2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing

Towards Integrated Simulation of Cyber-Physical Systems: A Case Study on Intelligent Water Distribution

Jing Lin, Sahra Sedigh, and Ann Miller
Department of Electrical and Computer Engineering
Missouri University of Science and Technology
Rolla, USA 65409
Email: {jlpg2,sedighs,milleran}@mst.edu

***Abstract*—In cyber-physical systems (CPSs), embedded computing systems and communication capability are used to streamline and fortify the operation of a physical system. Intelligent critical infrastructure systems are among the most important CPSs and also prime examples of pervasive computing systems, as they exploit computing to provide "anytime, anywhere" transparent services. The existing body of knowledge includes techniques for assessment, modeling, and simulation of physical and cyber infrastructures, respectively, but such isolated analysis is incapable of fully capturing the interdependencies that occur when they intertwine to create a CPS. Fundamental differences exist between the attributes of cyber and physical components, significantly complicating representation of their behavior with a single comprehensive model or simulation tool. This paper articulates the challenges present in integrated simulation of a CPS, where the goal is to accurately reflect the operation and interaction of the cyber and physical networks that comprise the system. A solution is presented for the CPS domain of intelligent water distribution networks, using EPANET and Matlab to represent the physical water distribution network and the decision support algorithms used to control the allocation of water, respectively.**

***Index Terms*—cyber-physical systems; simulation; water distribution networks; EPANET; MATLAB**

I. INTRODUCTION

In *cyber-physical systems* (CPSs), embedded computing and communication capability are used to streamline and fortify the operation of a physical system [1], [2]. Information collected by sensors from the physical infrastructure is fed to the cyber components, which use computing hardware and software and communication links to intelligently control the physical components. There is typically no one-to-one correspondence between the elements of the two networks, which complicates the understanding of their interaction.

Critical infrastructure systems reliant on intelligent monitoring and control are among the most important CPSs and also prime examples of pervasive computing systems, as they exploit computing to provide "anytime, anywhere" transparent services. While the added intelligence offers the promise of increased utilization, its impact must be assessed, as unrestricted cyber control can actually lower the reliability of existing infrastructure systems.

As a practical example, *water distribution networks (WDNs)* are an emerging CPS domain. Physical components, e.g., valves, pipes, and reservoirs, are coupled with the hardware and software that supports intelligent water allocation. An example is depicted in Fig. 1. The primary goal of WDNs is to provide a dependable source of potable water to the public. Information such as demand patterns, water quantity (flow and pressure head), and water quality (contaminants and minerals) is critical in achieving this goal, and beneficial in guiding maintenance efforts and identifying vulnerable areas requiring fortification and/or monitoring. Sensors dispersed in the physical infrastructure collect this information, which is fed to algorithms (often distributed) running on the cyber infrastructure. These algorithms provide decision support to hardware controllers that are used to manage the allocation (quantity) and chemical composition (quality) of the water. As WDNs become larger and more complex, their reliability comes into question. The increasing use of Supervisory Control and Data Acquisition (SCADA) systems also raises concerns about the vulnerability of WDNs [3].

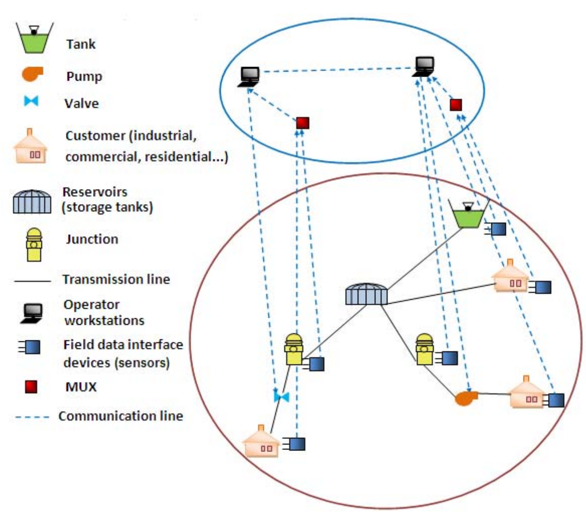

Fig. 1. Cyber and physical components of an intelligent WDN

978-0-7695-3929-4/09 $26.00 © 2009 IEEE
DOI 10.1109/DASC.2009.140

Modeling of any CPS is hampered by the need to model both the cyber (software, communication network, computing hardware) and the physical infrastructure (physical components and their interactions). Ideally, a single model would encompass both the physical and cyber system semantics in a meaningful way such that the effects of a specific event are reflected in the reaction of either a cyber or a physical component of the system. Interdependencies among the cyber and physical components, in operation and failure, present a major challenge, as they invalidate simplified models that assume components fail independently. Moreover, existing models focus on the physical infrastructure, often neglecting the sophisticated interactions that occur among complex intelligent components and their effect on the operation of the system as a whole.

Simulation of CPSs presents a similar challenge. Accurate representation of a CPS encompasses three aspects: computing, communication, and the physical infrastructure. Fundamental differences exist between the attributes of cyber and physical components, significantly complicating representation of their behavior with a single comprehensive model or simulation tool. Specialized simulation tools exist for the engineering domains represented in critical infrastructure, including power, water, and transportation. These tools have been created with the objective of accurately reflecting the operation of the physical system, at high spatial and temporal resolution. As is the case with specialized models of physical systems, intelligent control is not reflected in these tools. Despite the existence of simulation tools for cyber aspects such as computing and communication, differences in temporal resolution and data representation and the lack of well-defined interfaces pose considerable challenges to linking these simulation tools in a fashion that accurately represents the CPS as a whole.

This paper articulates the challenges present in integrated simulation of CPS, where the goal is to accurately reflect the operation and interaction of the cyber and physical networks that comprise the system. A solution is presented for the CPS domain of intelligent WDNs. The proposed solution utilizes EPANET and Matlab to represent the physical water distribution network and the decision support algorithms, respectively. Communication between the two simulators replicates the interactions between cyber and physical components of WDNs, and facilitates the observation of physical manifestations of intelligent control decisions. This communication between the simulators takes place without user intervention, as all information relevant to each simulator has been identified and extracted from the output of the other. Information flows from the physical simulator to the cyber simulator, replicating the operation of sensors in the physical infrastructure. The cyber simulator processes this data, and provides decision support for water allocation, in the form of setting for control elements in the physical infrastructure. This information is provided to the physical simulator, which applies these settings. This process repeats for the duration of the simulation, as it would in the actual operation of a CPS.

The interfacing technique developed overcomes challenges that arise from differences in spatial and temporal resolution, syntax, and data formatting between the cyber and physical simulators, enabling the strengths of each specialized simulator to be leveraged in representing the behavior of a CPS. The simulation technique will be instrumental in developing and validating models for CPSs, which is the objective of our research. Insights gained from the WDN domain will be used to extend the models and techniques developed to other CPS domains, with the ultimate goal of creating CPS models that are broadly applicable, yet capable of reflecting attributes specific to each physical domain.

The remainder of this paper is organized as follows. Section II presents background information, including a review of related work. A number of simulation tools for the physical and cyber networks, respectively, and challenges to linking them are described in Section III. The proposed approach to overcoming these challenges and simulating the CPS as a whole is presented in Section IV, as is an example simulation for an intelligent WDN. Section V concludes the paper and describes future research directions.

II. Related work

The simulation technique presented in this paper has been developed as a tool for validation of models for CPSs, an emerging research area related to pervasive computing. The focus of the majority of studies related to CPSs, e.g., [4]–[7] is on interdependencies among different components of critical infrastructure. A relatively comprehensive summary of modeling and simulation techniques for critical infrastructure systems, an important category of CPSs, is provided in [6]. System complexity has been identified as the main challenge in characterizing interdependencies in CPSs [5]. Other challenges include the low probability of occurrence of critical events, differences in the time scales associated with these events, and the difficulty of gathering data needed for accurate modeling.

Several challenges to the development of a generic framework for the design, modeling, and simulation of CPSs are articulated in [8]. Features described as desirable for such a framework include the integration of existing simulation tools, software reusability, and graphical representation of the modeling and simulation environment. The work presented in this paper meets these criteria.

The study most closely related to the work presented in this paper is [9], where a method is proposed for integration of the ns-2 network simulator with the Modelica framework, a modeling language for large-scale physical systems. The paper highlights the challenge of two-way synchronization of the simulators. The key difference between this study and our work is that we link to a specialized simulator capable of accurately representing the operation of the physical infrastructure, in this case a WDN, at high resolution. The WDN simulator, and other related simulation tools are described in the next section of this paper.

III. Simulation Tools and Integration Challenges

Our approach to simulation of a CPS is based on the use of existing simulation tools for the cyber and physical networks, respectively. This choice is due to the powerful capabilities of specialized tools in representing their domain (cyber or physical), which allows the focus of our work to shift to accurate representation of the interactions between the cyber and physical networks.

A. Simulation tools for the physical infrastructure of WDNs

Several tools are available for simulation of the physical water distribution infrastructure. Examples include EPANET, which can capture both quantity and quality of water throughout a distribution network [10]; RiverWeb, which is focused on river basin processes [11]; Water Quality Analysis Simulation Program (WASP), which provides watershed, water quality, and hydrodynamic models [12]. Also considered for our study was Waterspot, which simulates water treatment plants [13]; the Ground Water and Rainmaker Simulators [14], which is mainly a teaching tool; and the General Algebraic Modeling System (GAMS), which provides a high-level modeling system for the mathematical programming and optimization [15].

Among these simulators, EPANET provides the most detailed representation, as it can capture the layout of a WDN and track the flow of water in each pipe, the pressure at each node, the depth of the water in each tank, and the concentration of a chemical substance throughout the network during a simulation period [10]. The simulator is provided at no charge by the Environmental Protection Agency. The extensive capabilities, ease of use, and lack of licensing fees motivated the choice of EPANET as the simulator for the physical infrastructure of WDN in our study.

The most recent release, EPANET 2.0, was the version used. Objects in EPANET can be classified as nodes, links, map labels, time patterns, curves and controls. Each node can in turn be a junction, reservoir, or tank, and each link can be a pipe, pump, or valve. The topology depicted in Fig. 2 is a very simple WDN as visualized by EPANET. It is composed of one reservoir, one tank, one pump, one valve, five junctions, and several pipes that connect these elements. A reservoir is a node that represents an infinite external source or sink of water [16], and is used to model an entity such as a lake, river, or groundwater aquifer. A tank is a node with storage capacity, where the volume of stored water can vary with time during a simulation. A junction is a point in the network where links join together and where water enters or leaves the network. When a junction has negative demand, it indicates that water is entering the network at that point.

Pumps and valves are two primary actuators that can be turned on and off at preset times, or in response to certain conditions in the network. Fluids possess energy, and the total energy per unit weight associated with a fluid is denoted as "head." On many occasions, energy needs to be added to a hydraulic system to overcome elevation differences, or losses arising from friction or other factors. A pump is a device to which mechanical energy is applied and transferred to the water as total head, so it can add more energy to the fluid. The flow through a pump is unidirectional. If the system requires more head than the pump can produce, the pump is shut down. Therefore, pumps can be turned on and off at preset times, when tank levels fall below or above certain set-points, or when the pressure at a certain node falls below or above specified thresholds.

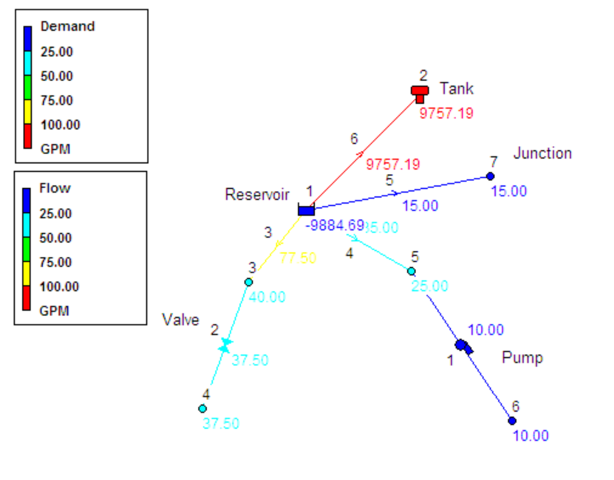

Fig. 2. A simple topology in EPANET

A valve is an element that can be opened or closed to different extents, to vary its resistance to flow, thereby controlling the movement of water through a pipe. The status of each valve can be specified for all or part of the simulation by using control statements. Pipes are links that convey water from one point in the network to another. The direction of water flow is from the end at higher hydraulic head to that at lower head, due to the effect of gravity. A negative label for a flow indicates that its direction opposes that of the pipe.

In the WDN depicted in Fig. 2, the reservoir is providing water to the tank and a number of different junctions. This topology can serve as a simple and abstract representation of a lake that provides water to consuming entities spread throughout a city. The reservoir in this figure always contributes water into the network, so its demand value is negative. The value of the demand indicates the amount of water contributed, in this case 9884.69 gallons per minute (GPM). The tank consumes the highest amount of water. Each junction is also labeled with its demand value, and each pipe with its flow speed. The entire graph is color-coded to simplify the categorization of demand or flow. The demand values of pumps and valves vary in accordance with the nodes they control.

A more complex topology is depicted in Fig. 3, which shows a screen capture at hour 8:00 of a 24-hour simulation period. This figure also depicts node groupings, circled in green, that can facilitate study of a subset of the nodes in the topology.

After simulating the system for the specified duration, EPANET can provide a report in graph, table, or text form. Among the various reports available, the full report provides the most comprehensive data, including the initial and updated

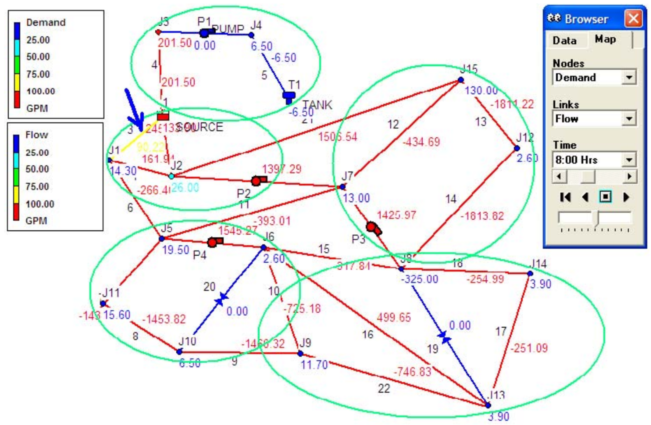

Fig. 3. A more complex topology and node groupings in EPANET

values of all properties of the nodes and links within each simulation time step (one hour by default). The water flow, pressure at each node, depth of water in tanks and reservoirs, and concentration of chemical substances can be tracked from the recorded data. Figs. 4 and 5 present snapshots of the link and node information, respectively, of the full report.

Link Results at 1:00 Hrs:

Link ID	Flow GPM	Velocity fps	Unit Headloss ft/Kft	Status
1	135.65	0.28	0.05	Open
2	69.84	0.20	0.03	Open
3	34.66	0.22	0.06	Open
4	75.00	0.48	0.25	Open
5	-2.50	0.02	0.00	Open
6	-175.31	0.50	0.16	Open
7	-1364.58	3.87	7.33	Open
8	-1370.58	3.89	7.39	Open

Fig. 4. Link information from full report

Node Results at 1:00 Hrs:

Node ID	Demand GPM	Head ft	Pressure psi	Quality
J1	5.00	1000.29	173.45	0.00
J2	10.00	1000.15	260.04	0.00
J3	75.00	998.77	216.12	0.00
J4	2.50	833.99	58.06	0.00
J5	7.50	1000.46	216.85	0.00
J6	1.00	1024.65	97.34	0.00
J7	5.00	1001.11	87.14	0.00
J8	-100.00	1023.99	391.70	0.00

Fig. 5. Node information from full report

B. Simulation tools for the cyber infrastructure of WDNs

Matlab R2008b was used to represent computational aspects of the CPS, due to its powerful mathematical tools and capability of supporting a diverse range of I/O formats, which is critical to successful interfacing to simulators for the physical and communication aspects. This version of Matlab provides support for parallel computing, which is essential for simulation of the cyber layer of a WDN, as the decision support algorithms used are typically implemented in a distributed fashion.

ns-2 [17], a public-domain discrete event simulator, is the tentative choice for representing the communication network, an aspect of the cyber infrastructure that is yet to be investigated.

C. Challenges in linking simulators for the cyber and physical networks

Accurate simulation of a CPS hinges on correctly recreating the information flow of Fig.1, through the following iterative procedure:

1) Simulating the operation of the physical infrastructure.
2) Extracting the data, e.g., water pressure in various pipes, required by the decision support algorithms from the report generated in Step 1, and converting this data to an acceptable input format for the simulator for the cyber infrastructure.
3) Simulating the operation of the cyber (computing) infrastructure, including the data of Step 2 as input. This data may be supplemented by other information, e.g., historical averages. The goal of this step is generation of settings for control elements, e.g., valves, in the physical layer.
4) Converting the output of Step 3 to a format acceptable as input by the simulator for the physical infrastructure.
5) Providing the data from Step 4 as input to the simulator for the physical infrastructure.
6) Repeat Step 1.

The procedure described above is repeated iteratively for the duration of the simulation. After the initial setup, all steps are expected to take place without user intervention, as would be the case with using a single simulator. As described in Section I, differences in temporal resolution and data representation, and the lack of interoperability, especially in interfaces, pose considerable challenges in linking cyber and physical simulators in a fashion that accurately represents the CPS as a whole. Our approach to overcoming these challenges is discussed in Section IV, which describes the simulation of an intelligent WDN using Matlab and EPANET.

IV. Integrated cyber-physical simulation of intelligent WDNs

The main contribution of this paper is in developing a procedure for simulation of an intelligent WDN, such that cyber (computing) and physical aspects of the CPS are accurately and precisely represented. As described in Section III, Matlab and EPANET, respectively, are used to simulate the computing and physical infrastructures of an intelligent WDN. The procedure described in Section III-C is necessary, as it would be for a CPS from any other domain. Fig. 6 depicts this procedure for the specific case of simulation of an intelligent WDN with EPANET and Matlab. The numbers identify the corresponding step from the procedure described in Section III-C.

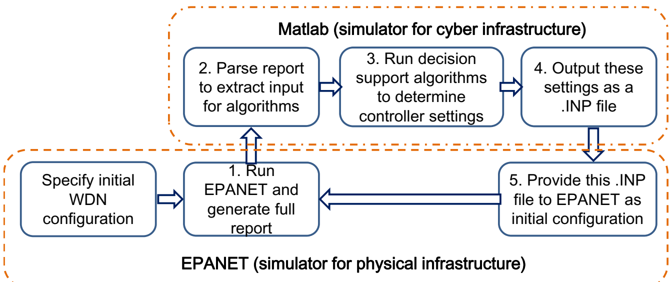

Fig. 6. Procedure for simulation of an intelligent WDN

The first step in simulating an intelligent WDN is to specify the duration to be simulated and the configuration of the physical infrastructure, e.g., topology and demand values, in EPANET. A 24-hour duration was selected for the simulation presented in this paper. After simulating the system for the specified duration, EPANET generates a full report that includes information for all links and nodes for each time step (one hour by default), as shown in Figs. 4 and 5. The full report generated as the output file of EPANET is automatically saved as a plain-text .NET file. This information includes values required as input by the decision support algorithms of the cyber infrastructure, which in turn determine settings for physical control elements such as valves.

To simulate the provision of sensor readings and other information about the physical infrastructure to the cyber control system, the full report generated as output by EPANET needs to be provided as input to Matlab. This necessitates pre-processing of the file, and parsing of the data into the matrix form required by Matlab. A script using the *textscan* and *cell2mat* commands can be defined within Matlab to carry out this pre-processing to generate a separate matrix from the EPANET data for each entity (node or link) for each simulation time step recorded in the full report, e.g., hour 1:00.

For simplicity, the simulation presented in this paper was focused on node flow. The controller (pump or valve) settings were determined by averaging the node demand within a node group, which is a subset of nodes defined in EPANET. Fig. 3 shows a number of groups. The same parsing approach can be used to extract additional data, e.g., water pressure or concentration of a given chemical, from the EPANET report, as required by more sophisticated decision support algorithms.

Each node group can reflect an associated group of consumers, such as residential nodes in the south of a city. The only requirement is that each node group include at least one controller (pump or valve), so controller settings determined by the cyber infrastructure can be utilized in water allocation. The focus of the research presented in this paper was integrated simulation of the CPS, and as such, a simplistic approach was taken to water allocation, with the goal of distributing the water as equitably as possible, subject to physical constraints on the nodes. More intelligent decision support can be achieved through game-theoretic approaches [18], which will be investigated in future stages of our research.

Matlab generates a matrix of controller settings, which need to be provided to EPANET, as they would be to the physical control elements in an actual WDN. A .INP file is required, in a format identical to the original input provided to EPANET in the first step of the simulation, with controller values updated to reflect the settings determined by the decision support algorithm. A Matlab script utilizing the *dlmwrite* and *fprintf* commands can be used to generate a .INP file with the format expected by EPANET.

```
[TITLE]
EPANET INPUT FILE

[JUNCTIONS]
; ID      Elev     Demand     Pattern
1         600      5          1
2         400      -4.5       1
3         500      25         1
4         700      2.5        1
5         500      4.3        1
6         800      1          1
7         800      -19.5      1
8         120      -100       1
```

Fig. 7. EPANET input file generated by MATLAB

In the final stage of the simulation, the .INP file generated by Matlab, which specifies settings for various control elements, is used to initiate another execution of EPANET, closing the physical-cyber-physical loop. The process can be repeated as necessary to simulate operation of the WDN over multiple cycles of cyber control. Fig. 7 shows the file resulting from execution of the water allocation algorithm for the node groups of Fig. 3. The result of executing EPANET with the .INP file generated by Matlab is shown in Fig.8. As an example of the manifestation of cyber control, the flow in the link connecting Junction1 (J1) and SOURCE, marked with an arrow, has been reduced from 75-100 GPM (yellow) in Figure 3 to 50-75 GPM (green) in Figure 8.

V. Conclusion and future work

CPSs are an emerging research area, and tools for their modeling and simulation are very limited. A number of related challenges were discussed in this paper, with focus on integrated simulation of CPSs, where the goal is to accurately reflect the operation and interaction of the cyber and physical networks that comprise the system. One of these challenges is accurate and precise representation of the features and

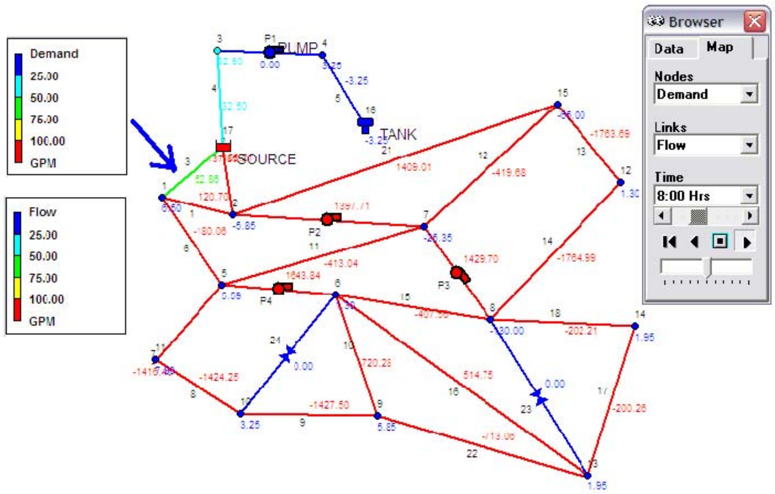

Fig. 8. Complex topology after applying cyber control

operation of the physical infrastructure, which necessitates the use of domain-specific tools such as EPANET, a simulator for WDNs. A method was described and illustrated for using Matlab and EPANET in integrated simulation of intelligent WDNs, which make use of intelligent decision support to control the quantity and quality of water. The proposed approach reflects the interdependencies between the physical and cyber infrastructures that comprise a CPS. Understanding these interdependencies is a critical precursor to any investigation of CPS, especially with respect to reliability.

The simulation technique presented in this paper is a preliminary step that will facilitate the broader goal of modeling CPS. Future extensions to this work include the simulation of more sophisticated decision support algorithms for WDN, including methods that utilize distributed computing. Insights gained from the WDN domain will be used to extend the models and techniques developed to other CPS domains, with the ultimate goal of creating CPS models that are broadly applicable, yet capable of accurately reflecting attributes specific to each physical domain.

ACKNOWLEDGMENTS

The authors would like to acknowledge the support of the Missouri S&T Intelligent Systems Center.

REFERENCES

[1] E. Lee, "Cyber physical systems: Design challenges," in *Proc. of the 11th IEEE Int'l. Symposium on Object Oriented Real-Time Distributed Computing (ISORC'08)*, May 2008, pp. 363–369.

[2] J. Sztipanovits, "Composition of cyber-physical systems," in *Proc. of the 14th Annual IEEE Int'l. Conference and Workshops on the Engineering of Computer-Based Systems (ECBS '07)*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 3–6.

[3] Department of Ecology, State of Washington, "SCADA systems for measuring and managing water," retrieved July 2009. [Online]. Available: http://www.ecy.wa.gov/programs/wr/measuring/images/pdf/scada_systems.pdf

[4] Y. Y. Haimes and P. Jiang, "Leontief-based model of risk in complex interconnected infrastructures," *Journal of Infrastructure Systems*, vol. 7, no. 1, March 2001.

[5] P. Pederson, "Critical infrastructure interdependency modeling: The survey of U.S. and international research," August 2006.

[6] S. M. Rinaldi, "Modeling and simulating critical infrastructures and their interdependencies," in *Proc. of the 37th Hawaii Int'l. Conf. on System Sciences*, 2004.

[7] N. K. Svendsen and S. D. Wolthusen, "Analysis and statistical properties of critical infrastructure interdependency multiflow models," in *Proc. of the IEEE Information Assurance and Security Workshop (IAW '07)*, June 2007, pp. 247–254.

[8] J. E. Kim and D. Mosse, "Generic framework for design, modeling and simulation of cyber physical systems," *ACM SIGBED Review*, vol. 5, no. 1, pp. 1–2, Jan. 2008.

[9] A. Al-Hammouri, V. Liberatore, H. Al-Omari, Z. Al-Qudah, M. S. Branicky, and D. Agrawal, "A co-simulation platform for actuator networks," in *Proc. of the 5th Int'l. Conference on Embedded Networked Sensor Systems (SenSys '07)*. New York, NY, USA: ACM, 2007, pp. 383–384.

[10] United States Environmental Protection Agency, "EPANET," retrieved July 2009. [Online]. Available: http://www.mathworks.com/products/parallel-computing/

[11] National Center for Supercomputing Applications, "RiverWeb," retrieved April 2009. [Online]. Available: http://destiny.mbhs.edu/riverweb/riverweb.html

[12] United States Environmental Protection Agency, "Water Quality Analysis Simulation Program," retrieved July 2009. [Online]. Available: http://www.epa.gov/athens/wwqtsc/html/wasp.html

[13] Dutch Ministry of Economics, "Waterspot," retrieved July 2009. [Online]. Available: http://www.waterspot.nl/

[14] United States Environmental Protection Agency, "Ground Water and Rainmaker Simulators," retrieved July 2009. [Online]. Available: http://www.epa.state.oh.us/ddagw/SWEET/sweet_simulators.html

[15] National Institute of Standards and Technology, "GAMS," retrieved July 2009. [Online]. Available: http://gams.nist.gov/

[16] United States Environmental Protection Agency, "EPANETmanual," retrieved July 2009. [Online]. Available: http://www.epa.gov/nrmrl/wswrd/dw/epanet/EN2manual.PDF

[17] USC Information Sciences Institute, "ns-2," retrieved Feb. 2009. [Online]. Available: http://nsnam.isi.edu/nsnam/index.php/User_Information

[18] W. Yu-Peng, X. Jian-Cang, C. Lintao, K. Ker, and T. Yew-Gan, "Game analysis in water resources optimal allocation," in *Proc. of the Int'l. Conference on Hybrid Information Technology (ICHIT '06)*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 443–446.