Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

DYNAMIC DATA DRIVEN APPLICATION SYSTEMS FOR SMART CITIES AND URBAN INFRASTRUCTURES

Richard M. Fujimoto

School of Computational Science
and Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA

Nurcin Celik
Haluk Damgacioglu

Department of Industrial
Engineering
University of Miami
Coral Gables, FL 33146, USA

Michael Hunter

School of Civil and Environmental
Engineering
Georgia Institute of Technology
Atlanta, GA 30332, USA

Dong Jin

Department of Computer Science
Illinois Institute of Technology
10 W. 31st Street
Chicago, IL 60616, USA

Young-Jun Son

Systems and Industrial Engineering
University of Arizona
1127 E. James E. Rogers Way
Tucson, AZ 85721, USA

Jie Xu

Systems Engineering and Operations Research
George Mason University
4400 University Drive, MS 4A6
Fairfax, VA 22030, USA

ABSTRACT

The smart cities vision relies on the use of information and communication technologies to efficiently manage and maximize the utility of urban infrastructures and municipal services in order to improve the quality of life of its inhabitants. Many aspects of smart cities are dynamic data driven application systems (DDDAS) where data from sensors monitoring the system are used to drive computations that in turn can dynamically adapt and improve the monitoring process as the city evolves. Several leading DDDAS researchers offer their views concerning the DDDAS paradigm applied to realizing smart cities and outline research challenges that lie ahead.

1 INTRODUCTION

Many of society's greatest challenges arise in its cities. It is widely recognized that the battle to achieve sustainable growth will be won or lost in urban regions. It is projected that by 2050 66% of the global population will reside in cities, compared to 54% today and 30% in 1950 (United Nations 2014). Global needs in infrastructure investment for land transport, telecommunications, electricity and water were estimated at $53 trillion from 2010 to 2030 (OECD 2011). In the U.S. the American Society of Civil Engineers gave the national infrastructure a grade of D+, and estimated $3.6 trillion will need to be invested by 2020 to address this problem (ASCE 2013). Resilience to natural and human-caused disasters continue to be an area of great concern; for example, the cost of disasters was estimated to be $55 billion in economic loss in the U.S. in 2011, and resulted in approximately 600 deaths (National Academies 2012). Global warming will have severe impacts on coastal cities and is another area of increasing concern.

At the same time, cities are undergoing numerous revolutionary changes due to advances in information exchange and communication technologies (ICT). Smart electric power systems are emerging that exploit ICT to improve efficiency and reliability of the energy production and distribution network. However, these systems are vulnerable to cyber-attacks raising new concerns that must be addressed. The electrification of the vehicle fleet will have important impacts on the power grid by placing new loads on the system while the introduction of photovoltaics in homes will transform some consumers of electrical energy into producers. Smart cars, autonomous vehicles and new vehicle sharing services are already beginning to revolutionize transportation. The emergence of commercial drones has the potential to revolutionize the transport and delivery of goods. At the same time manufacturing is undergoing its own revolution with new advances in data analytics, the Internet of Things, and additive manufacturing complementing increased exploitation of robotics. There are strong interdependencies among these infrastructures and changes in one infrastructure will inevitably impact others. It is clear that cities will be very different in the years ahead compared to the cities of today.

Smart cities refers to the development and integration of ICT to effectively manage the major infrastructures and municipal services making up the city in order to improve efficiency and the quality of life of its inhabitants. The smart city vision relies on extensive use of sensors to monitor the city and information technologies to adapt activities and operations of the city to meet its objectives. Realization of the smart city vision requires solution of many technical challenges.

Dynamic Data Driven Application Systems (DDDAS) is a paradigm that involves dynamically incorporating real-time data into computations in order to steer the measurement process of an application system (Darema 2004). It is immediately apparent that the DDDAS paradigm is directly applicable to realizing smart cities. For example, DDDAS provides an approach to address issues such as creating more efficient and reliable electric power grids and to mitigate traffic congestion. This paper brings together the views of several leading DDDAS researchers to discuss the applicability of this paradigm to create more resilient and sustainable cities.

The remainder of this paper is organized as follows. First Celik and Damgacioglu discuss the exploitation of DDDAS in smart power grids and many of the research challenges that must be addressed. Jin then discusses a related and important area – the protection of energy distribution networks from cyberattacks. Son discusses approaches to apply DDDAS to food, energy, and water (FEW) systems as well as its application in transportation. Hunter continues and focuses this discussion on DDDAS applications in transportation. Finally Xu discusses methodological issues in DDDAS, especially in the context of smart manufacturing.

2 SMART ENERGY FOR SMART CITIES (CELIK AND DAMGACIOGLU)

In this section, we discuss the emerging role of DDDAS-based methodologies in control and management of energy infrastructures of future smart cities, and its associated research challenges.

2.1 Smart Grid and Buildings in Smart Cities

A smart city is a holistic network of integrated urban infrastructures, where the electric power system serves as a pillar infrastructure that many other infrastructures depend on. The fundamental role of an electrical energy system in a smart and sustainable city is to improve energy efficiency and reliability of the grid, and decrease emission and energy consumption while satisfying the needs of citizens. To accomplish that, electric power infrastructure needs a technology for timely information exchange and communication (ICT). Here, the concepts of *smart grid* as a combination of the electrical power grid and ICT (Morvaj et al. 2015) and *smart building* as a way of managing and controlling renewable energy sources, house appliances, and energy consumption using a wireless communication technology (Morvaj et al. 2015) must be seamlessly integrated into the smart city energy infrastructure.

Today's energy infrastructure operates at one-way communication where demand points are passive, and utilities are the only main stakeholders that can control the whole system. With current communication technology, this system provides historical data for forecasting and decision making for utilities. However historical patterns are not enough for adapting a course of action to real-time changes in the demand in a timely manner and thus results in inefficiencies and even reliability problems in the power grid. In fact, 20% of total capacity of electricity generation is used for only 5% of time in today's infrastructure due peak demand (Lugaric et al. 2010). On the other hand, smart grid technologies promise bi-directional near- or real-time communication between the energy consumer and providers. Through this infrastructure, utilities can make use of smart houses' micro renewable energy sources, batteries and demand side management capabilities to supply energy in peak demand periods, which is so-called virtual power plant. Therefore, this controllability and real time data promise improving reliability of the power grid and energy efficiency. However, due to computational resource scarcity, complicity of making decision, etc., an intelligent architectures are needed to be utilized to enhance the value of ICT.

2.2 Yet Another Simulation Algorithm –or How is DDDAS Different?

Based on the historical data obtained from traditional energy infrastructures, researchers developed several models for efficient energy supply while satisfying energy demand and preserving and considering emission (Veit et al. 2006; Morais et al. 2010; Omu, Choudhary, and Boies 2013). In most of these studies, the challenge is to deal with the uncertainty and nonlinearities in demand and generation. In coping with these issues, *simulation* is proven as a valuable tool for assessment, comparison and optimization of the complex energy systems as a decision support and analysis platform. However, the obtained solutions from the usage of historical data in developing of these simulation models, can be far from real optimal or even infeasible due to real time changes in demand and generation. To this end, smart energy infrastructures can dramatically improve the performance of simulation model by feeding the model with near- or real-time data. However, they come at the expense of new challenges related to processing of the big data, shorter decision window and steering measurements. Here, the revolutionary concept of DDDAS investigates new algorithms and instrumentation methods for the real-time data acquisition and timely control of complex smart energy infrastructures in smart cities via incorporation of dynamic data into adaptive simulations. Such an application of the DDDAS in smart grids promises more accurate information for monitoring and controlling purposes through obtaining real-time measurements from the smart buildings (including home appliances), power supply resources, and other sensors (for weather conditions, battery, and even electric vehicle). An overview of a generic DDDAS framework for smart grids is demonstrated in Figure 1.

DDDAS performs timely monitoring, planning, and control of distributed operations at smart grids. This architecture allows us to mimic the system operations in the most accurate manner while enabling efficient usage of computational resources, access to the dynamically-changing data in the system, and superior enterprise integration in terms of communication and information synchronization. The components of system architecture and methodologies include real time data driven adaptive multi-agent simulation, smart algorithms for state estimation, fault detection and localization, model fidelity selection, demand side management decisions, pricing and operation planning.

2.3 Research Challenges and Future Directions

While DDDAS has been widely adopted in several problems such as supply chain systems (Celik et al. 2007; Celik et al. 2010), controlling and operation planning of microgrids (Thanos et al. 2013; Thanos et al. 2015; Shi, Damgacioglu and Celik 2015), controlling aerospace vehicles (Alleire et al. 2013; Alleire et al. 2013) and etc., the DDDAS efforts on smart grids is still at its infancy. Many existing issues have not been fully addressed, while new research challenges keep emerging from smart grid applications. In this section, we discuss some of the research challenges related to DDDAS approaches in smart grid operation and control.

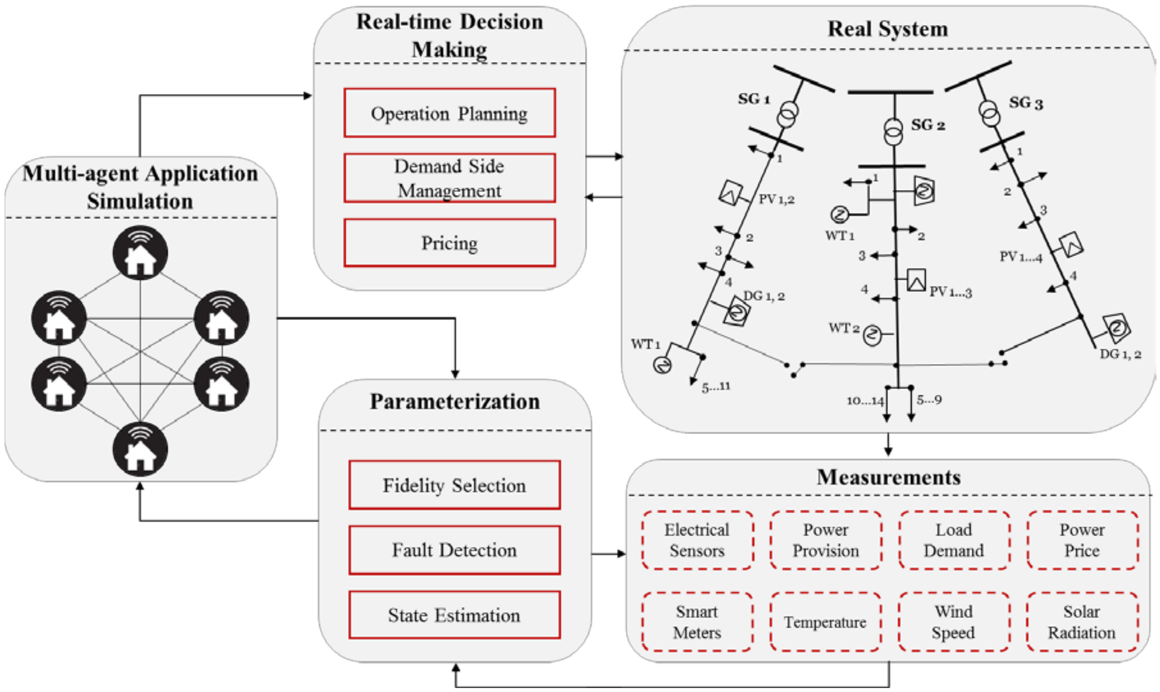

Figure 1: Overview of DDDAS application for operation planning and control of smart grids**.**

Self-configuring adaptive simulation. Since energy supply is more predictable and controllable than demand, predicting demand plays a vital role in operational planning and controlling the power grid. While aggregated demand forecasting or energy consumption of typical household appliances (Suganthi and Samuel 2012) are well-studied in the literature, forecasting of schedules of the hard hitters (e.g. dish washers, washing machines, vacuum cleaners) is not practical. Here, DDDAS addresses the issue by feeding real time data from smart meters into simulation model to adapt the model to the changes in the real system. However, the number of parameters, their associated inter-dependencies and the need for high frequency update for each parameter grows significantly at the operational level across these networks. Therefore, given the enormous amount of dynamically changing data that exists in the system, information needs to be updated wisely (and only when it is needed) in the model in order to prevent unnecessary usage of computing and networking resources. In such an architecture, the possibility communication failures and reliability of the information must also be taken into account. Therefore, adequate fault detection and localization mechanisms are needed to be developed to avoid cascading events and to improve the quality of the solution.

Multi agent modelling and multi objectivity. Power grids depend upon many interacting dynamic systems (referred to as systems of systems). Conflicting objectives of maximizing system reliability, quality of electricity flow (or minimizing number and extent of power outages), and compliance to governmental regulations as well as minimizing total cost, electricity prices, and environmental effects add further challenges to the optimal control of these systems. Through two way communication, each customer (smart building) has information about time varying prices and demand profile of the power grid and using this information, each customer may behave based on its own objective characteristics. For example, a customer may reschedule the utilization of the washing machine according to time varying prices. Therefore, in order to obtain a global optimum solution, each customer should be modeled as an agent in the simulation model.

Modular modeling. In a simulation-based planning and control framework of large-scale systems, computational efficiency is necessary not to disrupt a dynamically changing system. However, this is not a trivial task since an accurate and complete representation of a real large-scale and complex system may end up being a remarkably detailed model, especially when it is aimed to support short-term decisions (e.g.

pricing, real-time operational planning, load shifting, direct load management in smart grids). Such a detailed simulation and finding the optimal decision require significant amounts of computational power and time. To this end, intelligent modeling techniques should be developed for efficiency in these simulations. The simulation models of these complex systems also need to be robust to handle integration of various technologies at a distributed setting, especially for those allowing dynamic data updates during the model execution, and prevent from crashes and system anomalies.

3 AN ATTACK-RESILIENT FRAMEWORK FOR SECURING CRITICAL ENERGY INFRASTRUCTURE (JIN)

Critical infrastructures like power grids are key components in smart cities and urban infrastructures. Modern energy systems are increasingly adopting Internet technology in their control systems to boost control efficiency, which unfortunately opens up a new frontier in cyber-security. Cyber incidents in the energy sector have significantly grown in both number and sophistication. For example, on December 23, 2015, 225,000 customers in western Ukraine lost their electricity due to a synchronized and coordinated cyberattack (ICS-CERT 2015). These emerging scenarios demand that the problem of securing energy systems require immediate attention.

People typically explore security solutions in power systems by proceeding in one of two directions. One is to apply commercial off-the-shelf security products, such as firewall or antivirus software; the other is to exploit redundancy in power grid applications (e.g., bad data detection in the power system state estimation). However, neither direction alone can achieve effective attack-resilient solutions. Our goal is to take a DDDAS approach to develop an effective cyber-attack detection framework across both the communication network layer and the power application layer. To achieve the goal is challenging, because the power systems are (1) *mission-critical*, i.e., they must be continuously-available and the security policies and operational requirements must hold at every instant of time, and (2) *time-critical*, i.e., any new security designs must operate in real-time.

The high-level design of the cross-layer attack detection framework is depicted in Figure 2. During the execution, the system takes dynamic input data from the network layer (e.g., topologies, forwarding tables) and the application layer (e.g., control events, state estimation results), and then detects and verifies the network and application behavior against security policies specified by the policy engine. Violations will indicate vulnerabilities and errors caused by cyber-attacks or misconfigurations. Below we present the essential components of the framework in the DDDAS paradigm and discuss the corresponding challenges.

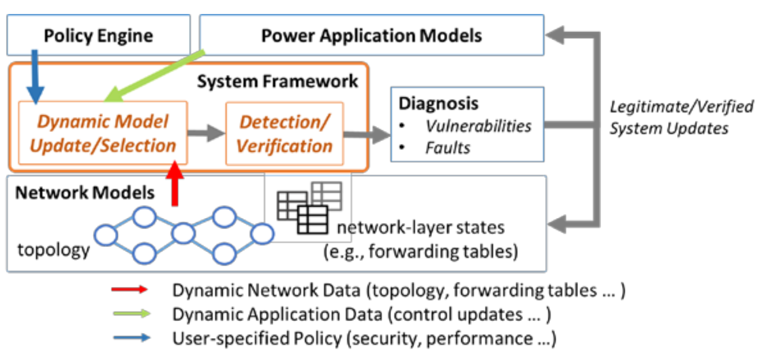

Figure 2: A Dynamic Data-Driven Cross-Layer Attack Detection Platform for Energy System Security**.**

3.1 Efficient Data-Driven Models

A key component is to design efficient models that precisely represent network-layer and application-layer behavior under dynamic system environment. We explore several models. For example, in the communication network layer, we develop an efficient graph-based networking model that uses the idea of equivalence class (i.e., a set of packets that undergo the same forwarding behavior) to significantly reduce the search space to achieve real-time verification; in the power application layer, we develop a principal component analysis (PCA)-based model for attack detection in optimal power flow (OPF). During run time, the PCA-based model analyzes real power flows and monitors different OPF outputs to determine whether the dynamic input data, including network configuration, generator capacity, and system loads, have been compromised.

One key feature is that the models are capable of accepting data at execution time (both real/synthesized data traces or real-time data) as system states evolve, and the input data drive the system to (1) select the appropriate models and the detection algorithms with different level of details, and (2) dynamically update the models, to steer the detection and verification process. For example, a system administrator plans to change the configuration of certain SELinux mandatory access control rules on a subset of end-hosts; if the primary concern is whether the desired end-to-end connectivity is preserved after the change, our system could avoid the invocation of complicated application models, and choose to perform a reachability test on a much simpler network model, with an optimized real-time verification algorithm.

3.2 Uncertainty Incorporation

It is critical to ensure security policy remains consistent as system states evolve in real time. However, the inevitable system uncertainty makes this objective very challenging. Uncertainty exists in multiple layers of the system, including uncertainty of message timing in the communication network layer, and uncertainty of measurement selection in a moving target defense (MTD)-based state estimation scheme in the power application layer. Therefore, no single point of the system is capable of controlling the exact timing and state of network/application events, ultimately leading to inconsistency between policy and actual desired behavior.

To address this issue, we incorporate uncertainty into the models and control algorithms for verifying the consistency between policy and actual network/application behavior. We develop an uncertainty-aware network model centering on the idea of "symbolic graphs", which represents the entire set of different possibilities of the network states in the manner that can be quickly analyzed (Zhou et al. 2015). We also explore ways to leverage the uncertainty-aware model to quickly perform accurate verification, including (1) algorithmic design to synthesize an update plan that (heuristically) minimizes transition latency, (2) tools to perform correct reasoning with out-of-date states, and (3) model optimization to enable real-time verification within latency bounds.

3.3 Cross-layer Detection and Interfacing

Cross-layer attack detection is useful, because carefully correlating measurements and outputs between the communication network layer and the power application layer can reveal attacks that are not seen by either layer and also enable early detection. For example, the redundant paths and traffic congestion information obtained from the communication network can be used to assist the selection of measurements and weights in state estimation attack detection, as well as the selection of elements and parameters in OPF attack detection. On the other hand, the compromised nodes and malicious application information derived from the power network can be used to ascertain the status of nodes and links for the communication network verification and to assist the selection the appropriate security policies.

The challenges lie in defining effective interfaces that enable decomposition of the system into measurement processes, administrative regions, and power application software layers that can be individually studied and verified; and in composing system-level metrics from metrics on multiple individual regions and heterogeneous types of devices and evidence. To respond to these challenges, we can leverage the recent advancement of model checking technology, such as NICE (Canini et al. 2012) with symbolic execution of event handlers, to perform a systematic joint exploration of an entire system to detect vulnerabilities and errors. This requires us to change NICE's API and internal algorithms to achieve a more robust solution and integrate its checking process in the detection framework.

4 MULTI-SCALE SIMULATION OF ENERGY AND TRANSPORTATION (SON)

Sustainable development of food, energy, and water (FEW) is vital for human well-being. An interdisciplinary team of systems and industrial engineers, plant scientist, bio-systems engineers, microbiologist, and economists at The University of Arizona has been working collaboratively under a system-based approach to develop a smart system that efficiently integrates microgrid-enabled energy, water, and food as an effective urban agriculture solution. One of the major research tasks led by the Systems and Industrial Engineering team is to design and develop a DDDAS-based, smart management system for the effective operation and management of the FEW system, while enhancing overall resilience and sustainability. The goal is to improve the use of water, energy, and labor resources, and the overall economic performance via real-time monitoring and control of a FEW system with distributed wireless sensors, data acquisition, and dynamic-simulation based decision support system.

To accomplish a DDDAS-based, smart management system for the FEW system, effective modeling of individual components is crucial. Among them, an energy system is emphasized in this section. The Son's group (a co-author) has extensive experiences in the modeling and management of photovoltaic (PV) generation, energy storage systems, and energy grids (Mazhari et al. 2011). For example, Figure 3 depicts a community-level simulator for a scenario involving distributed generators and grid, where the data was gathered from National Aeronautics and Space Administration (NASA), Energy Information Administration (EIA), National Renewable Energy Lab (NREL), Photovoltaic (PV) manufacturers, and utility companies. The demand is based on the energy consumption of individual households within the region considered. The user interface of the simulator allows the specification of various types and configurations of storage devices (battery, compressed air, super-capacitor), generators (PV, grid), and demand. Later, the Son's group has extended this dynamic data driven, simulation-based optimization approach to various problems involving integrated energy systems, such as (1) an integrated renewable energy system with a Plug-in-Hybrid Electrical Vehicle and demand response program (Zhao et al. 2013), (2) a smart energy management platform for homes (Qin et al. 2016), (3) optimal sizing and location of PV panels in a campus environment (Kucuksari et al. 2014), and (4) policy evaluation of a PV system (Zhao et al. 2011). Currently, various simulators at multi-scale (e.g. household, substation, community) are being enhanced to enable dynamic data-driven, multi-scale real-time decision support capabilities.

As another complementary effort for smart cities and urban infrastructures, a team of systems and industrial engineers and civil engineers at the University of Arizona has been developing various agent-based simulation models for transportation systems. For example, Figure 4(a) depicts a snapshot of the Phoenix road network in DynusT (Chiu et al. 2010) integrated with the Extended Belief-Desire-Intention (E-BDI) human decision-modeling framework (Lee et al. 2010). DynusT is a Simulation-Based Dynamic Traffic Assignment (SBDTA) software which performs mesoscopic simulation and dynamic traffic assignment for large-scale, regional networks (Chiu et al. 2010). The E-BDI framework was developed extending the original belief, desire and intention (BDI) concept (Bratman 1987), which includes a range of decision factors such as (1) a Deliberator, (2) a Real-time Planner, and (3) a Decision Executor in the decision-making (intention) process (Lee et al. 2010).

Figure 4(b) depicts the trajectories of diverted vehicles in the case of an accident in the considered Phoenix road network, where the concept of multi-scale modeling (a key feature enabling DDDAS) was applied. During the simulation execution, en-route planning of some agents (i.e. drivers) are performed by Time Dependent Shortest Path (TDSP) algorithm (aggregated scale), which generates the shortest route without considering inference and reasoning behaviors based on driver's preference. On the other hand, other agents perform their en-route planning via the E-BDI module (detailed scale). Extensive experiments were conducted to understand the trade-off relationships between the modeling accuracy and the computational complexity. Currently, various simulators at multi-scale for the FEW system and the transportation system are being enhanced or further developed toward a smart management system of FEW and transportation nexus in cities.

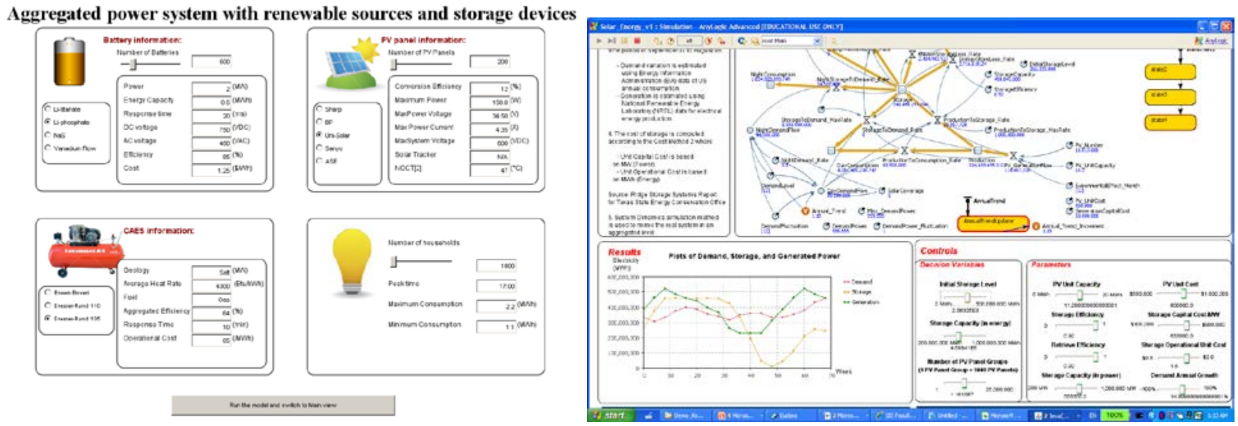

Figure 3: Simulator for community-level energy generation and grid (Mazhari et al. 2011).

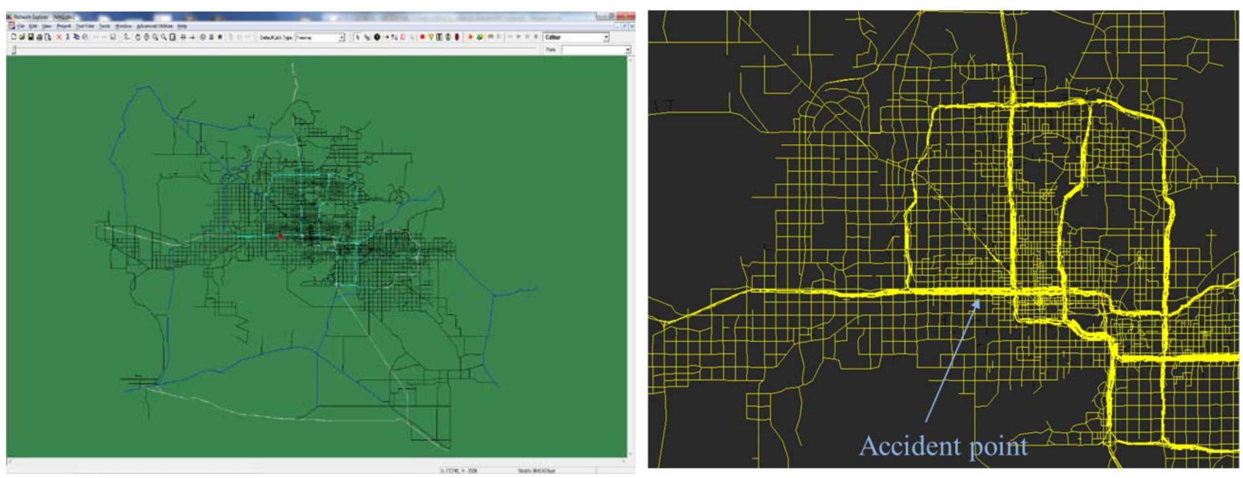

Figure 4: (a) Snapshot of a Phoenix road network in DynusT; (b) Trajectories of diverted vehicles (Kim et al. 2014).

5 TRANSPORTATION SYSTEM MANAGEMENT AND CONTROL (HUNTER)

As seen in the energy and transportation discussion, transportation systems management and control functions have employed aspects of Smart City data driven concepts for some time. Under the umbrella of Intelligent Transportation Systems (ITS) there are numerous examples of data-driven applications in transportation operations and control (Gibaud at al. 2011; Yang and Recher 2005; Fujimoto et al. 2007; Hunter et al. 2009; Skabardonis 1991; Bretherton et al. 1998). One of the earliest applications of ICT concepts was the area of intersection signal control. Since the middle of the last century, traffic control at many intersections has been driven by real-time data (Klein 2006). Embedded vehicle detection on intersection approaches have allowed implementation of control algorithms with objectives such as minimizing user travel time, minimizing emissions and energy use, or maximizing vehicle throughput. These techniques, under the moniker of coordinated actuated or adaptive signal control, have sought to optimize individual intersections, corridors, and networks. While more limited, freeway operations have seen similar efforts through the use of ramp metering technology.

However, the past several years have begun to see several developments that will significantly influence the trajectory of transportation operations and control in a smart cities environment. These developments include the inclusion of the vehicle in active data collection and vehicle control through vehicle-to-vehicle (V2V) or vehicle-to-infrastructure (V2I) technology (Vehicle Infrastructure Integration 2016); changes in the form of data available, and a broad re-imagining of mobility itself.

V2V and V2I – Historically, the detection of vehicles in the traffic system has been passive. That is, vehicles were detected externally through the use inductive loops, video, radar, etc. However, V2V and V2I, and more broadly any connected vehicle technology, is allowing the vehicle to self-identify in the system. Where previously the vehicle had to travel through a detection zone to be recognized, it may now announce itself when it enters the system or corridor. This creates significant new opportunities in scheduling arrivals at specific locations (e.g. intersections), providing user-specific trajectory guidance to minimize energy and emissions, real-time routing, etc. The implementation of new control systems incorporating data enabled by vehicle connectivity is well matched to the capabilities of the DDDAS environment. The need for simulation and analytic approaches, capable of utilizing real-time dynamic data streams, providing *faster than real-time* solutions, is critical to the utilization of these new vehicle initiated data.

Furthermore, connected vehicle technology offers the exciting opportunity for the vehicles to also be active participants in the development and implementation of control strategies. For instance, Figure 5, shows a paradigm by (Suh et al. 2014) in which ad-hoc distributed simulations were implemented onboard the vehicles. It was shown that the participation of vehicles in system traffic control strategies and system optimization could be undertaken without the need for extensive infrastructure-based detection systems or a centralized computational architecture.

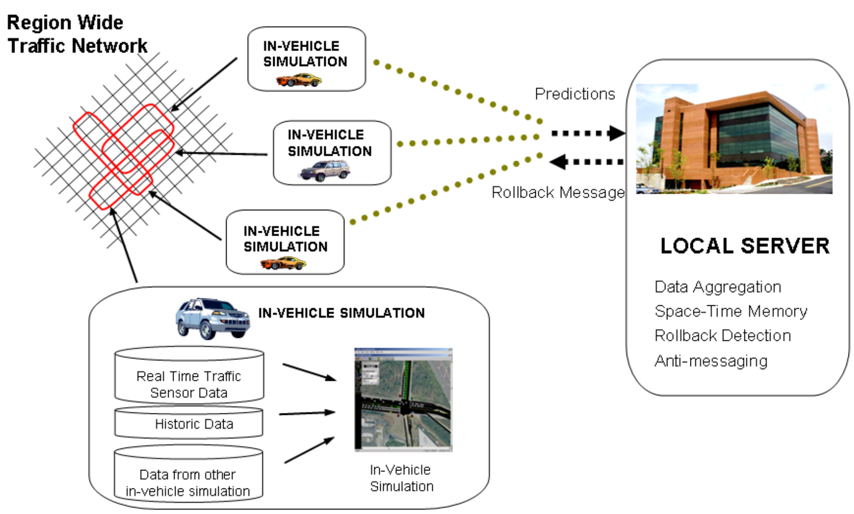

Figure 5: Ad Hoc Distributed Simulation for Transportation System Monitoring.

Data Form – On-board vehicle technology is rapidly creating a fascinating shift in the source of data for transportation operations and control. Traditionally most operational data (roadways speed, flow, densities, etc.) were the domain of the public transportation agency. Using agency operated detection devices in the roadway, performance was measured and the subsequent data streams were available for control applications. Today, vehicle's now routinely collect their individual data through the use of smart phone apps or proprietary onboard devices. While embedded infrastructure-based detection tends to be aggregate and point based (e.g. a count of all vehicle crossing a point on the roadway) vehicle based data is provided over the entire trip but is limited to the vehicle itself. Thus, data sets gathered from connected vehicles tends to be a sparse relative to the overall vehicle fleet but rich in detail. These vehicle trace data are opening new opportunities for DDDAS based management and control, particularly approaches that are able to fuse the traditional point based detection with the trace data.

Mobility – Smart cities and new technology are reimagining the very concept of mobility. Today, the vast majority of trips in many cities are in personally owned vehicles with an occupancy of one. The rapid increases in economical and easily accessed communication is creating a shift to on-demand shared mobility. Uber and similar services are only the tip of this trend. As driverless vehicles enter the vehicle fleet the concept of mobility through ownership will likely shift to mobility as a service. The impact of such a transform is still largely unknown. However, DDDAS will have a significant role in this transformation. As vehicle fleets become connected and shared, moving both people and goods, effective real-time management of vehicle services will be critical to efficient use of the vehicle fleet, optimization for trip sharing, and distribution of demands across the roadway infrastructure.

6 EFFICIENT MULTI-FIDELITY DECISION MAKING FOR DDDAS (XU)

The fundamental objective of ubiquitous sensing and control is to understand, analyze, and optimize operational conditions of systems. Although the classical feedback control theories lay a solid foundation to enable meeting operating goals and constraints based on the assessed system states, the traditional control paradigm has limited applicability to the modern dynamic data driven application systems. The fundamental challenge is the efficiency of processing and computing of data to arrive at a timely and optimal decision, which is critical due to the unprecedented dynamic interactions of multiple entities' multimodal and multi-fidelity data collection activities, and the computing of systems' operational conditions at different levels and scales. This close interaction offers tremendous potential to transform the ability to understand and control the operation of systems in the format of drastically improved real time situational awareness. But it also poses new and significant challenges on multiple fronts. A unique and critical challenge among them is the intelligent allocation of limited data acquisition and computing resources in order to make (near) real-time decisions to optimize the execution of these systems.

We consider a flexible manufacturing system as an illustration (Xu et al. 2016). There are two types of products and five work stations. Each product type has a processing sequence and needs to re-enter some work stations multiple times. Each station has multiple flexible machines. Inter-arrival and service time distributions are non-exponential. Figure 6 shows the flow of jobs through this manufacturing system. When more than one type of products are waiting for the same work station, product 1 has higher priority over product 2. The machine can perform serial batches with two products of the same type to save the setup time. The system operator needs to determine the number of machines at each work station with an objective to minimize the average cycle times. When the total number of machines in the system is 37 and the number of machines in each work station must be between 5 and 10, there are a total of 780 feasible solutions that the system operator needs to evaluate.

The optimal allocation depends on the mix of products 1 and 2. When there is a change in the product mix, the system operator needs to re-compute an optimal machine allocation in real-time. The complex characteristics of the system (e.g., re-entrance, non-exponential service times) make simulation necessary to accurately determine the cycle times for products 1 and 2. However, a full-featured discrete-event simulation model may be too time-consuming for real-time decision making with 780 feasible solutions to evaluate.

To alleviate this computational challenge, one promising paradigm is to leverage low-fidelity models that provide approximations to the cycle times for a given machine allocation decision. One such low-fidelity model can be obtained by assuming that all inter-arrival and service times are exponentially distributed, neglecting all complicating characteristics such as batching and re-entrance, and thus the average cycle times can be computed using M/M/c equations. Obviously, computing these closed-form equations are orders of magnitude faster than running discrete-event simulations. In addition, this approximation model only requires knowing the mean processing times at each work station, while the discrete-event simulation model requires full distributional information. The simplification also inevitably leads to significant bias. Nevertheless, the low-fidelity model may provide useful information in determining the relative ordering of the feasible allocation decisions. We introduce a new concept known as *ordinal transformation* (OT) to make use of low-fidelity estimates to rank all feasible decisions and transform the original decision space into a new ordinal space.

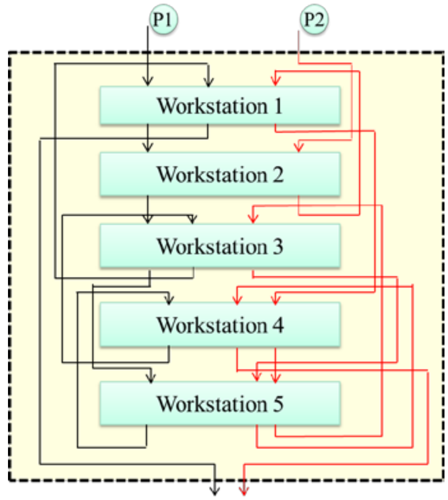

Figure 6: A flexible manufacturing system with two products.

In Figure 7(a), we plot cycle times estimated by high-fidelity discrete-event simulations. Because this is a 5-dimensional problem, we cannot display the results in the original 5-dimensional space. Instead, we indexed solutions based on their positions in the original space and then placed them on one axis using the indices. This represents one possible way to partition the original solution space. We then show in Figure 7(b) both the low- fidelity (the blue curve above) and the high-fidelity (the red curve below) simulation estimates of all 780 solutions after OT. The horizontal axis gives the rank of a solution as determined by the low-fidelity model. The left side represents solutions that are ranked to be better.

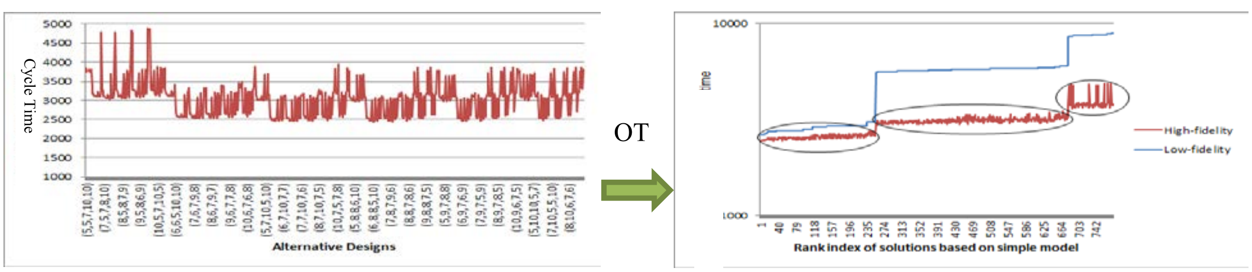

Figure 7: Cycle times in (a) the original decision space; and (b) the transformed decision space.

Despite the large bias in low-fidelity results, the relative order among solutions was highly reliable as seen by the monotonic trend in the high-fidelity model result in Figure 7(b). It is also obvious that we can partition solutions in Figure 7(b) into three groups. Solutions within the left and the middle groups have quite similar performance and thus these two groups have small group variances. While the right group shows substantial variability, it is less a problem because this group is clearly much worse than the other two groups. So one can safely sample within the left and middle groups to search for the optimal solution. In comparison, the partition in Figure 7(a) would only lead to groups with high group variances and very small group distances. Therefore, a sampling strategy would have to keep sampling from all groups.

For this particular example, the low-fidelity model results agree very well with the high-fidelity results in terms of ordinal ranking and thus the partitioning of the decisions after OT turns out to be a very good one. Based on this partitioning, one may be tempted to conclude that the middle and right groups can be thrown away and sampling should only focus on the left group. In general, we would not know *a priori* whether the partition based on the low-fidelity model is good or not. Therefore, it is important to design an *optimal sampling* (OS) strategy that focuses on more promising groups and at the same time also sample other groups to avoid being misled by the unknown bias in the low-fidelity model. In our preliminary experiments with this machine allocation problem as well as other test functions, we illustrate that this novel multi-fidelity scheme based on OT and OS leads to significant improvement in computational efficiency and thus provides a very promising real-time decision making framework for DDDAS.

7 CONCLUDING REMARKS

Smart cities offer an approach to addressing many of the most important challenges facing society today. The DDDAS paradigm is clearly applicable to realizing smart cities, and offers new technologies to address issues such as effective management of city resources such as energy distribution and cybersecurity, water management, food production and distribution, transportation, and manufacturing.

ACKNOWLEDGEMENTS

Research by the authors was supported by the following grants: NSF/AFOSR 1462503 (Fujimoto and Hunter), AFOSR FA9550-15-1-0190 (Jin), FA9550-13-1-0105 (Celik and Damgacioglu), AFOSR/DDDAS FA9550-12-1-0238 and FHWA EAR DTFH61-11-H-00015 (Son), and NSF/AFOSR ECE-1462409 (Xu).

REFERENCES

Allaire, D., J. Chambers, R. Cowlagi, D. Kordonowy, M. Lecerf, L. Mainini, and K. Willcox. 2013. "An Offline/Online DDDAS Capability for Self-aware Aerospace Vehicles." *Procedia Computer Science* 18: 1959-1968.

Allaire, D., D. Kordonowy, M. Lecerf, L. Mainini, and K. Willcox. 2014. "Multifidelity DDDAS Methods with Application to a Self-aware Aerospace Vehicle." *Procedia Computer Science* 29: 1182-1192.

ASCE 2013. *Report Card for America's Infrastructure, American Society of Civil Engineers*. www.infrastructurereportcard.org.

Bratman, M. E. 1987. "Intention, Plans, and Practical Reason" Cambridge, Massachusetts: Harvard University Press. doi:10.2307/2185304.

Bretherton, D., K. Wood, and N. Raha. 1998. *Traffic Monitoring and Congestion Management in SCOOT Urban Traffic Control System.* Transportation Research Record 1634, Transportation Research Board, National Research Council, Washington, D.C., 118-122.

Canini, M., D. Venzano, P. Perešíni, D. Kostić, and J. Rexford. 2012. "A NICE Way to Test OpenFlow Applications." *In Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 127-140.

Celik, N., S. Lee, K. Vasudevan, and Y. J. Son. 2010. "DDDAS-based Multi-fidelity Simulation Framework for Supply Chain Systems." *IIE Transactions* 42(5): 325-341.

Chiu, Y. C., L. Zhou, and H. Song. 2010. "Development and Calibration of the AnisotropicMesoscopic Simulation Model for Uninterrupted Flow Facilities." *Transportation Research Part B: Methodological* 44(1): 152-174.

Darema, F. 2004. "Dynamic Data Driven Applications Systems: A New Paradigm for Application Simulations and Measurements." *International Conference on Computational Science*.

Fujimoto, R., M. Hunter, J. Sirichoke, M. Palekar, H. Kim, and W. Suh. 2007. "Ad Hoc Distributed Simulations." *In Proceedings of the 21th International Workshop on Principles of Advanced and Distributed Simulation*, 15–24.

Gibaud, A., P. Thomin, and Y. Sallez. 2011. "FORESEE, a Fully Distributed Self-organized Approach for Improving Traffic Flows." *Simulation Modelling Practice and Theory* 19: 1096–1117.

Hunter, M., H. K. Kim, W. Suh, R. Fujimoto, J. Sirichoke, and M. Palekar. 2009. "Ad Hoc Distributed Dynamic Data-Driven Simulations of Surface Transportation Systems." *Simulation* 85: 243–255.

Industrial Control Systems Cyber Emergency Response Team (ICS-CERT). 2015. "Cyber-Attack Against Ukrainian Critical Infrastructure," *Dec-2015*. https://ics-cert.us-cert.gov/alerts/IR-ALERT-H-16-056-01, accessed April 2016.

Kim, S., Y. Son, Y. Tian, and Y. Chiu. 2014. "Drivers' En-Route Divergence Behavior Modeling Using Extended Belief-Desire-Intention (E-BDI) Framework." In *Proceedings of 2014 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Rhyzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 344–355. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Klein, L., M. Mills, and D. Gibson. 2006. *Traffic Detector Handbook: Third Edition—Volume II* Publication No. FHWA-HRT-06-139, Oct 2006, 394.

Koyuncu (Celik), N., S. Lee, K. K. Vasudevan, Y. J. Son, and P. Sarfare. 2007. "DDDAS-based Multi-fidelity Simulation for Online Preventive Maintenance Scheduling in Semiconductor Supply Chain." In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 1915-1923. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Kucuksari, S., A. Khaleghi, M. Hamidi, Y. Zhang, F. Szidarovszky, G. Bayraksan, and Y. Son. 2014. "An Integrated GIS, Optimization and Simulation Framework for Optimal PV Size and Location in Campus Area Environments." *Applied Energy* 113: 1601-1613.

Lee, S., Y. Son, and J. Jin. 2010. "Integrated Human Decision Making and Planning Model under Extended Belief-Desire-Intention Framework." *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 20(4): 23(1)-23(24).

Lugaric, L., S. Krajcar, and Z. Simic. 2010. "Smart City—Platform for Emergent Phenomena Power System Testbed Simulator." In *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, 1-7. IEEE.

Mazhari, E., J. Zhao, N. Celik, S. Lee, Y. Son, and L. Head. 2011. "Hybrid Simulation and Optimization-based Design and Operation of Integrated Photovoltaic Generation, Storage Units, and Grid." *Simulation Modelling Practice and Theory* 19: 463-481.

Morais, H., P. Kadar, P. Faria, Z. A. Vale, and H. M. Khodr. 2010. "Optimal Scheduling of a Renewable Micro-grid in an Isolated Load Area Using Mixed-Integer Linear Programming." *Renewable Energy* 35(1): 151-156.

Morvaj, B., L. Lugaric, and S. Krajcar. 2011. Demonstrating Smart Buildings and Smart Grid Features in a Smart Energy City. In *Proceedings of the 2011 3rd International Youth Conference on Energetics (IYCE)*, 1-8. IEEE.

National Academies, Committee on Increasing National Resilience to Hazards and Disasters. 2012. *Disaster Resilience: A National Imperative*. The National Academies Press.

OECD International Futures Programme. 2011. *Strategic Transport Infrastructure Needs to 2030*.

Omu, A., R. Choudhary, and A. Boies. 2013. "Distributed Energy Resource System Optimisation Using Mixed Integer Linear Programming." *Energy Policy* 61: 249-266.

Qin, X., L. Lin, S. Lysecky, J. Wang, Y. Son, and J. Sprinkle. 2016. "A Modular Framework to Enable Rapid Evaluation and Exploration of Energy Management Methods in Smart Home Platforms." *Energy Systems* 7(2): 215-235.

Shi, X., H. Damgacioglu, and N. Celik. 2015. "A Dynamic Data-driven Approach for Operation Planning of Microgrids." *Procedia Computer Science* 51: 2543-2552.

Skabardonis, A. 1991. *A Determination of Timings in Signal Systems with Traffic-Actuated Controllers.* Transportation Research Record 1554, Transportation Research Board, National Research Council, Washington, D.C.: 18-26.

Suganthi, L., and A. A. Samuel. 2012. "Energy Models for Demand Forecasting—A review." *Renewable and Sustainable Energy Reviews* 16(2): 1223-1240.

Suh, W., M. Hunter, and R. Fujimoto. 2014. "Ad Hoc Distributed Simulation for Transportation System Monitoring and Near-term Prediction." *Simulation Modelling Practice and Theory* 41: 1-14.

Thanos, A. E., M. Bastani, N. Celik, and C. H. Chen. 2015. "Dynamic Data Driven Adaptive Simulation Framework for Automated Control in Microgrids." *IEEE Transactions on Smart Grid* PP(99): 1.

Thanos, A. E., X. Shi, J. P. Sáenz, and N. Celik. 2013. "A DDDAMS Framework for Real-time Load Dispatching in Power Networks." In *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 1893-1904. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

United Nations, Department of Economic and Social Affairs, Population Division. 2014. *World Urbanization Prospects: The 2014 Revision, Highlights* (ST/ESA/SER.A/352).

Vehicle Infrastructure Integration Proof of Concept Executive Summary — Vehicle. 2009, VII Consortium. http://www.its.dot.gov/research.htm, accessed April 2016.

Veit, D. J., A. Weidlich, J. Yao, and S. S. Oren. 2006. "Simulating the Dynamics in Two-settlement Electricity Markets via an Agent-based Approach." *International Journal of Management Science and Engineering Management* 1(2): 83-97.

Xu, J., S. Zhang, E. Huang, C.-H. Chen, L.-H. Lee, and N. Celik. 2016. "MO^{2}TOS: Multi-fidelity Optimization via Ordinal Transformation and Optimal Sampling." *Asia-Pacific Journal of Operational Research* 33:1650017.

Yang, X., and W. Recker. 2005. "Simulation Studies of Information Propagation in a Self-organizing Distributed Traffic Information System." *Transportation Research Part C: Emerging Technologies* 13(5-6): 370-390.

Zhao, J., E. Mazhari, N. Celik, and Y. Son. 2011. "Hybrid Agent-based Simulation for Policy Evaluation of Solar Power Generation Systems," *Simulation Modeling Practice and Theory* 19(10): 2189-2205.

Zhao, J., S. Kucuksari, E. Mazhari, and Y. Son. 2013. "Integrated Analysis of High-Penetration PV and PHEV with Energy Storage and Demand Response." *Applied Energy* 112: 35-51.

Zhou W., D. Jin, J. Croft, M. Caesar, and P. B. Godfrey. 2015. "Enforcing Customizable Consistency Properties in Software-Defined Networks." In *Proceedings of the 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 73–85.

AUTHOR BIOGRAPHIES

RICHARD FUJIMOTO is Regents' Professor in the School of Computational Science and Engineering at the Georgia Institute of Technology. He received a Ph.D. in Computer Science and Electrical Engineering from the University of California-Berkeley in 1983. His email address is fujimoto@cc.gatech.edu.

NURCIN CELIK is an associate professor at the Department of Industrial Engineering at the University of Miami. She received her M.S. and Ph.D. degrees in Systems and Industrial Engineering from the University of Arizona. Her research interests include architectural design and application of dynamic data-driven adaptive simulations for distributed systems. She has received several awards, including AFOSR Young Investigator Research Award (2013), UM Provost Award (2011), IAMOT Outstanding Research Project Award (2011), ICCS2015 Best Paper Award, and ISERC 2015 CIS Best Paper Award. Her email address is celik@miami.edu.

HALUK DAMGACIOGLU is a Ph.D. candidate in the Department of Industrial Engineering at the University of Miami. He received a B.S. and a M.S in Industrial Engineering from the Middle East Technical University in Turkey. He has received several awards including the 2016 UM GSA Academic Excellence, Leadership and Service Award and 2015 ICCS Best Workshop Paper Award. His research

interests lie broadly in application of dynamic data driven adaptive simulations and simulation optimization under uncertainty. His email address is haluk.damgacioglu@miami.edu.

MICHAEL HUNTER is an Associate Professor at the School of Civil and Environmental Engineering at the Georgia Institute of Technology. He received his Ph.D. in Civil Engineering from The University of Texas at Austin. His email address is michael.hunter@ce.gatech.edu.

DONG JIN is an Assistant Professor in the Department of Computer Science at the Illinois Institute of Technology. He holds a Ph.D. degree in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign. His research interests lie in the areas of trustworthy cyber-physical critical infrastructures, cyber-security, simulation modeling and analysis, and software-defined networking. His email address is dong.jin@iit.edu.

YOUNG-JUN SON is a Professor and Head of Systems and Industrial Engineering at the University of Arizona and Director of Advanced Integration of Manufacturing Systems and Technologies Center. His research focuses on the coordination of a multi-scale, networked-federated simulation and decision model needed for design and control in manufacturing enterprise, renewable energy network, homeland security, agricultural supply network, and social network. He is an IIE Fellow and has received several research awards such as the SME 2004 Outstanding Young Manufacturing Engineer Award, the IIE 2005 Outstanding Young Industrial Engineer Award, the IERC Conference Best Paper Awards (2005, 2008, 2009, 2016), and Best Paper of the Year Award in 2007 from IJIE. He can be reached at son@sie.arizona.edu.

JIE XU is an Assistant Professor in the Department of Systems Engineering and Operations Research at George Mason University. He received his PhD degree in Industrial Engineering and Management Sciences from Northwestern University. His research interests include Monte Carlo simulation, simulation-based optimization, computational intelligence, and applications in risk management and aviation. His email address is jxu13@gmu.edu.