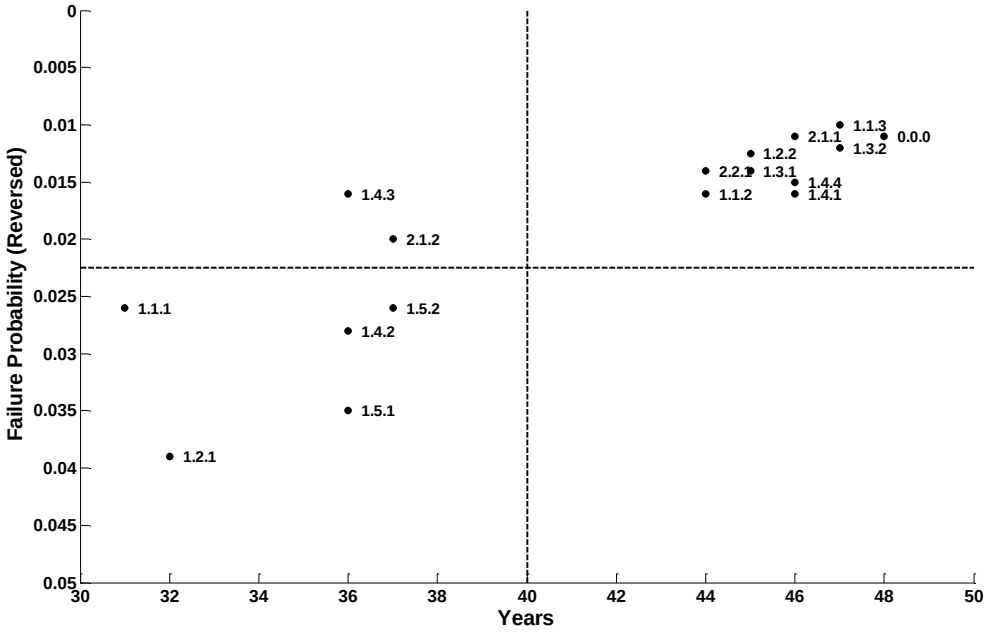

FIGURE 6-11. TWO-DIMENSIONAL FILTERING OF PRECURSORS BASED ON FAILURE PROBABILITY AND URGENCY

6.4.3 Precursor Detection

In the precursor detection phase, we are interested in detecting whether certain conditions or relationships exist or not among state variables and other building blocks of the system such as inputs, outputs, and decisions. Thus, the task of the detection phase is to quantify the detection probability Pr(Precursor | Evidence) of the precursors resulting from the filtering and prioritization process.

For example, precursor 1.5.2 in Table 6-2 indicates a potential system failure caused by an unmatched deterioration model used for the actual deterioration process of bridge components. If the chloride concentration level is lower than the threshold, the corrosion process of rebar is not initiated and the deterioration rate is relatively slow. Once the chloride concentration level is higher than the threshold, a different deterioration model incorporating the effects of reduction in rebar cross section area due to corrosion

reactions must be used. To detect this precursor, the chloride concentration level, which is also a state of the system, needs to be estimated by a chloride diffusion model or measured directly from a sample specimen. Then the chloride concentration level is compared with the pre-determined threshold to determine whether the precursor is valid. In general, the detection process aims to estimate the values of some specific system parameters or states, in this case, the chloride concentration level on rebar surface. To quantify the uncertainties in the detection results, the distribution parameters of the results such as mean and variance also need to be estimated. A well-designed experiment and data analysis are often employed for this task.

Research [Thoft-Christensen, 2000] shows that corrosion of steel rebar in reinforced concrete beams initiates when the chloride concentration level at the interface of rebar reaches 0.3% by weight of cement. After this corrosion process starts, the cross section area of the rebar decreases which causes the loss of pre-stress of the concrete beam. Different deterioration models should be used before and after the corrosion process initiates. If there is a mismatch between the model and the actual process, for example, the model which doesn't account for the corrosion process continues to be used after the corrosion process actually starts, the deterioration states of the bridge beam may not be correctly identified and this situation is considered as a precursor to bridge failure.

To detect this precursor, the actual chloride concentration level around the rebar needs to be measured or estimated. A measurement method using actual samples from the structure is a destructive testing approach which may reduce the integrity of the structure in the long run. A theoretical diffusion model such as shown in Equation 6-19 based on Fick's law of diffusion is commonly used to estimate this variable.

EQUATION 6-19

$$C(x,t) = C_0\left\{1 - \text{erf}\left(\frac{x}{2\sqrt{D_c \cdot t}}\right)\right\}$$

where $C(x,t)$ is the chloride ion concentration at a distance of x cm from the concrete surface after t seconds of exposure of the chloride source. D_c is the chloride diffusion coefficient expressed in cm^2/sec. C_0 is the equilibrium chloride concentration on the concrete surface, and erf is the error function. Although Equation 6-19 is a deterministic function, uncertainties exist in some parameters such as C_0 and D_c which depend on material, construction approach, and environment. In practice, the estimated chloride concentration is a random variable. Suppose for time t, the model estimates that the chloride concentration at distance x follows a normal distribution $C \sim N(0.25, 0.05)$, we have:

EQUATION 6-20

$$\Pr(Precursor \mid Evidence) = \Pr(C > 0.3 \mid C \sim N(0.25, 0.05)) = 0.159$$

and

EQUATION 6-21

$$\Pr(no\ Precursor \mid Evidence) = \Pr(C \leq 0.3 \mid C \sim N(0.25, 0.05)) = 0.841$$

where the precursor is expressed as a relationship between a system state variable C and a constant (0.3), and the evidence is the results in detection ($C \sim N(0.25, 0.05)$). This result shows that given the uncertainties in the detection, the probability of the existence of this precursor is relatively low.

Given Equation 6-20 and Equation 6-21, we are now able to calculate the system failure probability given the detection results (evidence).

Pr(system failure within a specific time domain | Evidence) =

*Pr(system failure within a specific time domain | Precursor)*0.159 +*

Pr(system failure within a specific time domain | no Precursor) 0.841=*

*1.4E-3*0.159 + 0.6E-3*0.841 = 0.73E-3*

where *Pr(system failure within a specific time domain | Precursor)* can be estimated through the simulation approach used in section 4.2.2 and *Pr(system failure within a specific time domain | Precursor)* is just the failure probability of the business-as-usual scenario. This result literally shows that based on the precursor detection result, the bridge may fail with a probability of 0.73E-3 in the next 50 years. However, as we posited, deriving the absolute value of this probability is not the purpose of this process, because no model is able to predict the probability of a complex system in 50 years due to so many unforeseeable factors not accounted in the model. Rather, the purpose of this process is to compare and prioritize different precursors so that risk management resources can be allocated to the most important ones thus actively reducing the failure probability of the system.

6.4.4 Precursor Evaluation

The system control meta-model developed in Equation 6-1 through Equation 6-4 can be readily used to incorporate new information whenever a precursor is detected along the timeline to fuse the information from multiple precursors and evaluate the likelihood of multiple failure modes. Instead of calculating the failure probability of the whole system which is the top event in Figure 6-6, the model can calculate the failure

probability of each failure mode, which is the intermediate event under the top event (SD, FO, and OL in this example). When the trend of likelihood of each failure mode and the likelihood of no failure is displayed along the timeline when each precursor is detected, the decision maker may have a better understanding of the current situation and possibly the root cause of the situation. Risk management actions targeted on a specific failure mode can be developed and evaluated. An example based on the accident investigation of Minnesota I-35W Mississippi river bridge collapse [NTSB 2007, 2008] is used to demonstrate the precursor evaluation process.

In this example, the probability of three failure modes – SD, FO, and OL – as well as the probability of no failure is calculated and monitored along a twenty-year time span. The initial superstructure and deck condition rating are assumed to be 6. Due to the fact that failure modes are not necessarily mutually exclusive, the probability of no failure is calculated as:

EQUATION 6-22

$$\Pr(no\ failure) = 1 - \Pr(SD) - \Pr(FO) - \Pr(OL) + \Pr(SD \cap FO) + \Pr(FO \cap OL) + \Pr(OL \cap SD) - \Pr(SD \cap FO \cap OL)$$

Four precursors are detected within above time span:

(i) Faster deterioration at year 4, , with P(existence of precursor = 0.5)

(ii) Rehabilitation project is not performed at year 16. (This is a known fact with no detection uncertainty.)

(iii) No inspection at year 18. (This is a known fact with no detection uncertainty.)

(iv) Increasing dead load due to maintenance throughout the 20-year period, with P(existence of precursor = 0.8).

The system model is reconfigured to incorporate information from these precursors and failure probabilities of each failure mode are plotted in Figure 6-12. "Reconfiguration" means changing system model parameters or structures to simulate the precursor scenario. In this case, at year 4, 16, and 18, the model is changed to accommodate the precursor. For example, from year 4, the parameters determining the deterioration rate are increased; at year 16, u(16) = 0; and at year 18, maintenance decision is based on s(16) instead of s(18); and the dead load has a linear increase throughout the 20-year period. In this example, the model is not used to predict the future probability of bridge collapse, as we did in the filtering of precursors. Instead, this model quantifies the failure probability at the current time stage as a monitoring and tracking tool.

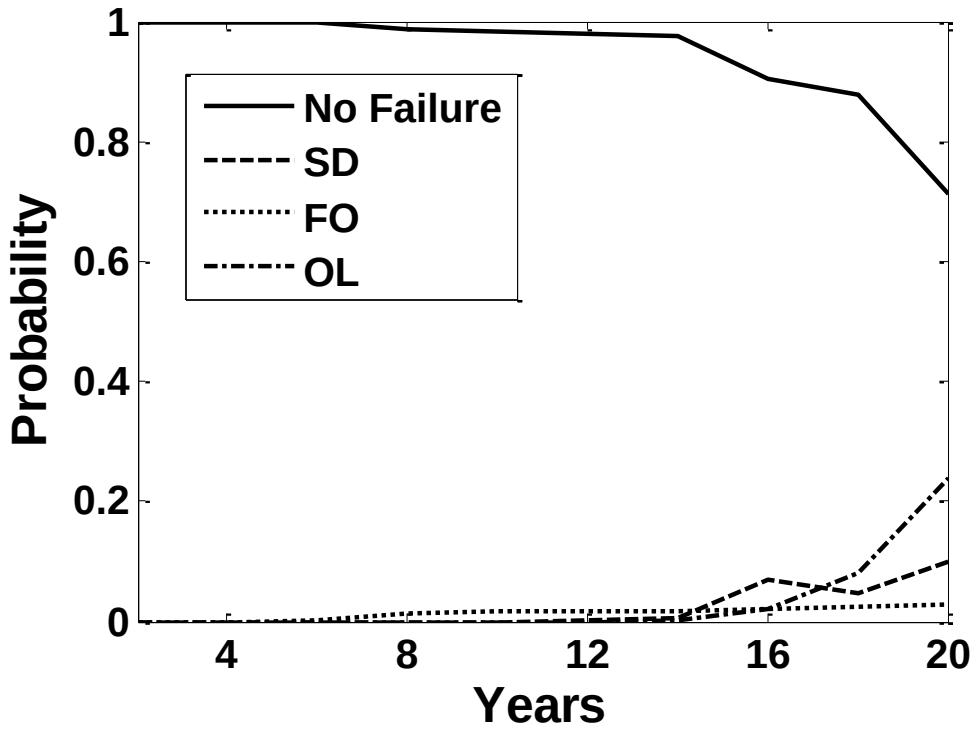

FIGURE 6-12. MULTI-PRECURSOR EVALUATION

Figure 6-12 shows that the failure probability of all failure modes is very low at the beginning. The first precursor at year 4 doesn't increase the failure probability significantly. Failure probabilities of SD and OL gain a sharp increase and surpass FO after the detection of the second precursor at year 16. The detection of the third precursor boosts the failure probability of OL well above that of the SD. However, it seems that none of the precursors has significant impacts on the failure probability of FO. At the end of the twenty-year time period, the failure probability of OL is highest among all three failure modes. In this case, the failure due to overload should be considered highly likely and risk management actions should be planned and executed accordingly. This graph delineates and directly compares the probability of all failure modes before and after the detection of each precursor, thus provides an intuitive and objective way to communicate

the current situation of the system to the decision maker so that they can choose appropriate risk management actions accordingly.

6.5. Conclusions

This chapter demonstrates the theories and methodologies developed in Chapter 3, 4, and 5 with a case study of bridge infrastructure SoS. Through traditionally being considered as engineering systems, the bridge infrastructure SoS encompasses different stakeholders, decision makers, functional organizations, and processes. A meta-modeling approach capturing the performance and safety objectives of the bridge are achieved above its natural deterioration process is developed. It demonstrates the existence of interdependencies and other conditions for certain systemic risks of bridge SoS. The precursor analysis framework is applied to identify and prioritize precursors to future bridge failure. Predicting the behaviors of a complex system is a challenging task. Although the resulting numbers in this example may not be practically useful by themselves, this precursor-based risk analysis approach provides a systemic and constructive way with which to explore and understand the sources of failure of complex systems and to develop proactive risk management.

7. Conclusions and Future Directions

Developing risk analysis theories and methodologies for complex infrastructure SoS, including various large-scale infrastructure systems, is a necessary and urgent requirement for system owners, decision makers, and users. It requires a systemic and holistic approach that integrates multiple perspectives, models and tools. The goal of this dissertation is not to compare or replace the existing tools with the tools developed here since each approach is based on different aspects and assumptions of the problem. Instead, it aims to add another tool in the toolbox for analyzing risks to complex infrastructure SoS, because it is believed that in the face of complex systems, each new approach provides a new perspective to the problem, and a combination of tools is needed to address the challenges of modeling and analyzing complex systems. This integration process is a continuous learning process which will enable us to better understand complex systems.

Complex infrastructure SoS possess unique characteristics, which distinguish them from traditional engineering systems. In this dissertation we posit that these

characteristics constitute major source of systemic risks that are inherent in complex SoS. We explore one such specific systemic risk in a nonlinear dynamic multi-objective decision process and demonstrate that: 1) subsystems with shared states can be decomposed and coordinated; 2) decision maker's preference on multiple objectives may cause system instability; and 3) perturbations can be introduced to the subsystems through shared states that further reduce the safety margin of the subsystem, and the system becomes susceptible to small perturbations when decision makers have a high preference for one objective.

This dissertation also provides a systemic framework for precursor analysis in complex infrastructure systems of systems. Three major phases in precursor analysis – (i) system modeling, (ii) precursor filtering and prioritization, and (iii) precursor detection and evaluation – are discussed respectively with a bridge infrastructure example. We demonstrate that all the characteristics of complex infrastructure SoS exist in real-world bridge systems. Based on the analysis results, we posit that the systemic risk discussed in this dissertation is most relevant to bridge systems and might be a primary cause for unforeseen and sudden bridge failures. With the advances in sensing and automatic monitoring technologies, new warning systems using multiple signals could be designed based on this improved understanding of systemic risks in the system. We demonstrate that through a systems engineering approach aimed at improving our understanding of the complex failure mechanism of systems, by designing efficient monitoring systems, and evaluating precursors, our quantitative precursor analysis is capable of objectively reducing the hindsight bias in precursor analysis and providing a sound theoretical basis for risk management and decision making.

This chapter summarizes our preliminary work in precursor analysis, where future work is expected in the following areas:

First, the precursor analysis framework needs to be validated based on data from a specific bridge system and demonstrate that the identified precursors do predict the failure of bridge if no risk management actions are taken. However, this task is difficult to achieve due to a couple of reasons. First, the actual deterioration of a bridge is a very complex process with many unknown factors and their interactions, and a nonlinear, time-variant model is required to capture this behavior. This means that a model that is built based on the data from earlier deterioration stage of the bridge may not be able to predict the later stage of the deterioration process due to system nonlinearity and time-variant parameters. If consider further on the very limited data points that can be collected from a bridge (generally new information is acquired every two years with bridge inspection, and a total of 50 data points are available in a one-hundred-year time span) and the quality of the data (mainly based on subjective visual inspection), the estimation of time-dependent system parameters is a big challenge for the modelers. The problem here is that there is not sufficient data to even train the model, not to mention validate the model. Secondly, each bridge is different in terms of the way it is designed, constructed, maintained and managed, the materials it uses, and the external environment under which the bridge is operated, such as the precipitation, humidity, temperature, traffic load and pattern. The applicability of directly applying a model built for one bridge to another bridge is very limited even if they have similar structure. It means that precursors for one bridge don’t necessarily have the same level of impacts on other bridges. And finally, as the dissertation suggests, the purpose of precursor analysis is to

support pro-active risk management. In reality, large repair and rehabilitation projects are often being implemented at certain time to prevent bridge from failure thus inevitably change the probability of bridge failure and possibly invalidate the model. To address this issue, we are proposing a research project to the National Science Foundation on a paradigm shift in modeling, understanding, and managing the risk and the lifecycle of bridge infrastructure as a complex system of systems [Chase, Haimes, Andrijcic, and Guo, working paper]. In this proposal, we posit that the current approach to bridge management is based upon condition and performance measures that are very subjective and heavily weighted toward ordinal assessments based upon non-quantitative criteria and methods. A combination of factors—technical, professional, managerial, economic, political, and institutional—has brought about a situation in which highway bridges do not provide adequate service life. Our paradigm shift calls for collecting and harvesting quantitative data on the actual factors that have limited the service lives of specific decommissioned highway bridges. We posit that a detailed examination of bridges that are being taken out of service will provide the ability to document deterioration mechanisms much earlier and provide early detection of those mechanisms responsible for the failure of the bridge. With access to decommissioned bridges provided by the Virginia Department of Transportation, the collected quantitative and objective data including: (i) the actual concrete cover on in-situ bridges; (ii) in-situ condition and performance of bridge bearings; (iii) in-situ condition and performance of bridge joints; (iv) quantitative measurements of diffusion and corrosion rates and metal loss; and (v) quantitative measurements of residual stress, will support objective identification, characterization, and definition of these factors and will allow the identification of the

systemic failures at the nexus of a complex system of systems of technical, institutional, organizational, economic, environmental, and political domains.

Secondly, the theories and methodologies developed in this dissertation need to be applied, tested, and validated for other complex SoS. Although the bridge infrastructure SoS is complex, it is not at the same level of complexity as compared with other complex systems, such as financial systems, in terms of the number of stakeholders and decision makers, the frequency of decision making, the complexity in system structure, the way information is exchanged, the level of nonlinearity, and the speed of system response. For example, our research demonstrates that (not included in this dissertation) for some highly nonlinear systems, some statistical measures can be used as precursors to system failure. However, as the bridge system is not such a highly nonlinear system, this finding cannot be verified by the bridge system.

Some statistical and data-driven approaches for detecting signals prior to system failure or abrupt state transition [Wolff, 1990] [Scheffer et al. 2009] [Drake & Griffen 2010] may be supplemented with the precursor analysis framework. Positive feedback is another factor causing system unstable. Approaches to detect the condition for the formation of positive feedback loops in the system include [Kyrtsou & Labys 2007].

In the modeling of SoS, this dissertation posits that shared states, which are the mechanism causing subsystem interdependency, are known. However, in many complex systems identifying these interdependencies may not be a trivial task. Both qualitative and quantitative approaches are needed to discover the interdependencies, which have the potential to cause system failure. Some data-driven approaches to detect couplings

[Romano et al. 2007] [Chicharro & Andrzejak 2009] [Smirnov & Andrzejak 2005] [Andrzejak et al. 2006] can be considered in the future.

Other improvements in the techniques related to this dissertation may include:

- Develop efficient system identification methods for interdependent subsystems through shared state variables.
- Select appropriate detection and estimation methods for hidden state variables.
- Improve system identification and state estimation methods when data are scarce and noisy.
- Incorporate nonlinear dynamic system stability analysis as another perspective to understand system failure.
- Identify related databases for bridge infrastructure, and propose potential data need to be collected in the future.

Other takeaways from this dissertation include

- Risk analysis of complex infrastructure SoS is an iterative, learn-as-you-go process.
- No single tool or method is sufficient to perform risk analysis of complex systems.
- State variables are the essential building blocks of system models.
- Meta-modeling is essential to capture the big picture of a large system without being overwhelmed by details of each component/subsystem; however, more

research in needed in this area to capture all important features or behaviors of all important subsystems of SoS.

The unique value of this approach compared to others is that it recognizes the bridge infrastructure as a complex SoS where physical bridge, human decision process, control and detection actions, and the interactions among these components, all contribute to the dynamics of the overall SoS. Risk management methods focusing only on part of the encompassing system will eventually miss critical factors leading to emergent forced changes to the system. With an emphasis on holism, this approach is expected to bring immense value to the risk assessment and management of complex infrastructure SoS.

8. References

AASHTO (2013). AASHTOWare Bridge Management. Available online at: http://aashtowarebridge.com/

Adeli, H., and Jiang, X. (2003). Neuro-fuzzy logic model for freeway work zone capacity estimation. *Journal of Transportation Engineering*, *129*(5), 484-493.

Aktan, A.E., Faust D. (2003). *A holistic integrated systems approach to assure the mobility, efficiency, safety and integrity of highway transportation*. Published in *Structural Health Monitoring* and Intelligent Infrastructure – Volume 1. Eds. Wu, Z. and Abe M. Sweets and Zeitlinger.

Al-Subhi, K., Johnston, D. W., and Farid, F. O. A. D. (1990). Resource-constrained capital budgeting model for bridge maintenance, rehabilitation, and replacement. *Transportation Research Record*, (1268).

Amaral and J. Ottino. (2004). Complex networks: Augmenting the framework for the study of complex systems. *The European Physical Journal B*, Vol. 38, pp. 147-162.

American Society of Civil Engineers (1998). Report Card for America's Infrastructure (1998 Report Card). *Available at: http://www.asce.org/reportcard/index.cfm?reaction=full&page=1*.

American Society of Civil Engineers (2003). Report Card for America's Infrastructure (2003 Report Card). *Available at: http://www.asce.org/reportcard/index.cfm?reaction=full&page=6.*

American Society of Civil Engineers (2005). Report Card for America's Infrastructure (2005 Report Card). *Available at: http://www.asce.org/reportcard/2005/index2005.cfm.*

American Society of Civil Engineers (2009a). Report Card for America's Infrastructure (2009 Report Card). *Available at: http://www.asce.org/reportcard/*.

American Society of Civil Engineers, (2009b). America's Infrastructure Crisis: Can we Come Back from the Brink? *Available at: http://www.asce.org.*

American Society of Civil Engineers (2013). Report Card for America's Infrastructure (2013 Report Card). *Available at: http://www.infrastructurereportcard.org*.

Andrijcic, E., Chase, S., Guo, Z., and Hwang, S. N. (2012). Exploring system interdependencies via a multi-disciplinary modeling approach: application to bridge management. *The 6th International Conference on Bridge Maintenance, Safety and Management (IABMAS).* Lake Maggiore, Italy.

Andrijcic, E., Haimes, Y. Y., and Beatley, T. (2013). Public policy implications of harmonizing engineering technology with socio-economic modeling: Application to transportation infrastructure management. *Transportation Research Part A: Policy and Practice*, 50, 62-73.

Andrzejak, R. G., Ledberg, A., and Deco, G. (2006). Detecting event-related time-dependent directional couplings. *New Journal of Physics*, 8(1), 6.

Balmer, M., Meister, K., and Nagel, K. (2008). *Agent-based simulation of travel demand: Structure and computational performance of MATSim-T*. ETH, Eidgenössische Technische Hochschule Zürich, IVT Institut für Verkehrsplanung und Transportsysteme.

Barabási, A., and Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*(5439), 509-512.

Bar-Yam, Y. (2003). When Systems Engineering Fails --- Toward Complex Systems Engineering. *International Conference on Systems, Man & Cybernetics*, Vol. 2, pp. 2021-2028.

Bertalanffy, L. von. (1968). *General System Theory: Foundations, Development, Applications, Revised Edition*. New York: George Braziller.

Blauberg I.V., Sadovsky V.N., Yudin, E.G. (1977). *Systems Theory: Philosophical and Methodological Problems*. New York: Progress.

Building America's Future Educational Fund, (2011). Building America's Future: Falling Apart and Falling Behind. *Available at: http://www.bafuture.org*.

Bouti, A., and Kadi, D. A. (1994). A state-of-the-art review of FMEA/FMECA. *International Journal of reliability, quality and safety engineering*. 1(04), 515-543.

Burmeister, B., Haddadi, A., and Matylis, G. (1997, February). Application of multi-agent systems in traffic and transportation. In *Software Engineering. IEE Proceedings-[see also Software, IEE Proceedings]* (Vol. 144, No. 1, pp. 51-60). IET.

Carlson, J. M., and Doyle, J. (1999). Highly optimized tolerance: A mechanism for power laws in designed systems. *Physical Review* E, 60(2), 1412.

Carlson, J. M., and Doyle, J. (2000). Highly optimized tolerance: Robustness and design in complex systems. *Physical Review Letters*, 84(11), 2529.

Chang, M. and Harrington, J. (2005). Agent-based models of organizations. In: *Handbook of Computational Economics II: Agent-Based Computational Economics*, K. Judd and L. Tesfatsion, Eds., pp. 1-66.

Chankong, V, and Y.Y. Haimes (1983). Multiobjective Decision Making: Theory and Methodlolgy. North Holland, New York

Chankong, V, and Y.Y. Haimes (2008). Multiobjective Decision Making: Theory and Methodlolgy. Dover Publications

Chase, S. B., and Gáspár, L. (2000). Modeling the reduction in load capacity of highway bridges with age. *Journal of Bridge Engineering*, 5(4), 331-336.

Chase, S. B., Small, E. P., and Nutakor, C. H. R. I. S. (1999). An in-depth analysis of the national bridge inventory database utilizing data mining, GIS and advanced statistical methods. *Transportation Research Circular*, 498, 1-17.

Chase, S. B., Haimes, Y. Y., Andrijcic, E., and Guo, Z. (Working paper) Modeling, Understanding, and Managing the Risk and the Lifecycle of Bridge Infrastructure as a Complex System of Systems: A Paradigm Shift Manifesto. Submitted to *ASCE Journal of Infrastructure*

Chicharro, D., and Andrzejak, R. G. (2009). Reliable detection of directional couplings using rank statistics. *Physical Review E*, 80(2), 026217.

Corcoran, W.R. (2003a). The Phoenix Handbook. Windsor, Conn.: Nuclear Safety Review Concepts Corporation .

Corcoran, W.R. (2003b). Firebird Forum 6(1). Available online at *http://groups.yahoo.com/group/Root_Cause_State_of_the_Practice/*.

Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, *28*(4), 269-287.

DeLaurentis, D., Callaway, R.K. (2004). A System-of-Systems Perspective for Public Policy Decisions. *Review of Policy Research* Vol. 21, No. 6: 829-837.

DeLaurentis, D. (2008). Appropriate modeling and analysis for systems of systems: Case study synopses using a taxonomy. *IEEE International Conference on System of Systems Engineering, June 2008.*

Drake, J. M., and Griffen, B. D. (2010). Early warning signals of extinction in deteriorating environments. *Nature*, 467(7314), 456-459.

Ferry, T. S. (1988). Modern accident investigation and analysis (Second Edition ed.). New York: Wiley.

FHWA (2011). Bridge Preservation Guide – Maintaining a State of Good Repair Using Cost Effective Investment Strategies. FHWA Publication Number: FHWA-HIF-11042

Fisher, L. (2011). Crashes, crises, and calamities: How we can use science to read the early-warning signs. New York: Basic Books.

Fischhoff, B. (1975). Hindsight = / = foresight: the effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance 1:* 288–299.

Glantz M.H. (2003). Usable Science: Early warning systems: Do's and Don'ts. Report of workshop, 20-23 October, Shangai, China.

Gros, C., and Markovic, D. (2012). Observing scale-invariance in non-critical dynamical systems. arXiv preprint arXiv:1210.3474.

Haimes, Y.Y. (1977). *Hierarchical Analyses of Water Resources Systems: Modeling and Optimization of Large-Scale Systems*. New York: McGraw-Hill.

Haimes, Y.Y. (1981). Hierarchical holographic modeling. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 11, No. 9, pp.606–617.

Haimes, Y. Y. (2007). Phantom system models for emergent multiscale systems. *Journal of Infrastructure Systems, 13*(2), 81-87.

Haimes, Y.Y. (2008). Phantom system models for risk management of systems of systems. *International Journal of Systems of Systems* Vol. 1, No. 1: 222-236.

Haimes, Y.Y. (2009). *Risk Modeling, Assessment, and Management*, Third Edition. New York: Wiley

Haimes, Y.Y. (2012). Modeling complex systems of systems with phantom system models. *Systems Engineering,* Vol. 15, No.3, 333-346

Haimes, Y.Y. and Jiang, P. (2001). Leontief-based model of risk in complex interconnected infrastructures. *ASCE Journal of Infrastructure Systems*, Vol. 7, No. 1, pp.1–12.

Haimes, Y.Y., Andrijcic, E., and Guo, Z. Demonstrating the Predominant Contributions of Shared State Variables in Modeling Interdependent Systems Using Phantom System Models. Working Paper.

Hawkins, S.A., and R. Hastie. (1990). Hindsight: biased judgments of past events after the outcomes are known. *Psychological Bulletin* 107: 311–327.

Hines et al, (2011) . Estimating Dynamic Instability Risk by Measuring Critical Slowing Down. *IEEE Power and Energy Society Meeting*, pp. 1-5.

Hollnagel, E., and Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International Journal of Man-Machine Studies*, *18*(6), 583-600.

Hollnagel, E. (2002). Understanding accidents-from root causes to performance variability. In *Human Factors and Power Plants*, 2002. Proceedings of the 2002 IEEE 7th Conference on (pp. 1-1). IEEE.

Hollnagel, E. (2004). Barriers and accident prevention. Ashgate Pub Limited.

Kaplan, S and Garrick, B. J. (1981). On the Quantitative Definition of Risk. *Risk Analysis*. 1(1):11-27.

Jiang, Y., and Sinha, K. C. (1989). Dynamic optimization model for bridge management systems. *Transportation Research Record*, (1211).

Johnson, C. (2006). What are emergent properties and how do they affect the engineering of complex systems? *Reliability Engineering and System Safety*, Vol. 91, pp. 1475–1481.

Kaplan S, Haimes YY, Garrick BJ. (2001). Fitting hierarchical holographic modeling into the theory of scenario structuring and a resulting refinement of the quantitative definition of risk. *Risk Analysis*; 21(5):807–815.

Kim, T., Lovell, D. J., and Paracha, J. (2001, January). A new methodology to estimate capacity for freeway work zones. In *Transportation Research Board Annual Meeting CD-ROM*.

Ko, J. M., and Ni, Y. Q. (2005). Technology developments in structural health monitoring of large-scale bridges. *Engineering Structures*, 27(12), 1715-1725.

Kyrtsou, C., and Labys, W. C. (2007). Detecting positive feedback in multivariate time series: the case of metal prices and US inflation. *Physica A: Statistical Mechanics and its Applications*, 377(1), 227-229.

Lakats, L. M., and Pate-Cornell, M. E. (2004). Organizational warning systems: A probabilistic approach to optimal design. *Engineering Management*, IEEE Transactions on, 51(2), 183-196.

Langford, J. W. (1995). Logistics: Principles and Applications. McGraw Hill. p. 488.

Lasdon, L.S., Schoeffler, J.D. (1966). Decentralized Plant Control. *ISA Transactions* 5 (175-183)

Leontief, W.W. (1951a). Input/output economics. *Scientific American*, Vol. 185, No. 4.

Leontief, W.W. (1951b). *The Structure of the American Economy 1919-1939*, 2nd edition, New York: Oxford University Press.

Leveson, N. (1991). Software safety in embedded computer systems. *Commun. ACM* 34, 2 (February 1991), 34-46. DOI=10.1145/102792.102799 *http://doi.acm.org/10.1145/102792.102799*

Leveson, N. (2004). A new accident model for engineering safer systems. *Safety Science, 42*(4), 237-270.

Lewe, J., DeLaurentis, D., Mavris, D. (2004). Foundation for Study of Future Transportation Systems Through Agent-Based Simulation. *24th International Congress of the Aeronautical Sciences (ICAS), Yokohama, Japan, August 2004.*

Lian, C. and Haimes, Y.Y. (2006). Managing the risk of terrorism to interdependent infrastructure systems through the dynamic inoperability input-output model. *Systems Engineering*, Vol. 9, pp.241–258.

Ljung, L. (2010). "Perspectives on system identification," *Annual Reviews in Control*, Vol. 34, pp. 1-12, April.

Lloyd, S. and Lloyd, T. (2003). Bits and bucks: Modeling complex systems by information flow, *MIT Engineering Systems Division, Working Paper Series ESD-WP-2003-01.17*.

Lundberg, J., Rollenhagen, C., and Hollnagel, E. (2009). What-You-Look-For-Is-What-You-Find–The consequences of underlying accident models in eight accident investigation manuals. *Safety Science*, 47(10), 1297-1311.

Mabsout, M. E., Tarhini, K. M., Frederick, G. R., and Tayar, C. (1997). Finite-element analysis of steel girder highway bridges. *Journal of Bridge Engineering*, *2*(3), 83-87.

Maier, M. W. (1998). Architecting principles for systems-of-systems. *Systems Engineering*, *1*(4), 267-284.

National Council on Public Works Improvement (1988). Fragile Foundations: A Report on America's Public Works. Washington D.C.: Government Printing Office.

National Research Council (2004). Accident Precursor Analysis and Management: Reducing Technological Risk Through Diligence. Washington, DC: The National Academies Press.

Neves, L. A., Frangopol, D. M., and Petcherdchoo, A. (2006). Probabilistic lifetime-oriented multiobjective optimization of bridge maintenance: Combination of maintenance types. *Journal of Structural Engineering*, *132*(11), 1821-1834.

Nowak, A. S. (1993). Live load model for highway bridges. *Structural Safety*, *13*(1), 53-66.

NTSB (2007): Update on NTSB Investigation of Collapse of I-35W Bride in Minneapolis. National Transportation Safety Board. August 8, 2007. Retrieved December 1.

NTSB (2008): Design Errors Factor in 2007 Bridge Collapse. Frederic J. Frommer (Associated Press). November 13, 2008. Retrieved November 13.

Ottino, J. M. (2003). Complex systems. *AIChE Journal*, Vol.49, No.2, February, pp. 292-299.

Ottino, J. M. (2004). Engineering complex systems. *Nature*, 427(6973), 399-399.

Page, S. E. (1999). Computational models from A to Z. *Complexity*, Vol 5, No. 1, pp. 35-41.

Parker, J.M. (2010). Applying a System of Systems Approach for Improved Transportation. *S.A.P.I.E.N.S.* Vol. 3, No. 3.

Paté-Cornell, M. E. (1986). Warning systems in risk management. *Risk Analysis*, 6(2), 223-234.

Perrow, C. (1984). *Normal accidents: Living with high-risk technologies*. New York: Basic Books.

Perrow, C. (2011). *The next catastrophe: Reducing our vulnerabilities to natural, industrial, and terrorist disasters (new in paper)* Princeton University Press.

Phares, B. M., Rolander, D. D., Graybeal, B. A., and Washer, G. A. (2000). Studying the reliability of bridge inspection. *Public Roads*, *64*(3).

Phares, B. M., Washer, G. A., Rolander, D. D., Graybeal, B. A., and Moore, M. (2004). Routine highway bridge inspection condition documentation accuracy and reliability. *Journal of Bridge Engineering*, 9(4), 403-413.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. (1992). Runge-Kutta Method. Numerical Recipes in FORTRAN: The Art of Scientific Computing, 704-716.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Tucker, P.K. (2005), Surrogate-based analysis and optimization. *Progress in Aerospace Sciences*, 41, 1-28.

Qureshi, Z. H. (2007). A review of accident modelling approaches for complex socio-technical systems. *In Proceedings of the twelfth Australian workshop on Safety critical systems and software and safety-related programmable systems*-Volume 86 (pp. 47-59). Australian Computer Society, Inc.

Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science, 27*(2), 183-213.

Reason, J. (1990). The contribution of latent human failures to the breakdown of complex systems. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences, 327*(1241), pp. 475-484.

Reason, J. T., and Reason, J. T. (1997). Managing the risks of organizational accidents. Ashgate Aldershot.

Romano, M. C., Thiel, M., Kurths, J., and Grebogi, C. (2007). Estimation of the direction of the coupling by conditional probabilities of recurrence. *Physical Review* E, 76(3), 036211.

Sage, A.P. (1977). *Methodology for Large Scale Systems*. New York: McGraw-Hill.

Sage, A.P. (1992). *Systems Engineering*. New York: Wiley.

Sage, A.P. (1995). *Systems Management for Information Technology and Software Engineering*. New York: Wiley.

Sage, A.P., Rouse W.B. (eds.) (1999). *Handbook on Systems Engineering and Management*. New York: Wiley.

Sage, A. P., and Cuppan, C. D. (2001). On the systems engineering and management of systems of systems and federations of systems. *Inf.Knowl.Syst.Manag*, *2*(4), 325-345.

Sage, A. P. (2003). Conflict and risk management in complex system of systems issues. *Systems, Man and Cybernetics, 2003. IEEE International Conference on, , 4* 3296-3301 vol.4.

Sage, A. P., and Biemer, S. M. (2007). Processes for system family architecting, design, and integration. *Systems Journal, IEEE*, 1(1), 5-16.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., and Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461(7260), 53-59.

Scherer, W. T., and Glagola, D. M. (1994). Markovian models for bridge maintenance management. *Journal of Transportation Engineering*, *120*(1), 37-51.

Shalizi, C. (2006). Methods and techniques of complex systems science: An overview. *arXiv:nlin/0307015* v4, 24 March, 96p.

Simon, H. A. (1979). Rational decision making in business organizations. *The American economic review*, 69(4), 493-513.

Smirnov, D. A., and Andrzejak, R. G. (2005). Detection of weak directional coupling: Phase-dynamics approach versus state-space approach. *Physical Review E*, 71(3), 036207.

Smith, B. L., and Demetsky, M. J. (1997). Traffic flow forecasting: comparison of modeling approaches. *Journal of transportation engineering*, *123*(4), 261-266.

Society for Automotive Engineers (1967). Design Analysis Procedure For Failure Modes, Effects and Criticality Analysis (FMECA). ARP926.

Stamatis, D. H. (2003). Failure mode and effect analysis: FMEA from theory to execution. Asq Press.

Testa, R. B., and Yanev, B. S. (2002). Bridge maintenance level assessment. *Computer-Aided Civil and Infrastructure Engineering*, *17*(5), 358-367.

Thissen, W.A.H., Herder P.M. (2009). System of Systems Perspective on Infrastructures. Published in *System of Systems Engineering*, Ed. Jamshidi, M.. New York: John Wiley & Sons.

Thoft-Christensen, P. (2000). *Modelling of the deterioration of reinforced concrete structures*. Dept. of Building Technology and Structural Engineering.

Thompson, P. D., Small, E. P., Johnson, M., and Marshall, A. R. (1998). The Pontis bridge management system. *Structural engineering international*, *8*(4), 303-308.

UNEP (2012). Early Warning Systems: A State of the Art Analysis and Future Directions. Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP), Nairobi

UN-ISDR (2005) Platform for the Promotion of Early Warning, Four Elements of People Centered Early Warning Systems, presented at the Virtual Symposium, Public Entity Risk Institute: Early Warning Systems – Interdisciplinary Observations and Policies from a Local Government Perspective. April 18-22, 2005.

U.S. Department of Transportation (U.S. DOT). (1995). Recording and coding guide for the structure inventory and appraisal of the nation's bridges, Ofc. of Engrg., Bridge Div., Bridge Mgmt. Branch, Federal Highway Administration, Washington, D.C.

Vesely, W. E., Goldberg, F. F., Roberts, N. H., and Haasl, D. F. (1981). Fault tree handbook (No. NUREG-0492). Nuclear regulatory commission washington dc.

Vincoli, J. W. (2006). Preliminary Hazard Analysis. *Basic Guide to System Safety*, Second Edition, 65-83.

Wang, W., Wang, J., and Kim, M. S. (2001). An algebraic condition for the separation of two ellipsoids. *Computer aided geometric design*, 18(6), 531-539.

Weeks, John A. III (2007). I-35W Bridge Collapse Myths And Conspiracies. John A. Weeks III.

Wiener, N. (1948). *Cybernetics, or Control and Communication in the Animal and the Machine*. MA: The Technology Press.

Wolff, U. (1990). Critical slowing down. *Nuclear Physics B-Proceedings Supplements*, 17, 93-102.