Approaches to Modeling and Simulation for Dynamic, Distributed Cyber-Physical Systems

Kurt Rohloff, Partha Pal,
Michael Atighetchi, Richard Schantz
BBN Technologies
Cambridge MA, USA
krohloff, ppal
matighet, schantz @bbn.com

Kishor Trivedi
Duke University
Durham NC, USA
kst@ee.duke.edu

Christos Cassandras
Boston University
Boston MA, USA
cgc@bu.edu

Abstract— **In this paper we discuss challenges and new directions in modeling and simulation for effects-based what-if and sensitivity analysis of dynamic, distributed cyber-physical systems. We are motivated on one hand by the critical need to reliably understand how mission-critical cyber-physical systems would respond to unanticipated effects, and on the other hand by the technology gap that has prevented us from doing so until now. Modern cyber-physical systems are very large distributed systems, covering wide geographic areas with time-critical operations, asynchronous state updating and implicit interactions through resource sharing. Canonical examples of such systems include the national electric power generation and distribution grid, computing infrastructure, communication networks and manufacturing systems that are built upon multiple layers of software and physical resources. We address three main aspects of challenges and next steps to the modeling and simulation of these systems based on 1) a revised foundation for model representation, 2) advanced model-analytic tools and 3) a general adaptable and reusable simulation environment. Our suggested approach to these challenges incorporates both the generalization and repurposed use of existing technologies to create a rigorous and justifiable foundation for survivability analysis in large-scale cyber-physical systems.**

Keywords-cyber-physical systems; modeling; simulation; verification; challenges

I. INTRODUCTION

The safety and survivability of cyber-physical systems and systems-of-systems are and will continue to be crucially important for the maintenance of modern society as we know it. Examples of these cyber-physical systems include the national electric power generation and distribution grid, computing infrastructure, transportation systems such as train networks, communication networks and manufacturing systems that are built upon multiple layers of software and physical resources.

Our society's reliance on cyber-physical systems requires us to be able to model, simulate and analyze these systems for achieving critical properties, including safety and survivability in the face of combinations of accidents, operator error, cyber attacks, physical attacks and "wrath of god" incidents such as hurricanes, earthquakes and blizzards. We need to be able to model, simulate and analyze both 1) the likely direct affects of these adverse events on cyber-physical systems and 2) indirect effects, caused by the ability of partial component system faults to propagate, leading to large-scale outages and potentially catastrophic events. The ultimate goal of these analyses is to take corrective action in a timely manner (either proactively or reactively) to remove or limit the affects of these adverse events.

Cyber-physical systems are particularly difficult to model and simulate because their components mix many different system modalities, and these system components interact both implicitly and explicitly. These systems, both in their physical and cyber domains blend time-driven behaviors (with discrete, continuous or hybrid state-updating) and event-driven behaviors (which may or may not be expressible as easily-modeled regular languages.)

In this paper we discuss challenges and next steps for the effective modeling and simulation of large-scale cyber-physical systems. The remainder of this paper is organized as follows. In the following section we outline challenges of modeling and simulation for cyber-physical systems where failures can propagate, leading to unexpected and disastrous results. In Section III we introduce our vision for the next steps for modeling and simulation of wide-scale, distributed cyber-physical systems. In each of the following sections after this we discuss aspects of these challenges.

II. CHALLENGES

Specific challenges in modeling and simulation for these cyber-physical systems arise from the fact that they are large and highly interconnected, with a very large or effectively infinite number of:

- States, internal interactions, and potential configurations, often coping with or embedding adaptive, non-deterministic, and evolving (e.g., learning) behaviors. Complicating matters further are

bugs and misconfigurations that are generally unavoidable in software systems.

- Inter-component interactions, not only including distributed control and data flow between components, but also including timing dependencies and indirect or even inadvertent coupling through shared resources (e.g., CPU, memory, and communications) which are frequently less explicit or hidden.
- Ranges of inputs, including sensor inputs of unpredetermined or unbounded ranges and very large or continuous valued inputs, such as external signals.
- External factors that can include wide-ranging and unanticipated effects on system behavior, arising from hostile actions such as cyber-attacks and bombings, extreme weather conditions, collision with physical elements (e.g., airplane crashes), and interaction (benign, accidental, or hostile) with external systems (e.g., signal noise, system components competing for the same resources in the face of conflicting operational requirements, etc.…)

Another challenge is that mathematically, the combination of time-driven and event-driven dynamics gives rise to *hybrid* systems. Not only do we need to combine different types of mathematical models but also to deal with issues of synchronization in these large state-space systems. It is highly nontrivial for a time-driven model to handle concurrency of events, and lack of developer understanding of these concurrent interactions (both cyber and physical) is the common source of many bugs in the cyber components.

A number of real, well-known failures illustrate the challenges associated with modeling and simulation in complex cyber-physical systems. One example of this from the military domain is the failure of the Alpha battery of the Patriot missile system at Dhahran, Saudi Arabia on February 25, 1991, which resulted in the deaths of 28 US Army reservists. This failure was attributed to the lack of understanding of the time-varying survivability of the system and component interactions which seriously degraded after 20 hours of continuous operation due to a known bug [20]. Another example from the civilian cyber-physical system space is the failure of ATT's 4ESS switching system on January 15, 1990 where the "initial address message" received in quick succession from a recovering switch caused a cascading failure in a series of switches, incapacitating a major part of ATT's network [15].

Our experience with very large, mission-critical, interdependent, cyber-physical system prototypes (with on the order of thousands of components), such as the ARMS naval distributed computing environment [14], platooned UAV avionics [10], protection for the JBI information environment [13] and TRANSCOMM military logistics and planning systems [19], informs us that such problems (in both the cyber and physical domains) of unexpected fault propagation are not uncommon and are exceedingly difficult to model by focusing on individual components. Furthermore, it is especially difficult to model how these propagating component faults impact the survivability of the larger system-of-systems.

The examples above of adverse events cascading and leading to large-scale system failures, potentially impacting society as a whole in unexpected ways, motivate the need for approaches to the modeling, simulation and analysis of complex cyber-physical systems and systems-of-systems. Although methods for modeling such systems have begun to emerge, these strategies have been developed primarily for specific domains and/or for addressing specific sub-aspects of the encompassing modeling and simulation problems.

An important limitation faced by researchers and system administrators for evaluating these large-scale cyber-physical systems is that there is currently no method to *estimate* the *effects* arising from interactions between cascading failures in the cyber components of cyber-physical systems without actually deploying them in the system and subjecting them to conditions expected to cause real cascading failures. This approach is prohibitively expensive and generally unfeasible for safety-critical one-of-a-kind systems such as the US electric grid. Unfortunately, as a cost-effective alternative to this "red-team" testing of networked information systems, system designers and network administrators commonly use mental exercises to predict the effects of cascading failures even though this approach is self-evidently incomplete, error-prone and subject to investigator bias.

III. VISION

The salient feature of the representative cyber-physical systems we mention above is the various ways local faults can propagate due to multiple types of interdependency. Although we can model many interdependent systems using knowable parameters, current modeling formalisms do not lend themselves to distributed fault propagation analysis for very large interdependent systems due to the lack of scalable mathematical or computational techniques to analyze fault propagation using the models. Consequently, current fault propagation and survivability analysis methods are infeasible for interdependent systems. Model abstraction and tuning are additional challenges. Improper model abstractions lead to models that are ineffectual due to a lack of detail or become too complex to analyze efficiently. Even if methods exist to analyze large, high-fidelity models for interdependent systems, practitioners are still prevented from using these models due to the economics of tuning the large number of parameters typically present in these models. As a result, models of interdependent systems rely on simpler measures conditioned on parameters with little observable reality. Additionally, there are few modeling

environments than can host diverse sets of models for the effective computational analysis of the effects of propagating faults. There is also the problem of integrating component models at various layers into an overall system-of-systems model. Although hierarchical model decomposition and refinement are easy to explain on paper, tools supporting these system-of-systems modeling paradigms have been weak and forced designers to use models at the wrong abstraction layer.

With this in mind, we identify the following three goals to address the specific challenges outlined above for the modeling, simulation and analysis of large-scale, distributed cyber-physical systems:

- A representation of components taking into account the effects of uncertainty in model structure and providing flexibility to select parameters based on desired level of fidelity.
- An analysis methodology for a more effective and computationally tractable approach to simulating these models both in isolation and when coupled with other component models of cyber-physical systems.
- A general adaptive and reusable environment to host the models when simulated.

Some current modeling and simulation approaches have the right ingredients to address the challenges we outline above, but these kernels of solutions have not yet been put together in the right way for successful use with real systems. In particular, a number of recent mini-successes lead us to believe that the approaches outlined in this paper are becoming much more viable.

As an example, work in the area of modeling uncertainty (both in practice and theory) has been used for the modeling of IBM clusters with rejuvenation [21], reliability modeling of SIP on IBM WebSphere [16], and reliability analysis of the Current Return Network in the Boeing 787 aircraft for the purpose of certification by the FAA [18]. SIP protocol reliability analysis resulted in a new interacting Markov chains-based method of calculating the number of dropped calls and was responsible for the sale of the system by IBM to a Telco customer.

Over the past several years the theory of Perturbation Analysis (PA) has been extended to hybrid systems, providing ample evidence of its power and general scope. The idea of a "calculus for perturbation propagation" has been applied in a number of domains including manufacturing systems, computer systems, transportation, and command-control systems [5]). The discrete event dynamic system modeling framework, developed in conjunction with Perturbation Analysis, was adopted by The MathWorks in developing SimEvents, a discrete event simulator which is now part of the MATLAB/Simulink suite of software products. [12]

Over the next sections we discuss in more detail the three goals and associated challenges outlined above and our vision for their resolution.

IV. Representation of Model Components

In this section we suggest mathematical bases for the computationally efficient modeling and simulation of distributed cyber-physical system components. We decompose this discussion into two subsections. In the first subsection we address the *aleatory uncertainty* - uncertainty in the possible structure of models. We focus on the problem of manipulating and analyzing models of very large interdependent systems. We describe new, generalized mathematical and computational approaches to modeling and model manipulation for these systems. In the second phase we address *epistemic* or *parametric uncertainty* - uncertainty in model tuning. We focus on the problem of identifying which model abstraction level and which parameters would be most effective to tune accurately for higher fidelity modeling and survivability analysis. In the cyber-physical systems-of-systems we discuss, we denote the models of individual subsystems as component models.

A. Addressing Aleatory Uncertainty: Generalizing Reliability and Availability Analysis Approaches to Very Large Systems

There have been recent promising theoretical developments of distributed modeling formalisms we can generalize and enhance to address the practical problem of modeling large systems with interdependent fault propagation for time-varying distributed cyber-physical system modeling and simulation. In this subsection we identify some of these approaches that show promise for the next steps of modeling and simulation for cyber-physical systems.

Recent work has shown the viability of using stochastic hybrid system models as formalisms to model distributed interdependent systems [4, 6]. Another promising modeling framework for interdependent survivability analysis is a max-plus algebra that capture the dynamic behavior through the two mathematical operations “max” (modeling precedence and general interdependence conditions) and “plus” (the passage of time) [GY94]. Additional promising distributed modeling formalisms include interacting Markov chains [17] via Kronecker algebra [2] and interacting semi-Markov processes via GSMP (Generalized Semi-Markov processes) [8].

B. Addressing Epistemic Uncertainty: Identifying Appropriate Abstraction and Parameters for Effective Modeling of Uncertainty In Model Components

Too often model parameters (if they are even based on physically measurable phenomena) for large, complex, interdependent system models are based on difficult to measure field data, data from systems with similar functionality, or even by guessing. This prevents effective

survivability analysis using these models. If we provide methodologies to identify which model parameters have the most effect on survivability (or on our ability to assess survivability of a system), practitioners can more effectively direct their efforts to measure or estimate these key parameters and identify effective model abstractions, leading to higher-fidelity assessments of survivability.

A parametric sensitivity analysis [ST07] approach is one avenue to determine the most effective parameters to be more accurately measured for survivability analysis. A promising approach is to develop numeric computation methods that analyze the propagation of epistemic uncertainty [YST01, STAM02] across system components to. This will allow us to model and analyze how uncertainty can propagate across system components during simulation, ultimately allowing us to more selectively tune model parameters for higher-fidelity, large-scale modeling and simulation..

V. Analysis Methodology

Despite the promise of increased capabilities from newer model analysis methodologies, current approaches are still insufficient for the simulation and analysis of fault propagation in modeling interdependent systems that contain both implicit and explicit interactions. Current model analysis techniques are frequently based on static development techniques, and testing according to some measures of acceptable coverage. These techniques are generally not applicable to dynamic, distributed cyber-physical systems, which have large (or potentially infinite) numbers of possible states and interactions, continuous (or approximately continuous)-valued inputs, and unknown effects from external factors. In this section we address some approaches to this specific challenge. We outline the theoretical bases for two general "next-steps" approaches for analyzing models of large-scale cyber-physical systems that show promise for use in large-scale practical settings. These approaches include the analysis of uncertainty propagation and a theory of software reliability.

A. A Calculus for Uncertainty Propagation Analysis

Since the critical parameters involved in survivability analysis models of cyber-physical systems are usually unknown, it is important to understand how they affect the behavior of the system, i.e., how a parametric perturbation generates a perturbation in the state of the system and then how this perturbation propagates throughout the system. A promising "calculus" for perturbation propagation analysis has been developed under the theory of Perturbation Analysis (PA) for discrete event systems [10] that is implemented numerically for simulation. In this approach, sensitivities of important performance metrics with respect to parameters of interest can be estimated based on observed system data and, therefore, the essential tradeoffs involved in assessing the operation of a system can be analyzed. Thus, PA can be a data-driven approach to select which model parameters survivability assessments are most sensitive to. This will help practitioners make informed decisions about what "real", observable parameters should be used in interdependent system models.

B. A Theory of Software Reliability:

In some sense, the hardest and most challenging aspect of modeling and simulating cyber-physical systems is accounting for the distributed, interacting software components. These software components follow few physical laws (such as Maxwell's laws) that slow the propagation of faults, and traditional modeling can be difficult to apply, making analysis exceedingly difficult. The existing theory of software reliability, based on the traditional Bohrbug model [9], needs to be completely revised to reflect two new types of software bugs: "Mandelbugs", which are endemic in interdependent systems and "Aging-Related Bugs", which are endemic to dynamic software systems. Mandelbugs activate faults in software due to complicated, unexpected conditions in interdependent systems, but do not immediately cause failures. Aging-related bugs cause errors which worsen over time. Both types of these bug types cause errors when propagated, but due to nondeterminism introduced by interdependence, errors caused by Mandelbugs and aging-related bugs do not repeatably manifest as failures. On top of this, there is typically a long delay between the fault activation and the final failure occurrence. Consequently, it is difficult to analyze fault propagation from faults that cause system failures due to these bugs. A useful approach to develop a new software survivability theory could start by acquiring and studying failure histories of existing systems to develop techniques for survivability assessment and mitigation methods for Mandelbugs and aging-related bugs by a combined approach of analytic modeling, simulations, measurements and optimization. The goal would be to develop a theory of software reliability that can be used to develop new analysis techniques specifically inclusive of interdependent software components.

VI. Adaptable and Reusable Simulation Environment

Although there has been a lot of work to develop modeling and simulation environments to host individual simulation and modeling technologies with their relevant datasets, there has been little work to develop comprehensive end-to-end environments to host many different types of modeling and simulation technologies with a wide array of associated knowledge stores. We propose a comprehensive system solution for high-grade integrated modeling, analysis and decision-support capabilities. The primary purpose of this modeling and simulation system is to support experimentation and planning by providing an environment for repeatable experiments in order to confirm analysis of complex systems (such as for cyber-physical systems.) This

system should support experimentation in the following important ways:

- It should provide modeling and simulation applications with a common data environment from which they can easily retrieve the data they need and an environment into which they can store and share their results.
- It should provide the means to rapidly insert new data, as requested by scientists, analysts and planners (including new live data feeds) in such a way that it is seamlessly integrated with existing data and made available through configuration managed data services. This provides an evolutionary (least disruptive) path for experimentation, analysis.
- It should provide the scientists, analysts and planners with a comprehensive user interface allowing them to inspect all of the data and results through a web browser from any location at any time.
- It should provide the means to capture and analyze all of the results obtained during experiments for comparison with new results obtained during subsequent experiments.

There is a long legacy of information systems developed to support interactive modeling environments that address some of the concerns outlined above [1]. Unfortunately, this work has primarily focused on integrated modeling environments that leverage *structural* data, and there has been little work on developing integrated modeling environments that can integrate wider ranges of unstructured, semi-structured, and traditional structural data in the Semantic Web context [3] often prevalent when attempting to model cyber-physical systems deployed in a "real" environment.

In conjunction with analytical methods and modeling formalisms outlined above, these formalisms enable the development of simulation tools that judiciously combine analytic, analytic-numeric and simulative solutions to assess the survivability of larger interdependent systems. Although modeling safety-critical systems in practice often involves modeling rare events driving faults that propagate, experimental methods lack the ability to explore such events and quantify failure likelihood. This suggests that methods to deal with rare-event issues, including the theory of large deviations and "fast simulation" methods [7] may also be useful. We feel that additional approaches need to be advanced for larger-scale systems. Theoretical or practical advancements in this area would be ground-breaking.

VII. CONCLUSION

An important realization in our outline of next steps for the modeling and simulation of large-scale cyber-physical systems is that it is not always essential to estimate all parameters accurately for effective high-fidelity modeling. A benefit of progressing along the steps we outlined for addressing modeling and simulation through the analysis of uncertainty due to failures and fault propagation is the possibility of developing a framework to understand how model uncertainty propagates for a more nuanced analysis of survivability distributed, safety-critical cyber-physical systems. We think our sensitivity analysis methods will be able to guide decision making on which component models should be iteratively refined to lower model abstraction levels. This effectively increases global model fidelity and enhances distributed modeling through lower-level "observable" parameters with physical meaning. Our approach is based on an anytime computing paradigm that iteratively improves model fidelity until the limits of available computational techniques are reached to get a "best effort" out of available modeling and survivability analysis tools.

REFERENCES

1. Anderson, B., & Flynn, J. CASES: A System for Assessing Naval Warfighting Capability. Proceedings of the 1990 Symposium on Command and Control Research. 1990
2. Bao, Y., Bozkurt, I., Dayar, T., Sun, X., and Trivedi, K., "Decompositional Analysis of Kronecker Structured Markov Chains", in Electronic. Trans. on Numerical Analysis, 2008.
3. Berners-Lee, T., Hendler, J., & Lassila, O. (2001, May). The Semantic Web. Scientific American .
4. Cassandras, C., and Lygeros, J., "Stochastic Hybrid Systems", Taylor and Francis, 2006.
5. Cassandras, C. "Joint Air Operations (JAO) Mission Planning Problem" http://vita.bu.edu/cgc/alpha/Scenario1.htm
6. Cassandras, C., Wardi, Y., Melamed, B., Sun, G., Panayiotou, C.G., "Perturbation Analysis for On-Line Control and Optimization of Stochastic Fluid Models", IEEE TAC, 2002.
7. Dembo, A., Zeitouni, O. Large Deviations Techniques. Jones and Bartlett 1993.
8. Glynn, P. W., "A GSMP Formalism for Discrete Event Systems", in Proc. IEEE, 1989.
9. Grottke, M., and Trivedi, K. S., "Fighting Bugs: Remove, Retry, Replicate and Rejuvenate," IEEE Computer, vol. 40, no. 2, 2007.
10. Ho, Y.C., Cao, X., "Perturbation Analysis of Disc. Event Dyn. Systems", Kluwer, 1991.
11. Loyall, J, Schantz R, Corman, D, Paunicka, J, Fernandez, S. "A Distributed Real-time Embedded Application for Surveillance, Detection, and Tracking." RTAS, 2005.
12. The Mathworks SimEvents event simulation software. http://www.mathworks.com/products/simevents/
13. Pal P, Webber, F and Schantz R "The DPASA Survivable JBI-- A High-Water Mark in Intrusion-Tolerant Systems", EuroSys Workshop on Advances in Intr. Tol. Sys., Lisbon, 2007.
14. Rohloff, K, Gabay, Y, Ye, J and Schantz, R. "Scalable, Distributed, Dynamic Resource Management for the ARMS Distributed Real-Time Embedded System." WPDRTS, 2007.
15. Travis, Paul. "Why the AT&T network crashed." Telephony, Jan 22, 1990 p11.
16. Trivedi, K. S., Wang, D., Hunt, D. J., Rindos, A., Smith, W. E., Vashaw, B., "Availability Modeling of SIP Protocol on IBM(c) WebSphere(c)", Proc. PRDC 2008.
17. Trivedi, K.S., "Probability & Statistics with Reliability, Queuing and Computer Science Applications", (2nd ed.), John Wiley, 2001.
18. Trivedi, K., Wang, D., Ramesh, A., Sharma. T., et al, "Reliability Estimation Methods for Large Networked Systems", US Patent, Sept. 2009.
19. Tustin, J. P., W. F. Ferguson, C. N. Van Groningen, C. J. Keyfauver, and E. R. Beeker. "Integrating MIDAS and ELIST into the AMP HLA Federation." Simulation Interoperability Workshop, Logistics and Enterprise Models Forum, Fall 2001.
20. US Government Accounting Office. "Patriot missile defense, Software problem led to system failure at Dharhan, Saudi Arabia; GAO report IMTEC 92-26". http://www.gao.gov/products/IMTEC-92-26.
21. Vaidyanathan, K., Harper, R. E., Hunter, S.W., and Trivedi, K. S., "Analysis and Implementation of Software Rejuvenation in Cluster Systems", ACM SIGMETRICS, 2001.