VASA: Interactive Computational Steering of Large Asynchronous Simulation Pipelines for Societal Infrastructure

Sungahn Ko, Jieqiong Zhao, Jing Xia, *Student Member, IEEE*, Shehzad Afzal, Xiaoyu Wang, *Member, IEEE*, Greg Abram, Niklas Elmqvist, *Senior Member, IEEE*, Len Kne, David Van Riper, Kelly Gaither, Shaun Kennedy, William Tolone, William Ribarsky, David S. Ebert, *Fellow, IEEE*

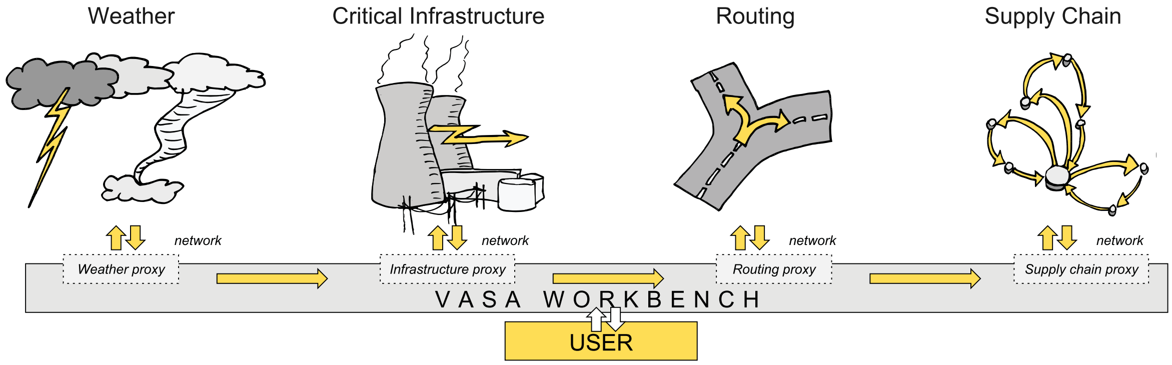

Fig. 1. Conceptual overview of the VASA system, including four simulation components for weather, critical infrastructure, road network routing, and supply chains, as well as the VASA Workbench binding them together.

Abstract—We present VASA, a visual analytics platform consisting of a desktop application, a component model, and a suite of distributed simulation components for modeling the impact of societal threats such as weather, food contamination, and traffic on critical infrastructure such as supply chains, road networks, and power grids. Each component encapsulates a high-fidelity simulation model that together form an asynchronous simulation pipeline: a system of systems of individual simulations with a common data and parameter exchange format. At the heart of VASA is the Workbench, a visual analytics application providing three distinct features: (1) low-fidelity approximations of the distributed simulation components using local simulation proxies to enable analysts to interactively configure a simulation run; (2) computational steering mechanisms to manage the execution of individual simulation components; and (3) spatiotemporal and interactive methods to explore the combined results of a simulation run. We showcase the utility of the platform using examples involving supply chains during a hurricane as well as food contamination in a fast food restaurant chain.

Index Terms—Computational steering, visual analytics, critical infrastructure, homeland security

1 Introduction

Highways, interstates, and county roads; water mains, power grids, and telecom networks; offices, restaurants, and grocery stores; sewage, landfills, and garbage disposal. All of these are critical components of the societal infrastructure that help run our world. However, the complex and potentially fragile interrelationships connecting these components also mean that this critical infrastructure is vulnerable to both natural and man-made threats: twisters, hurricanes, and flash floods; traffic, road blocks, and pile-up collisions; disease, food poisoning, and major pandemics; crime, riots, and terrorist attacks. How can a modern society protect its critical infrastructure against such a diverse range of threats? How can we design for resilience and preparedness when perturbation in one seemingly minor aspect of our infrastructure may have vast and far-reaching impacts across society as a whole?

Simulation, where a real-world process is modeled and studied over time, has long been a standard tool for analysts and policymakers to answer such questions [10]. Using complex simulations of critical infrastructure components, expert users have been able to create "what-if" scenarios, calculate the impact of a threat depending on its severity, and find optimal mitigation measures to address them. In fact, analysts have gone so far as to name simulation as the "new innovation" [33]: instead of endeavoring to produce the perfect solution once and for all, this new school of thought is to create a whole range of possible solutions and determine the optimal one using modeling and simulation. For example, during the Obama reelection campaign, it was reported that Organizing for Action data analysts ran a total of 62,000 simulations to determine voter behavior based on data from social media, political advertisements, and polling [42]. Basically, the philosophy with big data analytics driven by simulation is not to get the answer perfectly right, but to be "less wrong over time" [32]. Put differently, while it would be inappropriate to state—as others have done [2]—that big data will ever overtake theory, it is clear that large-scale simulation is a new and powerful tool for making sense of the world we inhabit.

Applying simulation to the scope of entire critical infrastructures—such as transportation, supply chains, and power grids—as well as the factors impacting them—such as weather, traffic, and man-

- *Sungahn Ko, Jieqiong Zhao, Shehzad Afzal, Niklas Elmqvist, and David S. Ebert are with Purdue University in West Lafayette, IN, USA. E-mail: {ko, zhao413, safzal, elm, ebertd}@purdue.edu.*
- *Jing Xia is with State Key Lab of CAD&CG, Zhejiang University in Hangzhou, China. E-mail: xiajing@zjucadcg.cn.*
- *Xiaoyu Wang, William Tolone, and William Ribarsky are with University of North Carolina at Charlotte in Charlotte, NC, USA. E-mail: {xiaoyu.wang, ribarsky}@uncc.edu.*
- *David Van Riper, Len Kne and Shaun Kennedy are with University of Minnesota in Minneapolis, MN, USA. E-mail: {vanriper, lenkne, kenne108}@umn.edu.*
- *Greg Abram and Kelly Gaither are with University of Texas at Austin in Austin, TX, USA. E-mail: {gda, kelly}@tacc.utexas.edu*

Manuscript received 31 Mar. 2014; accepted 1 Aug. 2014. Date of publication 11 Aug. 2014; date of current version 9 Nov. 2014. For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.
Digital Object Identifier 10.1109/TVCG.2014.2346911

made threats—requires constructing large *asynchronous simulation pipelines*, where the output of one or more simulation models becomes the input for one or more other simulations arranged in a sequence with feedback. Such a *system-of-systems* [11, 28] (SoS) will enable leveraging existing high-fidelity simulation models without having to create new ones from scratch. However, this approach is still plagued by several major challenges that all arise from the complexity of chaining together multiple simulations in this way: (C1) *monolithic simulations* that are designed to be used in isolation, (C2) *complex configurations* for each model, (C3) *non-standard data exchange* for passing data between them, (C4) *long execution times* for each individual simulation that are not amenable to interactive visual analytics, and (C5) *uncertain and inaccurate* simulations compounded by their composition.

To address these challenges, we present **VASA** (Visual Analytics for Simulation-based Action), a visual analytics platform for interactive decision making and computational steering of these types of large-scale simulation pipelines based on a visual analytics approach. The VASA Workbench application itself is an interactive desktop application that binds together a configurable pipeline of distributed simulation components. It enables the analyst to visually integrate, explore, and compare the inter-related and cascading effects of systems of systems components and potential final alternative outcomes. This is achieved by visualizing both intermediate and final results from the simulation components using a main spatiotemporal view as well as multiple secondary views. The tool provides an interface for the analyst to navigate in time, including stepping backwards and forwards, playing back an event sequence, jumping to a particular point in time, adding events and threats to the timeline, and initiating mitigation measures. Moreover, it allows them to select between or combine different ensemble outputs from one simulation to be fed to other SoS components and explore consequences. Using this interface, an analyst could for example add a weather event (e.g., either an existing hurricane from a historical database, the union of several output paths, or simulation of a new one) to a particular time, and then step forward a week to see its impact on roads, power, and food distribution.

The simulation components provide the main functionality to the VASA platform. Each simulation component communicates with the Workbench using a representational state transfer (REST) API that standardizes the data and parameter exchange. The data flows and parameters passed in the pipeline can be configured in the Workbench application using a graphical interface. Furthermore, the Workbench also includes a local *simulation proxy* for each remote simulation component that provides real-time approximations of each simulation model to enable using them for interactive visual discourse. This feature also provides the computational steering functionality of the Workbench: after configuring a simulation run in an interactive fashion, the analyst can launch the (possibly lengthy) execution from the Workbench. The Workbench then provides tools to manage the simulation pipeline, for example to prematurely shut down a simulation component to accept a partial result, skip a run, or rerun a component with new parameters.

Our work on the VASA project has been driven by stakeholders interested in supply chain management of food systems, with an initial working example of a food production to restaurant system. For this reason, other than the VASA Workbench application and the protocols and interfaces making up the platform, we have also created VASA components for simulating weather (including storms, hurricanes, and flooding), the power grid, supply chains, transportation, and food poisoning. We describe these individual components and then present an example of how the VASA platform can be used to explore a what-if scenario involving a major hurricane sweeping North Carolina and knocking out a large portion of the road networks and power grid. We also illustrate how the tool can be used to simulate food contamination outbreaks and how this information can be used to track back the contaminated products to the original distribution centers.

2 Background

Visual analytics [36], can be a powerful mechanism to harness simulation for understanding the world. Below we review the literature in visual analytics for simulation and computational steering, as well as appropriate visual representations for such spatiotemporal data.

2.1 Simulation Models

The potential for applying visual analytics to simulation involves not only efficiently presenting the results of a simulation to the analyst, but also building and validating large-scale and complex simulation models. For example, Matkovic et al. [26] show that visual analytics can reduce the number of simulation runs by enabling users to concentrate on interesting aspects of the data. Maciejewski et al. [23] apply visual analytics techniques to support exploration of spatiotemporal models with kernel density estimation and cumulative summation. This approach has also been applied to epidemic modeling and decision impact evaluation [1]. Similarly, Andrienko et al. [5] propose a comprehensive visual analytics environment that includes interactive visual interfaces for spatial modeling libraries, including selection, adjustment, and evaluation. Our work is different from this prior art in that our approach combines multiple components in a simulation pipeline, where each stage in the pipeline provides visualization for analysis.

Supply chain management is another multi-decisional context where what-if analyses are often conducted to capture provenance and processes of supplies. Simulation is recognized as a great benefit to improve supply chain management, providing analysis and evaluation of operational decisions in the supply process in advance [35]. With the IBM Supply Chain Simulator (SCS) [9] and enterprise resource planning (ERP), IBM is able to visualize and optimize nodes as well as relations in the supply chain [20]. Perez also developed a supply chain model snapshot [29] with Tableau. However, existing visualizations of supply chains are mostly limited to either local supply nodes or a metric model rather than managing the overall supply process.

2.2 Computational Steering

Computational steering refers to providing user control over running computations, such as simulations. Mulder et al. [27] classify uses of computational steering as model exploration, algorithm experimentation, and performance optimization. Applications include computational fluid dynamics (CFD) [12], program and resource steering systems [38], and high performance computing (HPC) platforms [7].

For all of the above applications, the user interface is a crucial component that interprets user manipulation for configuring data, algorithms, and parameters. Controlling, configuring, and visualizing such computational steering mechanisms is an active research area. Waser et al. proposed World Lines [39], Nodes on Ropes [40], and Visdom [31] as well as an integrated steering environment [31] to help users manage *ensemble simulations*—multiple runs of the same or related simulation models with slightly perturbed inputs—of complex scenarios such as flooding. Endert et al. [13] show how to embed analysts in the analytics loop using computational steering. In the business domain, Broeksema et al. [8] propose the Decision Exploration Lab to help users explore decisions generated from combined textual and visual analysis of decision models rooted in articifical intelligence.

2.3 Spatiotemporal Data

Spatiotemporal visual analytics systems enable users to investigate data features over time using a visual display based on geographic maps [3]. In these systems, color, position, and glyphs display features of different regions by directly overlaying the data on the map.

Many approaches to visual analytics for spatiotemporal data exist [6]; a relevant sampling follows. Andrienko and Andrienko [4] use value flow maps to visualize variations in spatiotemporal datasets by drawing silhouette graphs on the map to represent the temporal aspect of a data variable. Hadlak et al. [16] visualize attributed hierarchical structures that change over time in a geospatial context. Fuchs and Schumann [15] integrate ThemeRiver [17] and TimeWheel [37] into a map to visualize spatiotemporal data. Finally, Ho et al. [18] present a geovisual analytics framework for large spatiotemporal and multivariate flow data analysis using bidirectional flow arrows coordinated and linked with choropleth map, histograms, and parallel coordinate plots.

Some approaches enable analysis of spatially-distributed incident data, which is of particular relevance here. Maciejewski et al. [22]

propose a system for visualizing syndromic hotspots, while Malik et al. [24] develop a visualization toolkit utilizing KDE (Kernel Density Estimation) to help police better analyze the geo-coded crime data. The latter system has also been extended [25] to historic response operations and assessment of potential risks in maritime environments.

3 Design Space: Steering System-of-System Simulations for Modeling Societal Infrastructure

Computational steering is defined as user intervention in an autonomous process to change its outcome. This approach is commonly utilized in visual analytics [36] when embedding a human analyst into the computation loop for the purpose of creating synergies between the analyst and computational methods. In our work, the autonomous processes we are studying are simulation models (often based on discrete event models) that are chained together into asynchronous simulation pipelines where the output of one or several simulations becomes the input to one or several other simulations. Such a simulation pipeline is also a *system-of-systems* [11, 28] (SoS): multiple heterogeneous systems that are combined into a unified, more complex system whose sum is greater than its constituent parts. Synthesizing all these components yields the concept of visual analytics for *steering system-of-system simulations*: the use of visual interfaces to guide composite simulation pipelines for supporting sense-making and decision-making. In this work, we apply this idea to modeling societal infrastructure, such as transportation, power, computer networks, and supply chains. Below we explore the design space of this concept, including problem domains, users, tasks, and challenges. We then derive preliminary guidelines for designing methods supporting the concept.

3.1 Domain, User, and Task Analysis

The concept of creating large-scale system-of-system simulation pipelines is applicable to a wide array of problem domains. Our particular domain is for business intelligence for supply chain logistics in the fast-food business, but we see multiple other potential applications:

- **Supply chain logistics:** Impact of large-scale weather events on the distribution of goods (particularly perishables, e.g., food).
- **Public safety:** Crime, riots, and terrorist attacks on critical infrastructure, such as on roads, bridges, or the power grid.
- **Food safety:** Incidence, spread, and causes of food contamination, often due to weather (power outage) or transport delays.
- **Cybersecurity:** Societal impact of cybersecurity attacks, such as on power stations, phone switches, and data centers.

The intended audience for computational steering of simulation models using visual analytics is what we call "casual experts:" users with deep expertise in a particular application domain, such as transportation, supply chain, or homeland security, but with limited knowledge of simulation, data analysis, and statistics. Their specific background depends on the problem domain; for example, they may be logistics analysts for supply chain applications, police officers for public safety, and homeland security officials for food safety and cybersecurity. Because of this "casual" approach—a term we borrow from Pousman et al.'s work on casual information visualization [30]—our intended users are motivated by solving concrete problems in their application domain, but are not necessarily interested in configuring complex simulation models and navigating massive simulation results.

Even if our primary audience is these casual experts, it is likely that the outcome of a simulation analysis will be disseminated to managers, stakeholders, or even the general public [36]. Thus, a secondary user group for consuming our analysis products is laypersons with an even more limited knowledge in mathematics, statistics, and data graphics.

In our particular application, we identified tasks for simulation steering by working with a group of analysts from a restaurant chain that has a very large number of restaurants across the U.S, as wel as a food supply chain involving farms, food processing centers, and food distribution centers. The two main concerns voiced by these analysts are better understanding and traceability of their supply chain and understanding resiliency/vulnerability of their food supply network, especially in relation to pertain to (C1) *severe weather* and (C2) *food poisoning*: understanding the impact of natural disasters (e.g., hurricanes) on their food supply chain, processing facilities, and restaurants, as well as determining the source and distribution of food contamination cases in relation to their restaurants. More specifically, our analyst audience wants to perform the following high-level tasks in relation to these two concerns:

- Increasing *preparedness* for potential scenarios;
- Improving the *resilience* of the restaurant chain; and
- Planning for *mitigation and response* to a situation.

A motivating example for our target analysts is to understand the impact of severe weather (e.g., hurricanes) on power plants and roads, which may directly or indirectly impact food processing centers, distribution centers, and restaurants. Direct impacts include power outages, flooding, and evacuation. Indirect impacts, on the other hand, occur due to direct impacts earlier in the supply chain, such as a farms, food processing or distribution centers going offline causing shortages and redistribution of products. Both types of impacts may cause closing of facilities, which in turn may lead to indirect impacts downstream in the supply chain. Detecting such closures allows the analysts to mitigate their impact, for example by rerouting deliveries from other distribution centers, or even transporting back frozen products from a restaurant lacking refrigeration due to an extended power outage. In a hurricane scenario, the primary task then becomes determining which facilities will be closed, which routes will be impassible, and the impacts and duration these will have throughout the supply chain. Similar effects can be caused by power failures caused by other events (e.g., severe summer demand, tornadoes, power grid cyberattacks). These failures can also impact food safety (C2) due to spoilage and conditions favorable for contamination. If this is not prevented, it leads to the second task named by our target analysts: the capability to model food contamination and backtrack to its source so that the contamination can be stopped. Similar to the hurricane example above, this also requires coordinating multiple interdependent simulation models. Unfortunately, our user group does not currently have tools for performing a series of simulations to explore these scenarios.

3.2 Challenges

Modeling the real world is a tremendously difficult and error-prone process. However, we leave concerns about the fidelity, accuracy, and quality of a simulation to researchers from the simulation field. Rather, in this subsection we concern ourselves with the challenges intrinsic to connecting multiple individual simulation models into large-scale pipelines. In the context of simulation steering for such pipelines, we identify the following main challenges from our analyst audience:

C1 **Monolithic simulations:** While individual high-fidelity simulation models exist for all of the above components and threats, these models are monolithic and not designed to work together.

C2 **Complex relationships:** Each high-fidelity simulation model consists of a plethora of parameters and controls that require expertise and training, which is exacerbated when several such models are combined into a single model.

C3 **Non-standard data:** No standardized data exchange formats exist for passing the output of one simulation model, such as for weather, as input to another model, such as supply chain routing.

C4 **Long execution times:** Most state-of-the-art, high-fidelity simulation models require a non-trivial execution time, often on the order of minutes, if not hours. Such time frames are not amenable for real-time updates and interactive exploration.

C5 **Uncertainty and fidelity:** Chaining together multiple simulations into a pipeline may yield systematically increasing errors as uncertain output from one model is used as input to another. This is compounded by the fact that heterogeneous simulation models may have different levels of fidelity and accuracy.

3.3 Design Guidelines

Based on our review of the problem domain, users, and tasks above, as well as the challenges that these generate, we formulate the following tentative guidelines for designing visual analytics methods for steering system-of-system simulation pipelines:

- G1 *Simulations as standardized network services:* Distributing simulation models as network services avoids the trouble of integrating a monolithic design with another system (C1) and automatically provides a data exchange format (C3). The simulations also become decoupled, which means they can be parallelized and/or distributed in the cloud to manage long execution times (C4).
- G2 *Simulation proxies for interactive response:* Meaningful sensemaking in pursuit of one of the high-level tasks in Section 3.1 requires real-time response to all interactive queries. This means that long execution times (C4) of simulation models in the pipeline should be hidden from the user. We propose the concept of a *simulation proxy* as an approximation of a remote simulation service that is local and capable of providing real-time response at the cost of reduced (often significantly) accuracy.
- G3 *Visual and configurable relationships:* The interactive visual interfaces routinely employed in visual analytics may help to simplify and expose the complex configurations necessary for many high-fidelity simulation models (C2), even for non-expert users.
- G4 *Partial and interruptible computational steering:* Once an analyst has configured a simulation run using simulation proxies (G2) and visual mappings (G3), the full simulation pipeline must be invoked to calculate an accurate result. A full-fledged simulation run may take minutes, sometimes hours, to complete. The computational steering mechanisms provided by the software should provide methods for continually returning partial results [14] as well as interrupting a run halfway through.
- G5 *Visual representations of both intermediate and final results:* To fully leverage the power of visual analytics, we suggest using interactive visual representations of simulation results. Such visualizations should be used for both intermediate data generated by a simulation component anywhere in the pipeline—which would support partial results and interrupting a run at any time—as well as for the final results. All visual representations should be designed with uncertainty in mind (C5), and providing intermediate visualizations should also help in exposing propagation of increasing error. Finally, it may also be useful to use visual representations for the approximations created by simulation proxies (G2), but these should be clearly indicated as such.
- G6 *Spatiotemporal focus:* The primary data dimension of interest for results from simulation pipelines has both spatial and temporal attributes; for this reason, we will base the visual analytics interface on spatiotemporal visualization [3, 6]. Secondary visualizations may focus exclusively on time, space, or quantities.

4 VASA: OVERVIEW

VASA (Visual Analytics for Simulation-based Action) is a distributed component-based framework for steering system-of-system simulations for societal infrastructure. Figure 1 gives a conceptual model of the system architecture. At the center of the system is the VASA Workbench (Figure 2), a user-driven desktop tool for configuring, steering, and exploring simulation models, impacts, and courses of action. The workbench provides a visual analytics dashboard based on multiple coordinated views, an event configuration view, and a computational steering view. The workflow of the workbench revolves around initiating, controlling, analyzing, exploring, and handling events from the remote simulation components as well as the local simulation proxies.

Within the dashboard, events are displayed in a selectable calendar view (a) where each event's name, dates and a user-selected representative attribute (e.g., storm's maximum wind speed) are shown. The selected events from (a) are listed chronologically in the event viewer (b), where a user can select times for investigation. A toolbar (b-1) provides buttons for initiating simulations (e.g., cyberattack, storm simulations, distribution re-routing), selecting combinations of events (union, intersection, difference), selecting event visualization modes (polygons, contours), and triggering chronological playback.

Users can select a time within an event for comparison (right-clicking on a event's black rectangle), causing a red mark to be shown in the upper right corner of the associated event (b-2), and the corresponding impact to be highlighted in the main geospatial view (d-1, Sandy, red). This allows for comparing across different events and effects. We provide a legend window (c) for selected properties and the geographical view (d) renders the simulation results, including event evolution, routing paths, and impacts on critical infrastructures. Several of the dashboard widgets are plugged in from the simulation components. For example, a food delivery schedule to each store within a supply chain is provided in (e) where the x-axis presents corresponds to different restaurants while the y-axis represents different food processing centers or different types of foods. Here, the darker the red, the larger the quantity of the delivered food. The quantity information is provided in a tooltip that helps a user to estimate possible losses. This view enables traceback analysis (e.g., which type of food was contaminated from which processing centers, how much contaminated food was delivered to which store, etc) for food contamination incidents.

5 VASA: COMPONENTS

The VASA suite currently provides four simulation components: weather, critical infrastructure, routing, and supply chains. Data from each component and proxy is processed and merged before being visualized in the Workbench. Each proxy not only processes and stores data for its own visualization but also communicates with other proxies upon request. For example, to visualize new delivery routes, the routing proxy asks the infrastructure proxy for impacted stores before approximating new routing information. In this way, the VASA system uses a loosely coupled state that is distributed across components and proxies. We review each of the VASA components below.

5.1 Weather Component

To provide analysts with a one-stop source for weather data, we implemented a server that asynchronously amasses data from on-line sources and presents it to clients through a RESTful web interface. This provides access to weather data—both historical, current, and modeled—through a singly authenticated VASA component. The service can be queried by the user or set into a push-mode to send new events to the VASA Workbench during severe weather season (e.g., hurricane, flood, tornado season, etc).

5.1.1 Simulation Model and Simulation Proxy

Beyond historical data, the VASA weather component currently collect both ADCIRC and NOAA weather models. The ADCIRC (ADvanced CIRCulation) model is a collaboration of several research centers off the East and Gulf coasts of the United States. Active during hurricane season, these models are run every four hours when storms are presents, producing ADCIRC-formatted datasets at fixed intervals forward from the start point. These results are made publicly available using THREDDS and OPeNDAP for cataloging, discovery and data access. Similarly, NOAA produces wind-speed probabilities along the tracks of many types of storms as contours at 34, 50, and 64-knot levels. When updated datasets appear on the respective dissemination sites, we import them onto the VASA weather server, which provides a simple API to access the data in convenient multi-resolution formats.

The proxy in this component has two roles. The first role is to prepare all event datasets from the remote event server. Therefore, the system first checks for new updates from the server. If there is a new update, it retrieves the data and caches it on the local workbench for faster loading. The second role is to visualize new status of an event on the date that a user selected and notify the status change of the event to other proxies. An example status change is a user changing the start date of a hurricane in the event viewer. When this happens, the proxy visualizes a new state of the hurricane on that date and notifies this change to other components, which initiates work by downstream proxies (e.g., estimating an area without power and impassable roads).

For visualizing weather data, the user can select the visual representation either as polygons or as contours as shown in Figure 2 (b-2, the last button). In polygonal mode, two probability models (blue with two different opacities) are projected as shown in the magnification view in Figure 2. Here, the smaller polygon represents a predicted path

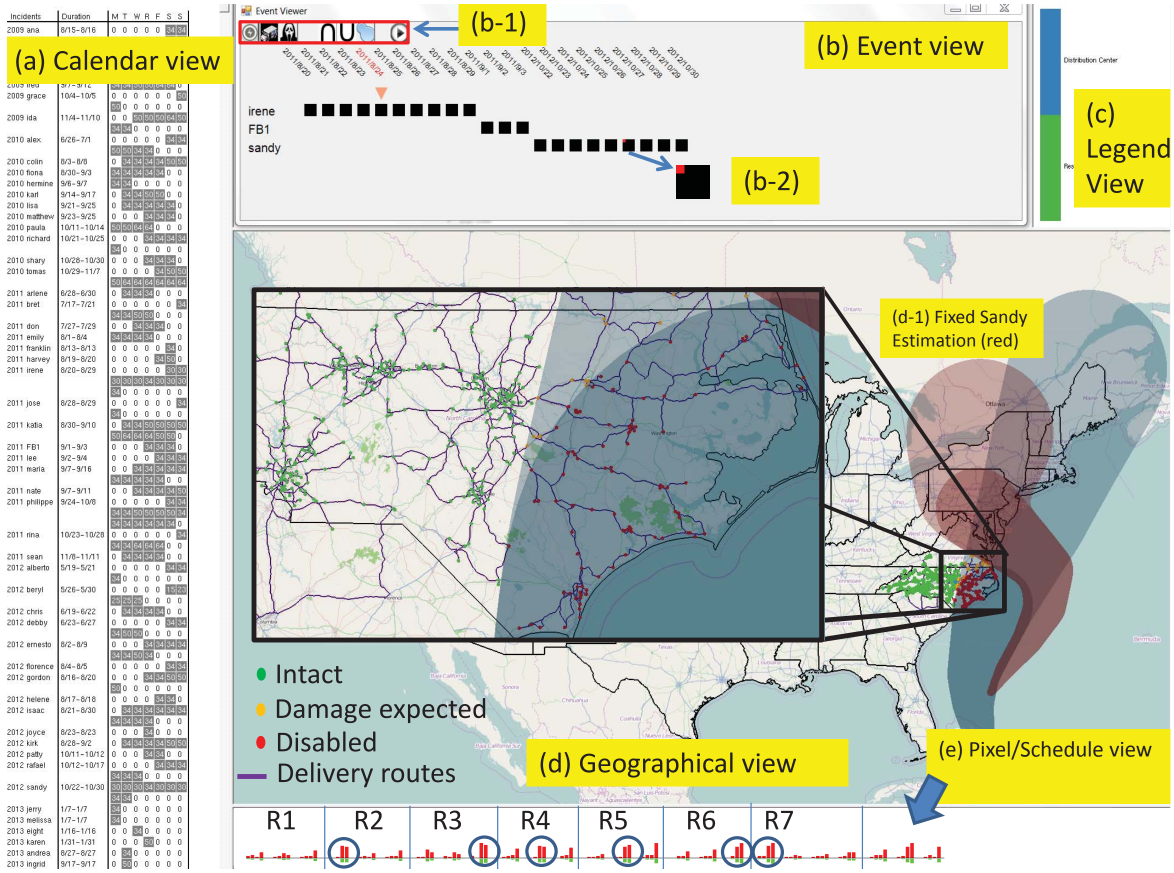

Fig. 2. Multiple coordinated views in the VASA Workbench. (a) Calendar view with available events (e.g., weather, food poisoning, cyberattack, etc). (b) Event timeline for configuring events. (b-1) Event buttons. (b-2) Fixed event. (c) Map legend. (d) Geographical map. (d-1) Fixed Sandy estimation (red). (e) Pixel/schedule view showing food deliveries. Each area divided by a blue line means a route that visits 3–4 restaurants, 3 times a week. This view also can be used for pixel-based visualization.

with high probability, and a larger one represents a predicted path with low probability. When a user selects a hurricane, the hurricane turns red for comparison to other hurricane paths. For example, in Figure 2 the paths of Hurricane Irene on August 24, 2011 (blue) and Hurricane Sandy on October 27, 2012 (red) are both rendered for comparison.

In contour mode, on the other hand, hurricanes are drawn using three different sizes of contours, each representing mean areas in different wind speeds (e.g., Hurricane Irene in our simulation model has 64 knot highest wind speed at the innermost contour, and 34 knot lowest wind speed at the outermost contour as shown in Figure 6). To utilize different wind speeds in simulation steering, a user can set up a threshold for infrastructures (e.g., a power generation unit is disabled if the wind hitting the plant has speed higher than 34 knot). In addition, a user can apply one of the contours for a time. For example, Figure 6 (top-right) presents which power generation units are affected when a contour with 34 knot hits the area. Here, red circles represent affected restaurants and purple circles represent power generation units supplying electricity to those restaurants.

5.1.2 Input and Output

The weather component often serves as a starting point for analysis by alerting severe weather conditions, and thus typically has no upstream component dependencies. Instead, simulation runs are often initiated by the analyst by adding weather events—current, modeled, or historical—to the timeline. Available weather events currently only include hurricanes, but are being expanded to other severe storm alerts, and are listed in a calendar view (Figure 2(a)).

5.1.3 Implementation Notes

The VASA weather component is implemented as a web service accessed using the common VASA RESTful API. All data objects are represented by URLs that encode procedures and parameters that, when issued, return JSON objects containing the results. This provides a very simple interface for use both by browser-based visualization UIs that use AJAX to issue requests asynchronously, and other platforms that provide access through language-specific interfaces.

5.2 Critical Infrastructure Component

Widespread emergencies such as hurricanes, flooding, or cyberattacks will often affect multiple societal infrastructures. High winds and flooding from a hurricane, for example, could knock out parts of the power grid, the effect of which would cascade to traffic signals, the communications network, the water system, and other infrastructures. The flooding might simultaneously make parts of the road network impassable. These breakdowns would affect critical facilities such as schools, hospitals, and government buildings. For longer-lived disasters, food distribution might break down due to power outage, route disruption, or other cascading effects. The purpose of VASA's critical infrastructure component is to simulate how such external emergencies, modeled in other components, will impact critical infrastructure.

5.2.1 Simulation Model and Proxy

To capture these complex, multifarious, and dynamic effects, the VASA critical infrastructure component takes into account the interrelationships between critical infrastructure systems. The simulation is built within the Vu environment [41] (Figure 3), which provides a rule-based framework for integrating multiple infrastructure submodels at a high level. This results in an interdependency ontology. Thus, for example, a breakdown of a power substation would immediately cascade to power loss at points on its distribution network. If a school were a node in the distribution network, it would be switched to backup power that, after a given time, would also shut down. Likewise, telecommunication nodes would switch to backup power that might also shut down

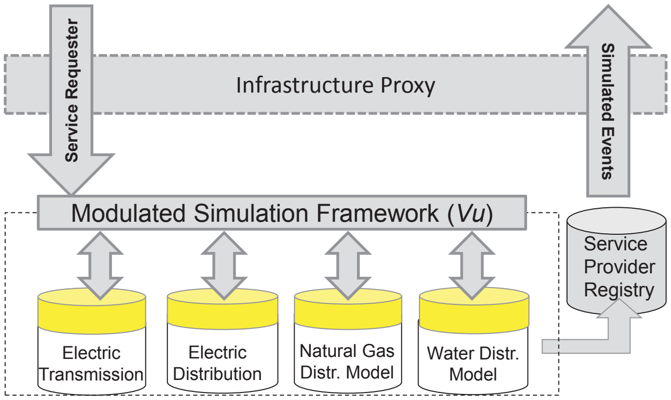

Fig. 3. Critical infrastructure server containing the pre-computed ensemble database and Vu environment with its simulation submodels.

after its prescribed duration. There could also be outages due to power load imbalances at other points in the grid.

The infrastructures we currently model include the electric grid; the communications network including TV stations, radio stations, cellular switch controls, and cell towers; transportation facilities including airports, bus terminals, rail lines and terminals, bridges, tunnels, and ports; the road network including main and secondary roads; natural gas pipelines and pumping stations; critical facilities including fire stations, police stations, schools, hospitals, emergency care facilities, manufacturing locations, government buildings, and hazmat facilities. Figure 4 shows the electric grid, which includes the complete transmission network down to substations for both North and South Carolina. Parts of the distribution network are also included, especially for critical installations. Figure 5 shows the transportation network for North Carolina, including roads, airports, rail lines, etc. For the purposes of VASA, we have also added store and distribution center locations for a large food chain in North and South Carolina. These facilities are linked to the power grid and road networks.

The proxy for the critical infrastructure component maintains a simplified connectivity network of critical infrastructure. In this graph, restaurants are connected to the nearest plant. When the proxy receives a signal of a new event (e.g., storm path change, new day for approximation), it computes which infrastructures are affected by the event. For example, when a user moves the hurricane simulation forward to a new day, our proxy checks which infrastructures are newly affected and produces an estimate and its corresponding visualization (e.g., color changes for restaurants affected by power disruptions).

5.2.2 Input and Output

The primary inputs of this component comes from the weather component represented as polygons of severe weather, such as wind speed, precipitation, and temperature data. Furthermore, the component also accepts direct manipulation of simulation parameters for particular facilities from the Workbench itself, such as the user manually shutting down a power substation. The outputs returned from the component is a list of facilities (e.g., restaurants, food processing centers and distribution centers) that are closed, and a duration of their closures.

Our prototype system currently uses data from the state of North Carolina, and the data collection and organization process involves locating and identifying components of the various infrastructures for the state. We use publicly available data sources, in some cases identifying infrastructure components by indirect means. For example, comprehensive information about the electrical grid is closely held by the utility companies. However, we have shown our results to utility company officials and received confirmation as to their high accuracy.

5.2.3 Implementation Notes

As for all the other VASA simulation components, we use a web service that can accept requests from the VASA Workbench and return simulation results ready to be presented in the user interface (see Figure 1). The critical infrastructure server itself has two components.

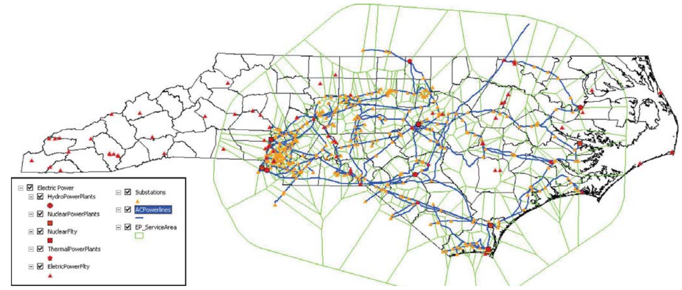

Fig. 4. Power transmission grid with parts of the distribution network.

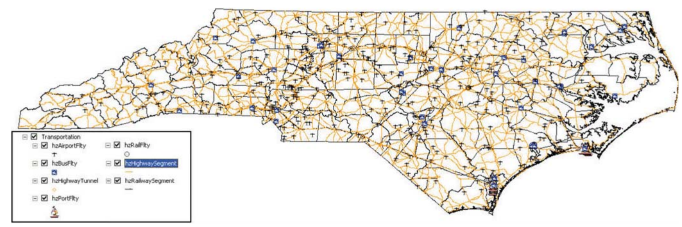

Fig. 5. Transportation network including transportation facilities.

One contains a searchable database of the pre-computed ensemble of simulation runs. The other accepts current storm path and other inputs from the weather component, converts them into courses of action, and computes a fresh set of cascading infrastructure disruption results via the Vu environment. When a request is issued via the user interface, the simulation proxy determines the weather inputs to send to the ensemble database component that immediately selects the closest ensemble simulation for use in the visual analysis. This proxy is then replaced by the more accurate result from the Vu model based on the current weather simulation as soon as it is available. Both the ensemble selection and the Vu model are depicted in Figure 3. Therefore, an emergency response manager can make initial decisions based on the proxy and then refine them once the up-to-date result is available.

5.3 Supply Chain Component

Most food systems involve a number of firms from on-farm production of inputs through processing, distribution and retail sales. For the fast food system in VASA, three different firms have collaborated to provide the data on normal system performance: a vertically integrated poultry firm (hatchery to processed chicken), a warehouse and distribution firm, and a fast food restaurant firm. Each firm contributed data from their portion of the supply chain to enable modeling of product movement from farm to restaurant. The type of data provided includes geospatial information on the facilities involved (e.g., feed mills, hatcheries, poultry farms, poultry processing facilities, distribution centers and restaurants), normal transportation routes and scheduled times from each facility to the next facility in the system and details on actual shipment quantities on average (hatchery through processing) or actual shipment records for a limited time frame (distribution centers to restaurants). As an illustration of the amount of data that drives these systems, one week of data on product delivered from the two distribution centers to the nearly five hundred individual restaurants alone constitutes more than 120,000 individual records.

Hurricanes pose significant risks for normal supply chain operation from impassable roads, power outages, and floodings disrupting facility operation and distribution of products in the system. Understanding which routes and locations are likely to be at risk from a storm would enable a firm to develop contingency plans in advance of a storm, thus reducing operational losses immediately after a storm. Given that daily sales at larger fast food restaurants can be $4,300-$7,400, losses can mount quickly. For an impending storm or immediate aftermath, rerouting could enable firms to most efficiently maintain their distribution systems for both maintaining product distribution and retrieving food from restaurants without power to minimize spoilage losses.

5.3.1 Simulation Model and Proxy

The primary objective of the supply chain component is to model distribution of product from food processing plant, through food distribution center, and to the restaurants. The routing of transports are handled in another component (Section 5.4); however, a primary concern of this component is to track product for the purpose of food safety.

Food contamination can occur both intentionally or as a malicious act at any point in the supply chain and can result in significant public health consequences, from morbidity to mortality. While firms are required to have information one step forward and one step back in their supply chain, they often have difficulty gaining visibility beyond that. By gathering data from each step in the supply chain, it is possible to trace product from farm through to restaurant and from restaurant back to farm. Using data on actual lot sizes from the firms involved, two illustrative contamination scenarios were constructed to illustrate how differently seemingly similar contamination scenarios would transpire. This system also illustrated a common problem of "hidden nodes" in the system, i.e., facilities that one firm in the system does not realize are part of its supply chain. One of the poultry slaughter and processing facilities ships raw poultry to a further processing facility that then ships the resulting product to the distribution centers. If there were a contamination at the "blind" facility, neither the distribution firm for the restaurant firm would initially know that it was part of their supply chain. A contamination scenario builder is now under development that would enable users to model a wide range of contamination events and see how they would propagate through the supply chain.

Our simulation model can generate food-borne illness data based on an approach similar to the Sydovat [21] system. There are two major components of the model for generating synthetic illness data: temporal and spatial data. A time series is contructed from its individual components (day-of-week, interannual, interseasonal, and remainder) similar to seasonal trend decomposition. To generate the time series of food-borne illnesses for a user-injected restaurant location, the user defines the mean daily count of illnesses along with seasonal and day of week components. If historical data is available, then seasonal and day of week components can be selected from this historical data. Spatial locations for temporal data are generated based on the population density distribition in that area. The analyst can also customize the grid size and density distributions.

Our simulation proxy for the supply chain component maintains a low-fidelity representation of the transport network. This is used together with the weather polygons to approximate when a distribution center and store must shut down. For food poisoning scenarios, this inherently contains spatially-distributed points of ill people simulated based on the simulation model (Section 5.3.1). To visualize the spatial distribution and the hotspots of the poisoned people, the proxy in this component uses a modified variable kernel density estimation technique with varying scales of the parameter of estimation based upon the distance from a patient location to the k_{th} nearest neighbor [34]. The model used for estimating the number of people poisoned is the same model utilized in Maciejewski et al. [1, 21], but we adjust parameters to consider different population densities in different regions.

5.3.2 Input and Output

This component accepts closures, including their durations, on supply chain facilities from the critical infrastructure component as well as severe weather polygons from the weather component. It then maintains and provides three types of information: (1) geo-information of all facilities of the supply chain, (2) delivery schedules, and (3) food products inventory in all locations (weight, size, and price).

5.3.3 Implementation Notes

The supply chain component is built in ArcGIS and Arc Network Modeler so that storm impacts can model solutions accounting for restaurants out of service (power, flooding) and impassable roadways.

5.4 Routing Component

The purpose of the routing component is to provide a mechanism for other VASA components to find appropriate routes from one facility to another given a dynamically changing world model, where roads may become impassable due to weather or other widespread emergencies.

5.4.1 Simulation Model and Proxy

The input to the routing component is a polygon representing an area impacted by severe weather (such as a hurricane). The component uses this input as a polygon barrier in the road network. Attributes of the road network are weighted to create a friction surface that iterates through routing options to determine the optimal route. The model does not currently include current traffic conditions or construction activity, but these factors could be added in the future. Each route minimizes the travel time between the distribution center and the first store or between stores. This set of routes represents the baseline scenario: how delivery trucks would travel under normal circumstances. Since delivery trucks can no longer reach outlets covered by the weather barrier, the routing service recomputes the routes with the barrier in place and returns new routes which avoid the outlets and roads covered by the barrier. If the barrier covers a distribution center, no deliveries will be made to outlets serviced by the center.

The main focus of the proxy in the routing component is on approximating the number of routes that will be replaced if a complete simulation result exists. The proxy investigates which nodes in routes are expected to be disabled when there is an event. Then, after the investigation, it builds a polygon by connecting outer-most nodes and visualizes the polygon. This gives awareness to a user that the routes in the polygon are likely to be changed after a complete simulation.

5.4.2 Input and Output

The severe weather data is ingested into the component as GeoJSON objects from the weather simulation component. Similarly, the calculated routes are output as a set of large GeoJSON objects and sent back to the caller (most often the supply chain component). One important input in this component is the impact area provided by the workbench that is presented by a polygon. Once this input is received, this component recalculate routes for the area in the polygon.

The geospatial database used by the component currently includes the addresses of two distribution centers and 505 fast-food outlets in our dataset, including the route information that links the centers to the outlets. This also includes the N shortest path routes, where N is the number of routes specified in the input data. The road network has a long list of attributes used to determine the shortest route, including road class, speed limit, number of lanes, and weight restrictions.

5.4.3 Implementation Notes

We implemented the routing component using ArcGIS Server 10.2 with the Network Analyst extension and the StreetMap Premium (TomTom North America data) road network. In general, the Esri suite of Geographical Information System (GIS) tools is widely used in a variety of industries and provides a robust set of tools and data. The server provides web-based services through REST endpoints and a robust API accessed with HTTPS GET or POST requests. The VASA workbench initiates a request to the routing service by providing a GeoJSON representation of the affected area. The affected area polygon is sent to Network Analyst Service to recalculate the route to traverse around the affected area. The response is two large GeoJSON objects containing a list of outlets no longer reachable, incremental travel time between stops, and the new route. Currently, the route processing requires 2-3 minutes to complete; this can be significantly improved in the future by commissioning a dedicated production server.

6 EXAMPLES

Here we demonstrate how the VASA system provides situational awareness using two examples: the impact of weather on macro-scale supply chains, and foodborne illness contamination and spread.

6.1 Supply Chains During Hurricane Season

Our first example is the potential impact of hurricanes on North Carolina's critical infrastructure, especially our food distribution network,

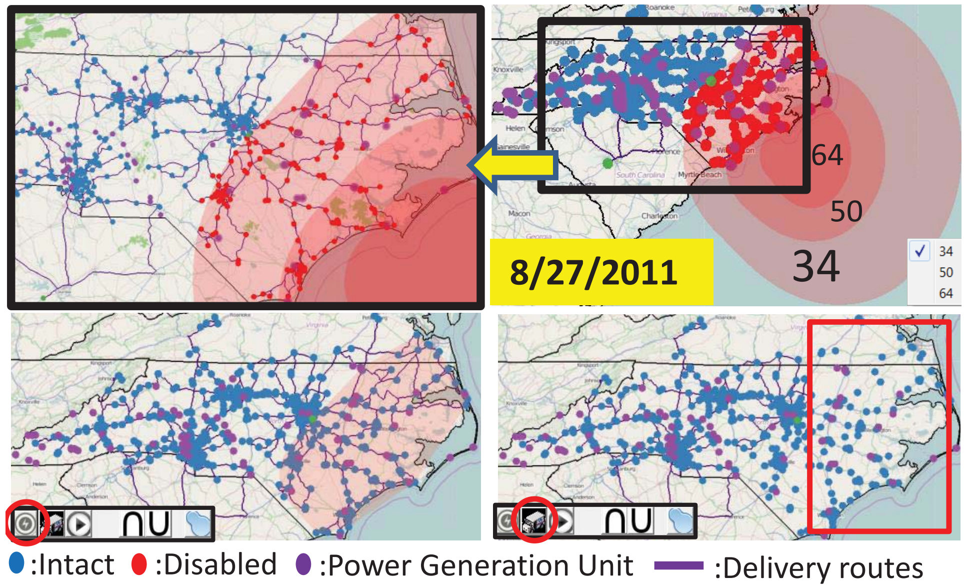

Fig. 6. In this simulation, power generation units were hit by up to 34 knot during Hurricane Irene on August 27, 2011. Our hurricane proxy instantly estimates the impacted restaurants (right-top, left-top). Note that one distribution center (green) is outside the hurricane. After a complete power-grid simulation run is finished (by clicking the circled lightning button), a polygon representing the power outage area is shown. Next, this polygon is sent for use in computing new food delivery paths. Note that food is not delivered to the power outage area (right-bottom, red box).

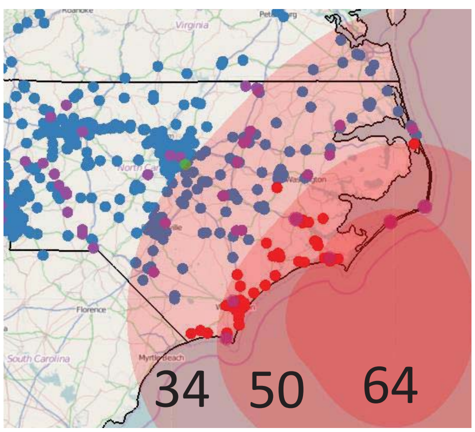

Fig. 7. If the power substations could have resisted up to 50 knot winds, the number of impacted restaurants would have been much smaller.

in North Carolina (NC). Our exploration begins by selecting appropriate historial hurricanes for examination using the calendar view as shown in Figure 2, where each hurricane name, duration, and selected summary attribute (e.g., maximum hurricane wind speed) are provided. While we investigate the paths of these historical hurricanes, we see that Irene in 2011 and Sandy in 2012 passed over NC. Because Sandy passed over only a small area in upper NC (Figure 2 (d), red polygon), we choose to focus on Irene for further investigation.

One interesting date is August 27, 2011 when Irene passed directly over eastern North Carolina, an area with many power generation facilities as shown in Figure 6 (top-right, purple circles). After we set up the wind tolerance value for these facilities to be 34 knots, our hurricane proxy instantly estimates which restaurants will be impacted based on the relationships between the units and the restaurants, coloring the impacted restaurants red. Here, we also initiated a complete simulation for power outages and transportation network damage. Next, a polygon is shown representing an area where restaurants are disabled and roads are blocked (bottom-left in Figure 6). To efficiently manage distribution, this impact requires the food provider to change its delivery schedule, and this new routing is computed based on the impacted restaurant polygon and road conditions (e.g., blocked by flooding). After a simulation to compute the new routes (by clicking the truck button in a red circle, right-bottom Figure 6), we see that the updated delivery paths do not include the affected restaurants. The economic loss caused by this event is estimated based on the model in Section 5.3 as being up to \$1.13 million. Another possible what-if question is "How different would the result be if the power generation units could resist winds up to 50 knots?" Figure 7 shows the first step of the analysis where we see many fewer restaurants affected compared to Figure 6 top-right (units are resilient to 34 knots). In this case, the estimated losses are less than \$333,000.

6.2 Fast Food Contamination

Food poisoning is an illness caused by eating contaminated food containing viruses, bacteria, and germ-generated toxins. There are many possible causes of food contamination including storage at inappropriate temperatures [19], improper food handling, and cross-contamination during processing or packaging. As unfortunately experienced several times per year, tracing back the cause of the contamination is a very difficult and lengthy process. In this example, we explore a hypothetical scenario demonstrating how VASA can be used to trace-back the root causes of an incident of foodborne illness.

To create the distribution of the ill population, we simulate the distribution of contaminated food to stores, then simulate the illnesses in the neighboring areas using the simulation model discussed in Section 5.3. This creates the common base scenario of reports of people who are ill, their date of illness, and their location to create the food contamination scenario for the trace-back investigation.

For example purposes, we simulated these illnesses occurring during a three day span (September 1, 2011 to September 3, 2011) as shown in Figure 8. Since this is almost one week after Hurricane Irene, one may assume that power outages during the storm could be the possible reason behind the contamination. To confirm this hypothesis, we looked at the hot spots in Figure 8 and identified the stores closest to these hot spots. On cross comparison, we can identify the common products/lots in those stores, their distribution center, as well as their delivery mechanisms. As shown in Figure 8 bottom matrices, the rows represent 3 food processing centers and 4 types of food, and there is a column for each restaurant. Each cell is colored such that the darker the red color, the higher the amount of each product provided. Here, the restaurants in the affected area that are selected in the box in the top-left are highlighted with light green boxes. For stores S9 and S12, only one food processing center provided products,

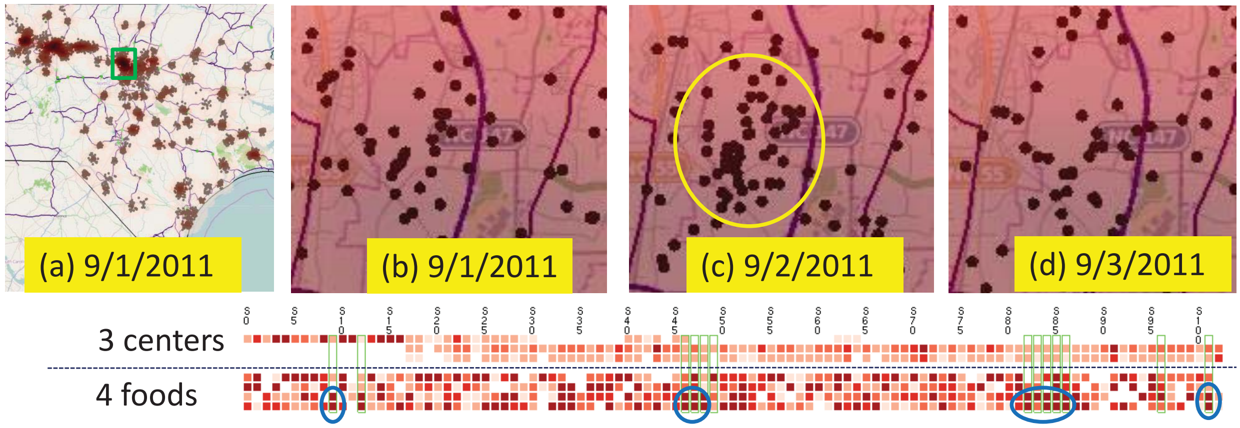

Fig. 8. Ill people caused by contaminated food is presented using a KDE hotspot visualization. In (a), the darker location has a larger number of poisoned people. Brown points mean ill people in the reported location. The locations highlighted by a green box in (a) is magnified in (b), (c) and (d) on different dates. As the timeline shows, the number of ill people increased until 9/2/2011, then started decreasing on 9/3/2011. The bottom matrices show which food processing centers (1–3) were involved and which foods (1–4) were delivered to which store in 8/30/2011, two days before the illness. Here, the restaurants in the light green boxes are the those selected by the thicker green box in (a). We see that a large quantity (darkest red pixels in blue circles) of two foods (third and fourth rows) are commonly provided to restaurants in the area.

while other processing centers supplied most of the food throughout the network. Upon further inspection, one can determining that product lots in row 3 and 4 are common in most of the restaurants yielding ill individuals. Some example routes are shown in Figure 2(e), where each route supplies 3-4 restaurants. A red bar represents the supplied food and the green bar represents the food consumed at a restaurant. Here, we see that a large amount of the third and fourth foods (blue circles in Figure 2(e)) are delivered and will all be consumed within a few days. Therefore, these two product lots are good candidates for further inspection in tracing back the contaminated food item.

7 Discussion

We have received some initial feedback from various user groups as to the value of the VASA system. Our food supply chain experts helped develop the pipeline and tailor it for their workflow. We have also had very positive feedback from numerous regional and federal government officials on the value of the VASA workbench for use in command centers at the local and regional level for increased situational awareness and the ability for plan for both resiliency and response before and during an event. Feedback from regional Federal Emergency Management Agency (FEMA) personnel is that this system is novel in that it could enable unprecedented work within their organization: visual investigation on large multiple simulation runs and instance approximations under severe weather conditions. They noted that the system enables “The Whole Community” approach to meet the actual needs of residents, emergency managers, organizational and community leaders, government officials, and the general public when extreme weather impacts various societal infrastructures. They felt that the VASA tool would enable each community to make informed and timely decisions about how to manage throughout an extreme weather event. They also suggested extending our system to real-time weather data to respond to all warnings and alerts from the National Weather Service. We have also received similar positive feedback from non-governmental aid organizations.

While the VASA system is full-featured, it may be overkill for simple analyses that only require using a few simulation components. Furthermore, sometimes which simulations to use is not clear a priori, and analysts may have to explore the problem in-depth before they can make a decision. This is also one of the strengths of the VASA system: the VASA Workbench does not stipulate a specific simulation pipeline, but leaves this choice to the analyst. It also provides proxies to estimate simulations prior to a run, and visual and interactive representations of intermediate results. However, it is also true that for a limited simulation involving only a single simulation, using the entire VASA system may be excessive and introduce a lot of overhead.

A more general question is how the VASA approach to interactive computational steering will impact the overall analysis process. Since we have yet to conduct formal user studies with our target audience for the VASA project, it is too early to conclusively answer this question. However, our intuition is that the core benefit of VASA is to introduce interactive visual analytics to a domain that is fundamentally asynchronous and off-line. We speculate that this, in turn, will yield the same kind of rapid, iterative exploration of simulation scenarios that Fisher et al. [14] observed when introducing visualization of partial results to large-scale database computations. We think that this will contribute to analysts wasting less time on configuring their simulation runs and will yield more informed and well-designed results.

8 Conclusion and Future Work

We have introduced the notion of visual analytics for simulation steering within the context of societal infrastructure. To our knowledge, ours is the first to study visual analytics for simulation from a *systems-of-systems* [11] perspective, where multiple heterogeneous—often physically distributed—systems are combined into a unified, more complex system in which the linkages between components provide a sum greater than its constituent parts. This notion transcends individual simulation models and instead chains together multiple high-fidelity simulations into large-scale asynchronous pipelines. The VASA system we presented as a practical example of such an approach is a distributed application framework consisting of a central Workbench controlled by an analyst and a set of loosely coupled simulation components implemented as distributed network services.

Big data simulation is a powerful new tool for data science, and while our work on applying visual analytics to this domain is conceptually complete, it really only scratches the surface of what is possible. Future work on the VASA system will involve integrating even more advanced and detailed simulation components, such as high-fidelity power grid models, gas pipelines, and power plants for energy infrastructure; bridges, tunnels, and causeways for transportation networks; and hospitals, police stations, and fire stations for societal infrastructure. In doing so, we envision designing additional novel visual representations and interactions for configuring these components as well as visualizing their proxy, intermediate, and final results.

Acknowledgments

This material is based upon work supported by the U.S. Department of Homeland Security under Grant Award Number 2009-ST-061-CI0001-06. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

REFERENCES

[1] S. Afzal, R. Maciejewski, and D. S. Ebert. Visual analytics decision support environment for epidemic modeling and response evaluation. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*, pages 191–200, 2011.

[2] C. Anderson. The end of theory: The data deluge makes the scientific method obsolete. *Wired Magazine*, June 2008.

[3] N. Andrienko and G. Andrienko. *Exploratory analysis of spatial and temporal data: A systematic approach*. Springer Verlag, 2006.

[4] N. V. Andrienko and G. L. Andrienko. Interactive visual tools to explore spatio-temporal variation. In *Proceedings of the ACM Conference on Advanced Visual Interfaces*, pages 417–420, 2004.

[5] N. V. Andrienko and G. L. Andrienko. A visual analytics framework for spatio-temporal analysis and modelling. *Data Mining and Knowledge Discovery*, 27(1):55–83, 2013.

[6] L. Anselin. From SpaceStat to CyberGIS: Twenty years of spatial data analysis software. *International Regional Science Review*, 35(2):131–157, 2012.

[7] J. Biddiscombe, J. Soumagne, G. Oger, D. Guibert, and J.-G. Piccinali. Parallel computational steering and analysis for HPC applications using a ParaView interface and the HDF5 DSM virtual file driver. In *Proceedings of the Eurographics Conference on Parallel Graphics and Visualization*, pages 91–100, 2011.

[8] B. Broeksema, T. Baudel, A. G. Telea, and P. Crisafulli. Decision exploration lab: A visual analytics solution for decision management. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):1972–1981, 2013.

[9] S. Buckley and C. An. Supply chain simulation. In *Supply Chain Management on Demand*, pages 17–35. Springer, 2005.

[10] L. Costa, O. Oliveira, G. Travieso, F. Rodrigues, P. Boas, L. Antiqueira, M. Viana, and L. Rocha. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Advances in Physics*, 3(60):319–412, 2011.

[11] D. DeLaurentis and R. K. Callaway. A system-of-systems perspective for public policy decisions. *Review of Policy Research*, 21(6):829–837, 2004.

[12] H. Doleisch, H. Hauser, M. Gasser, and R. Kosara. Interactive focus+context analysis of large, time-dependent flow simulation data. *Simulation*, 82(12):851–865, 2006.

[13] A. Endert, P. Fiaux, and C. North. Semantic interaction for sensemaking: inferring analytical reasoning for model steering. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2879–2888, 2012.

[14] D. Fisher, I. O. Popov, S. M. Drucker, and M. C. Schraefel. Trust me, I'm partially right: incremental visualization lets analysts explore large datasets faster. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 1673–1682, 2012.

[15] G. Fuchs and H. Schumann. Visualizing abstract data on maps. In *Proceedings of the International Conference on Information Visualization*, pages 139–144, 2004.

[16] S. Hadlak, C. Tominski, H.-J. Schulz, and H. Schumann. Visualization of attributed hierarchical structures in a spatiotemporal context. *International Journal of Geographical Information Science*, 24(10):1497–1513, 2010.

[17] S. Havre, B. Hetzler, and L. Nowell. ThemeRiver: visualizing theme changes over time. In *Proceedings of the IEEE Symposium on Information Visualization*, pages 115–124, 2000.

[18] Q. Ho, P. H. Nguyen, T. Åström, and M. Jern. Implementation of a flow map demonstrator for analyzing commuting and migration flow statistics data. *Procedia - Social and Behavioral Sciences*, 21:157–166, 2011.

[19] B. C. Hobbs. *Food poisoning and food hygiene*. Edward Arnold and Co., London, United Kingdom, 1953.

[20] A. Kamran and S. U. Haq. Visualizations and analytics for supply chains. Technical report, IBM, February 2013.

[21] R. Maciejewski, R. Hafen, S. Rudolph, G. Tebbetts, W. Cleveland, S. Grannis, and D. Ebert. Generating synthetic syndromic-surveillance data for evaluating visual-analytics techniques. *IEEE Computer Graphics & Applications*, 29(3):18–28, May 2009.

[22] R. Maciejewski, R. Hafen, S. Rudolph, G. Tebbetts, W. S. Cleveland, S. J. Grannis, and D. S. Ebert. Generating synthetic syndromic-surveillance data for evaluating visual-analytics techniques. *IEEE Computer Graphics and Applications*, 29(3):18–28, 2009.

[23] R. Maciejewski, S. Rudolph, R. Hafen, A. M. Abusalah, M. Yakout, M. Ouzzani, W. S. Cleveland, S. J. Grannis, and D. S. Ebert. A visual analytics approach to understanding spatiotemporal hotspots. *IEEE Transactions on Visualization and Computer Graphics*, 16(2):205–220, 2010.

[24] A. Malik, R. Maciejewski, T. F. Collins, and D. S. Ebert. Visual analytics law enforcement toolkit. In *Proceedings of the IEEE Conference on Technologies for Homeland Security*, pages 222–228, 2010.

[25] A. Malik, R. Maciejewski, B. Maule, and D. S. Ebert. A visual analytics process for maritime resource allocation and risk assessment. In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*, pages 221–230, 2011.

[26] K. Matkovic, D. Gracanin, M. Jelovic, A. Ammer, A. Lez, and H. Hauser. Interactive visual analysis of multiple simulation runs using the simulation model view: Understanding and tuning of an electronic unit injector. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1449–1457, 2010.

[27] J. D. Mulder, J. J. van Wijk, and R. van Liere. A survey of computational steering environments. *Future Generation Computer Systems*, 15(1):119–129, 1999.

[28] C. Ncube. On the engineering of systems of systems: key challenges for the requirements engineering community. In *Proceedings of the IEEE Workshop on Requirements Engineering for Systems, Services and Systems-of-Systems*, pages 70–73, 2011.

[29] R. Perez. Supply chain model, April 2011.

[30] Z. Pousman, J. T. Stasko, and M. Mateas. Casual information visualization: Depictions of data in everyday life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1145–1152, 2007.

[31] H. Ribicic, J. Waser, R. Fuchs, G. Bloschl, and E. Gröller. Visual analysis and steering of flooding simulations. *IEEE Transactions on Visualization and Computer Graphics*, 19(6):1062–1075, 2013.

[32] G. Satell. Why our numbers are always wrong. *Digital Tonto*, October 2012.

[33] G. Satell. Why the future of innovation is simulation. *Forbes*, July 2013.

[34] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 1986.

[35] S. Terzi and S. Cavalieri. Simulation in the supply chain context: a survey. *Computers in industry*, 53(1):3–16, 2004.

[36] J. J. Thomas and K. A. Cook. *Illuminating the Path: The R&D Agenda for Visual Analytics*. IEEE Press, 2005.

[37] C. Tominski, J. Abello, and H. Schumann. Axes-based visualizations with radial layouts. In *Proceedings of the ACM Symposium on Applied Computing*, pages 1242–1247, 2004.

[38] J. S. Vetter and K. Schwan. Progress: A toolkit for interactive program steering. Technical Report GIT-CC-95-16, Georgia Institute of Technology, 1995.

[39] J. Waser, R. Fuchs, H. Ribicic, B. Schindler, G. Bloschl, and E. Gröller. World lines. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1458–1467, 2010.

[40] J. Waser, H. Ribicic, R. Fuchs, C. Hirsch, B. Schindler, G. Bloschl, and E. Gröller. Nodes on ropes: A comprehensive data and control flow for steering ensemble simulations. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1872–1881, 2011.

[41] X. W. William Tolone and W. Ribarsky. Making sense of the operational environment through interactive, exploratory visual analysis. In *Proceedings of the NATO/OTAN Symposium on Visual Analytics*, pages 9–1–9–13, 2014.

[42] A. Woodie. Big data analytics give electoral edge. *Datanami*, June 2013.