Visual Analytics Decision Support Environment for Epidemic Modeling and Response Evaluation

Shehzad Afzal* Ross Maciejewski† David S. Ebert‡

Purdue University Visualization and Analytics Center

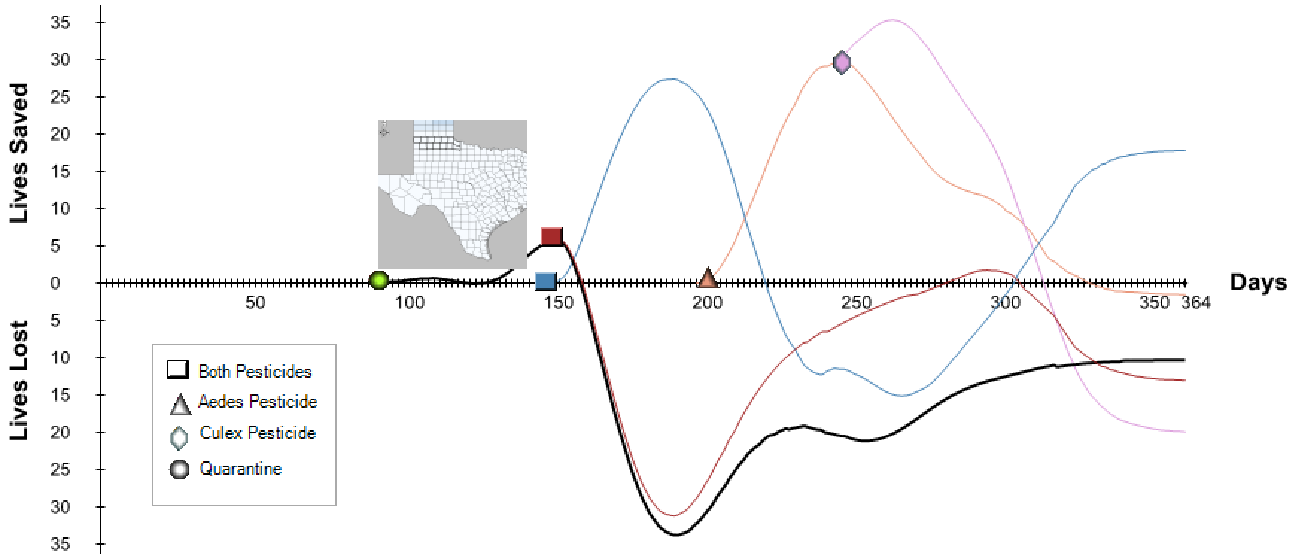

Figure 1: The decision history tree view. As users interact in the model view, the decisions made generate a history tree. Paths of the tree are plotted over time on the x-axis, with the y-axis representing the cumulative deviation from the baseline simulation. Mousing over on a node brings up a thumbnail view of the decision measures implemented at that point in the simulation. Legend symbols represent mitigative response measure types. Each symbol in the decision history tree represents the insertion point of the decision path. A unique color is assigned to each symbol and the corresponding decision path.

Abstract

In modeling infectious diseases, scientists are studying the mechanisms by which diseases spread, predicting the future course of the outbreak, and evaluating strategies applied to control an epidemic. While recent work has focused on accurately modeling disease spread, less work has been performed in developing interactive decision support tools for analyzing the future course of the outbreak and evaluating potential disease mitigation strategies. The absence of such tools makes it difficult for researchers, analysts and public health officials to evaluate response measures within outbreak scenarios. As such, our research focuses on the development of an interactive decision support environment in which users can explore epidemic models and their impact. This environment provides a spatiotemporal view where users can interactively utilize mitigative response measures and observe the impact of their decision over time. Our system also provides users with a linked decision history visualization and navigation tool that support the simultaneous comparison of mortality and infection rates corresponding to different response measures at different points in time.

*e-mail: safzal@purdue.edu
†e-mail:rmacieje@purdue.edu
‡e-mail:ebertd@purdue.edu

IEEE Symposium on Visual Analytics Science and Technology
October 23 - 28, Providence, Rhode Island, USA
978-1-4673-0013-1/11/$26.00 ©2011 IEEE

1 Introduction

Federal, state, and local community public health officials must prepare and exercise complex plans to deal with a variety of potential mass casualty events [1, 7, 13]. The planning stages often utilize knowledge gained during tabletop exercises [7], or summary details based on information and trends provided via very complex modeling. Moreover, such plans are developed with only a few specific scenarios or pre-event concepts in mind and often ignore the fact that the solutions dealing with a disease outbreak are very dependent on its underlying traits and actual characteristics, which may not be known *a priori*.

In order to better prepare and plan for events, analysts and decision makers have begun incorporating computer based simulations to model potential disease. These models employ a variety of parameters and output complex multivariate data which needs to be analyzed and explored. Furthermore, as parameters are modified, new outcomes occur, requiring further analysis to compare various scenarios. In regards to infectious disease simulations, the results need to be compared across space and time to evaluate decision measures as they are implemented over a variety of state spaces. Analysts need to work in an environment where they can explore the impact mitigative measures (i.e., school closures during a pandemic, spraying pesticides for insect borne illnesses). These decision measures are not used to determine the best solution to the model; instead, these decision measures are placed at different points to help analysts understand and illustrate the effects that certain responses will have. In this way, decision makers can better understand the effects of delaying responses, lack of supplies for implementing a response and the potential impact of an outbreak under varying conditions. Figure 1 illustrates such a decision tree analysis structure.

In this paper, we present a suite of predictive visual analytic tools that not only provides insight into the effectiveness of a decision, but also provides an interactive visual analytics environment for the investigation of multiple courses of responses as well as comparisons of the effectiveness of each component of a response plan. These tools can be utilized during training exercises to help decision makers and first responders better understand the impact of their decisions, as well as during crisis management where model parameters can be adjusted to model the current spread and predict potential future outcomes.

In order to facilitate enhanced model exploration and decision analysis, we have developed a linked spatiotemporal visual analytics tool (Figure 2) designed for advanced model simulation and exploration for epidemiologists, local public health officials and other healthcare officials. The system was designed from its inception in collaboration with health experts, state healthcare officials and epidemiologists to address their needs. Our system features include the following:

- Flexible decision history support trees that can link to multiple simulation runs (Figure 1);
- Interactive controls for exploring decision measures and decision points within the simulation;
- Simulation replay and path exploration visualizations for decision analysis.

As part of the simulation, users may interactively deploy various resources as a means of lessening the disease impact or preventing further spread. These decision points include both spatial and temporal locations, creating a large and complex decision space.

In order to understand this decision space, our work focuses on advanced interactive visualization and analysis methods providing linked environments of geospatial data and time series graphs that allow end users to explore infectious disease outbreak models, as shown in Figure 2. In this view, the map can be interactively filtered to show a variety of statistical measures about the disease spread (e.g., percentage ill, percentage dead), and plots in this view provide temporal details of the spread over a user selected geographic region.

Furthermore, in the geospatial view, users are able to interactively explore the simulation and insert decisions (e.g., quarantine counties, spray pesticides, enforce social distancing). The effects of these decisions are captured in the decision support tree space. The decision support space shows how the simulation paths vary over time where the height of the path is based on a user-defined metric of the decision impact as compared to the baseline metric (i.e., the simulation path in which no interdictions have taken place). In this manner, users are able to explore path choices and analyze the global impact over time. This allows users to explore both short and long-term ramifications of the decision measures employed.

2 Related Work

The exploration and visualization of simulation models and outputs has been a topic of much exploration. Matkovic et al. [20] proposed a simulation model view that provides a visual outline of the simulation process and the corresponding simulation model in an effort to bridge the gap between simulation model behavior and the dataset. Bruckner et al. [5] introduced a visual exploration approach for investigating parameter spaces for visual effects design utilizing sampling and spatiotemporal clustering techniques to generate an overview of resultant variations and temporal evolution. Kohlhammer et al [16] presented an overview of work on situational awareness, naturalistic decision making and decision-centered visualization. Guo [11] proposed a visual analytics approach to discover interesting patterns in spatial interaction data in order to design effective pandemic mitigative strategies and facilitate decision making process. The proposed approach consists of a new graph partitioning method to segment large interaction graph, a reorderable matrix to visualize major patterns and a modified flow map linked with reorderable matrix. Waser et al. introduced World lines [29] which integrated simulation, visualization and computational steering into a single unified system enabling user exploration of the entire solution space searching. Their steering mode enables the user to generate and control multiple simulation runs while visualization mode facilitates comparison between simulation runs while searching for an optimal solution. Our work follows many of the conventions employed in world lines, such as collapsing decision spaces and formatting the decision history tree to be temporally aligned. The major difference between the two works is the manner in which our decision history tree is formatted to represent the overall impact of a decision at each time point. This visualization creates paths which allow users to quickly explore the impact of decision at a given time point as well as the total impact at the end of the simulation run.

In each of the previous simulation analysis works, a key component is the tracking, evaluation and exploration of user events and interactions. Previous work in this area includes applications such as TimeTree [6], Grasparc [4], HyperScribe [30] and VisTrails [24]. TimeTree [6] is an interactive visual analytic tool that supports browsing large data sets while keeping the exploration process cognitively tractable. Brodlie et al. [4] introduced Grasparc, a framework that helps manage the problem solving process through the use of history trees to represent the solution exploration process. HyperScribe [30], an extension to Grasparc, provides a data management facility to organize and retrieve solution data in computational experiments. VisTrails [24] supports exploratory visualization by maintaining a detailed record of changes made to the workflows during the parameter exploration process and provides side by side comparison of results. Again, a history tree structure is utilized in order to manage the provenance data.

In conjunction with history tree structures, a variety of other visualization tools and systems have been developed to track other types of historical data. Baur et al [2] proposed an interactive visualization tool for displaying music listening histories along with contextual information and support to learn and understand these histories. LifeLines [22] provided a visualization environment for personal medical histories summarized in the form of lines and events highlighting relationships via search tool. Lifelines 2 [28] introduced several extensions to their earlier work which emphasized visualizing temporal categorical data for multiple records. AsbruView [17] extended the LifeLines concept to implement temporal view used to display hierarchical plan structures in medical therapy planning.

Heer et al. [12] performed the design space analysis of interaction histories and contributed a prototype graphical history interface for Tableau [19, 25] visualization system. This interface not only provides tools for history navigation and management but also supports sense-making, search and communication through some additional operations. This history information can be used to evaluate visualization design. CzSaw [15] captures the analysis process and corresponding user interactions, represented in the form of a scripting language which can then be viewed by analyst to identify repetitive patterns. Analysts can then look for different avenues without losing track of the previous analysis. Suntinger et al. [26] proposed an event-tunnel visualization and analysis framework which is based on two views of the cylinder: the side view

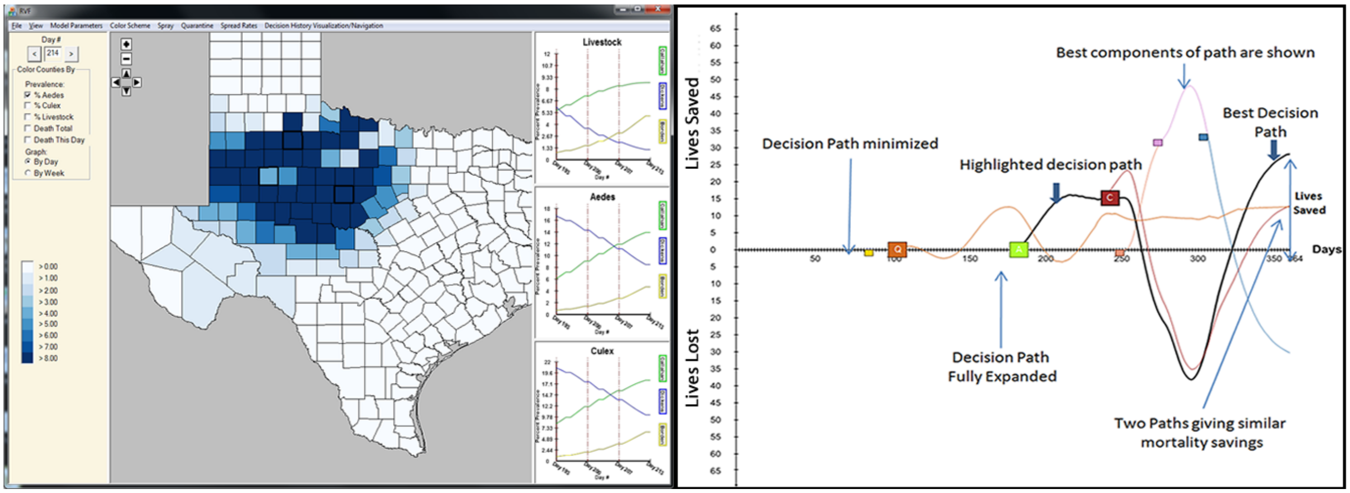

Figure 2: Visual analytics decision support environment. (Left) The spatiotemporal model view display. In this view, users can watch the spread of the model over space and time and introduce changes to the simulation as well as incorporate mitigative response measures to try and slow the disease spread. (Right) The decision history tree view. As users interact in the model view display, the different paths the simulation can take are calculated and visualized. The decision paths are plotted over time on the x-axis, with the y-axis representing the cumulative deviation from the baseline simulation.

(plotting the events in temporal order) and the top view (that looks into the stream of events along the time axis). These two interactive views of the event data can be embedded into a configurable workspace that supports analysis and mining. Work by Shrinivasan and van Wijk [23] also suggest the incorporation of history views into the analytical reasoning process for information visualization. Similar to the previous work in history visualization and analysis, our work also incorporates history trees and navigational structures; however, we modify the decision history space to allow for quick comparison between not only the decisions made but also the outcomes of the decision.

3 System Overview

Our visual analytics decision support environment consists of two main views as shown in Figure 2. The first view (Figure 2-(Left)) is the spatiotemporal model view in which users can interactively adjust model parameters and explore the effect of decision measures over space and time. As users interactively scroll through time and explore the epidemic spread model over the underlying population structure, and detailed views of the impact over time for a user selected region is shown in the small multiples plots on the right-hand side of Figure 2-(Left). In this window, users can interactively choose to employ decision measures for mitigating the outbreak. Decision measures are based on the model under investigation, and detailed examples of use cases are given in Section 4.

In the spatiotemporal model view, users may only interact with their current scenario, and modifications in this scenario are captured in our second view, the decision history view (Figure 2-(Right)). In the decision history view, users may explore the different decision paths and compare the cumulative outcomes over time. The decision history view supports path highlighting and branch collapsing as a means of reducing clutter and enabling effective exploration. Each of our two views supports a particular form of analysis and enables users to evaluate the effectiveness of various disease mitigation decisions.

These views are driven by an epidemic spread simulator in which a given model is integrated into the system. The model input parameters are fed into the simulator and the results of the simulation are modeled based on user-defined decisions. As users apply mitigative decision measures, the decision history tree is stored and a visualization of the user's choices can be displayed and analyzed. Our framework is based upon recommendations in the work by Jankun-Kelly and Ma [14] and Shrinivasan and van Wijk [23], both of which suggest that capturing metadata of the visual exploration process is a key component of the analysis process. Our history tree visualization is able to record users' decisions and allow them to compare, modify and insert new decisions. In this section, we describe the details of our system components and the related control structures.

3.1 Epidemiological Spread Simulator

In order to allow our system to be fully functional and adaptable to other models, our system contains an interchangeable epidemiological spread simulator component. This component generates a large scale spatial simulations over census tract, zip code and/or county level populations. Population and demographic[27] data is provided as input to the back-end simulation functions. The simulation then outputs information on the number of sick and dead within a given population by areal unit (e.g., county, zip code) and provides color coded geographical representations of the data. System control menus are then defined to incorporate the appropriate mitigative response measures that users can apply. Specific simulations used as case studies are described in Section 4. For a given model, the epidemiological spread simulator generates the epidemic spread data for specified number of days based on the given scenario and model parameters.

3.2 Spatiotemporal Model View

The spread data is then mapped into an interactive spatiotemporal view shown in Figure 2-(Left). This view facilitates the exploration of the disease spread through an interactive time spinner control seen at the upper left corner of the window. However, such exploration only provides slices of spatial data at a given time or an aggregate thereof. In order to understand these slices, users need to know the trends of previous data (and, if possible, model future data trends).

Users may interact with the system through a variety of viewing and modeling modalities. As shown in Figure 2-(Left), the main viewing area is the spatiotemporal view, and the three windows

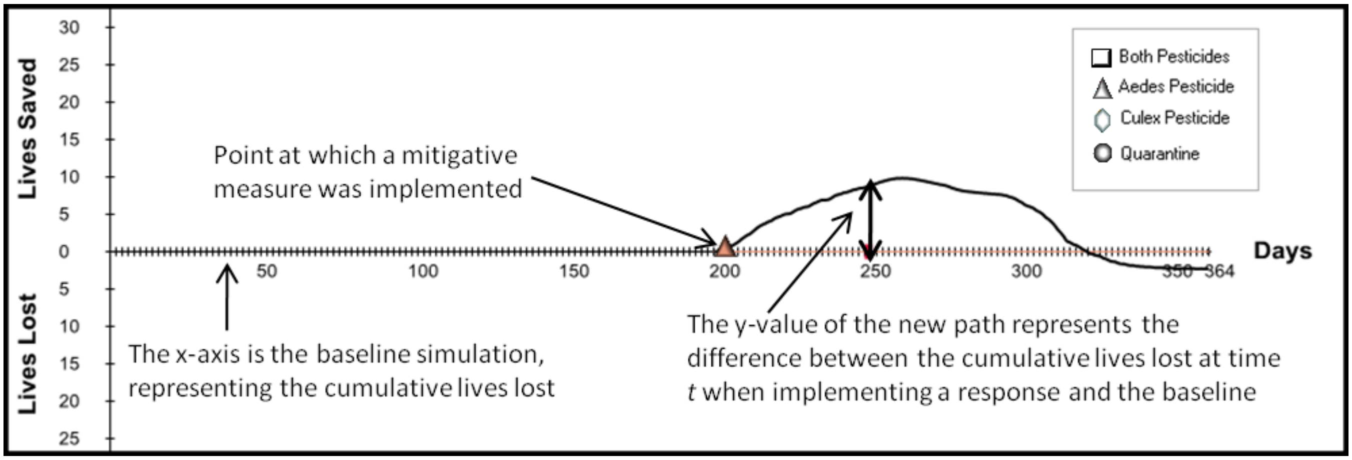

Figure 3: Creating a decision history tree. In this figure, we show the insertion of a decision point on day 200 of a Rift Valley fever simulation. The insertion of the decision points adds a node, and the resulting colored line shows the cumulative effects of this decision (as compared to the baseline) over time. If a path is above the x-axis, that decision has cumulatively performed better than the baseline up to that point in time. In this manner, users may track the magnitude of the disease spread with respect to the global impact.

on the right provide a time series view of the population statistics based on the underlying population and model parameters. These graphs provide a detailed view of the areal unit selected (with selections being indicated by a darker border) in the main viewing area, where each graph shows a different population statistic with respect to time. Both the geospatial and time series viewing windows are linked to the time control at the upper left portion of the screen. These linked views allow for a quick comparison of trends across various spatial regions.

As users explore the disease spread over time, they can also introduce mitigative response measures into the scenario. After each mitigative measure, the epidemiological spread model updates the scenario from that point in time. This form of exploration, which involves the human analyst inserting decision points into the scenario, provides a means for both creating training scenarios, as well as predicting possible future outcomes during an ongoing epidemic.

3.3 Decision History View

As the user inserts decisions points, scrolls through time, and revisits other scenarios, these interactions are tracked and displayed in the decision history view. This view keeps track of all the mitigative response measures performed and the corresponding mortality/sickness rate in a form of single visualization. This tool not only keeps track of the decision histories but also shows the consequences of each decision in terms of net gain or loss over time, creating the branching paths structures seen in Figure 2-(Right) and Figure 3. Currently the comparison can only be done according to a single criterion, and future work will explore ways of visualizing multivariate outcomes within a decision history tree. Note that all simulations used are designed to run until the disease has run its natural course, thus, the limits on the x-axis of our history view are derived from the model itself.

3.3.1 Path Building

In order to build the paths of Figure 3, a cumulative summation of the overall magnitude of the outbreak (in terms of lives lost or another user-defined variable) is calculated. In the original path, we consider the cumulative summation of lives lost on day t of the scenario to be the baseline. All other paths branch from this baseline such that the decision history view is visualizing the overall decision impact:

$$P_v(t) = \sum_{i=0}^{N} P_b^i(t) - \sum_{i=0}^{t} P_0^i(t) \qquad (1)$$

Here, $P_v(t)$ is the overall impact of the current decision path (P_b^i) in geographical area i with respect to the impact of the original un-mitigated simulation (P_0^i) at time t overall all N geographical areas in the simulation space. $P_v(t)$ is plotted on the decision history tree branching off from the last active decision path.

When the first mitigative response is added by the user, a decision path originates from the x-axis (representing time). The symbols shown in the legend represent the different types of mitigative response measures. Each decision path and corresponding symbol is assigned a unique color. The color of the decision path is temporarily changed to black whenever the user mouses over a decision path. In Figure 3 the brown triangle represents the point in time in which the user deploys some predefined resource as a means of mitigating the response. After this point, the height of the new path is calculated using Equation 1 for each time point left in the simulation. The height of the decision path along the y-axis represents the difference between the original (un-mitigated path) and the current branch.

Users may return to the original simulation as well and add other decision paths for comparison. Path selection is done through a right click within the interface and users may quickly move between scenarios. For each decision taken, a unique color is assigned to the path in order to facilitate analysis, as shown in Figure 1.

The height of these branches are again plotted with respect to the original decision path (as opposed to the parent path), thereby maintaining a consistent basis for comparison. If the branch is above the x-axis (positive), then the current decision path has helped mitigate the spread of the disease. In Figure 1, we can see one path that falls below the x-axis for a portion of time; however, it ends above the x-axis when the simulation run completes. Since the y-axis values represent a cumulative summation, this indicates that while this decision path may have seemed to be detrimental for a time during the simulation, it eventually proved effective in reducing the overall impact. However, in Figure 1, we can also see three other decision paths that end below the x-axis. Two of these paths would have initially appeared to be highly mitigative responses as they remain above the x-axis for a long period of time; however, we see that by the end of the simulation, this path actually would have worsened the overall impact of the spread. By using the decision history view, analysts may now see the end result as well as the path that it took to reach there. Furthermore, in the decision history view, analysts can see the effect of the decision and then utilize the model view display to explore periods of poor performance to look for other potential mitigative measures that could be added.

3.3.2 Path Analysis Tools

In order to explore a given decision path, users simply click on a path and choose the 'Activate Path' option from a menu. This path is then the path being explored in the spatiotemporal model view. Path activation requires us to introduce persistence support in the system and our system saves each previous state of the simulation before performing any mitigative measure. This saved state can be reloaded once any deactivated decision path is activated. Unfortunately, as the number and complexity of the scenario created by the user increases, the complexity of the visualization also increases. Mouse over on a node also provides a thumbnail view of what the geographical state space looked like at the time of the decision, as seen in Figure 1.

In order to reduce clutter due to the addition of a large number of decision measures, each decision path can be expanded or collapsed to any level. Whenever a path is expanded or collapsed, the last expansion/collapsed state is preserved for all sub paths. However, in some cases, users may wish to view a portion of the decision path and hide other obstructing nodes or lines. After introducing several branches within a given path, users may wish only to see the optimal path (the path with the highest cumulative score). In order to perform this operation, a user can right click on any path and select option 'Show Only Best Path Components'. Furthermore, in the presence of large number of decision paths/branches, it may become difficult to differentiate decision lines. Thus, our decision history tree also supports path highlighting to provide additional cues about path recognition.

4 Use Case Studies

In order to demonstrate our tools, we present two use case studies for two unique epidemiological spread models. Both models were adapted into the epidemic spread simulation component and minor changes to the user interface spatiotemporal model view were added to allow for the appropriate mitigative response measures. In this section, we briefly discuss the underlying model for each infectious disease scenario and then explore the decision analysis process.

4.1 Pandemic Influenza

Our first epidemiological spread model utilizes a Gaussian mixture model that simulates the spread of a pandemic influenza across the United States starting from a user defined point source location and incorporating airport traffic models [18]. The model makes use of a person-to-person contact model spread with a constant rate of diffusion in order to simulate a spatiotemporal outbreak. The model was designed to determine the number of influenza outbreak infections, hospitalizations, and deaths on a daily basis. As input, it requires the pandemic influenza characteristics, county data including population, demographics, and hospital beds, and decision measure anticipated impact. Spread vectors based on the point of origin and distance traveled per day are calculated, and effects on different age groups and population densities are taken into account.

Default parameters to our model are based on information from the U.S. National Strategic Plan [13]. In this plan, states are charged with the task of preparing for a pandemic influenza wave under the prediction that up to 35% of the population could be infected, 50% of the infected population will seek medical care, 20% of those seeking care will require hospitalization, and up to 2% of the infected population will die. These numbers are based on rates from the 1918 influenza pandemic [3, 8].

4.1.1 Mitigative Response Measures

Within our modeling tool, users are able to choose from three different global decision measures: (1) school closures; (2) media alerts; and (3) strategic national stockpile deployment (SNS). These decision measures were decided on based on requirements from the Indiana State Department of Health in order to best accommodate their training exercises. The choice of these decision measures is also influenced by previous work. Historical records of past pandemics illustrate the efficacy of social distancing with regards to lessening the impact of a pandemic. Furthermore, other researchers have noted the expected reduction of influenza transmission based on school closures or quarantines, and the effects of containing pandemic influenza through the use of antiviral agents and stockpiles have been well documented. Detailed descriptions of the effects of various decision measure strategies can also be found in [10] and [21], along with others.

Figure 4 shows how a user can simply toggle on and off decision points within the model view display to see their effects on the pandemic impact. Figure 4 (Left) shows the model on Day 36 with no decision measures employed. Using the controls on the lower left portion of the screen, the analyst chooses to deploy the SNS antivirals on day 3. The control widget shown in Figure 4 (Right) allows the user to set the day of the simulation on which the decision measure was enacted, the number of days it will take the decision measure to reach full effect, and the impact the decision measure is expected to have in reducing the infection. In the graphs of Figure 4 (Right), the user can immediately see how the use of the (SNS) has helped mitigate the magnitude of the pandemic.These graphs represent the number of sick cases, people who need hospitalization and number of deaths for selected counties in the spatiotemporal view. Through these controls, the user can interactively toggle decision points on and off and explore the effects that decisions taking place in the past would have on the current situation. Interactive toggling allows the user to understand the magnitude of the change by watching both the graphs and map display colors change for a given day as decision measures are implemented.

In this model, all decision points are designed to mitigate the spread, and each decision measure may only be deployed once. As such, the decision history view will show that all paths improve the outcome when compared to the base scenario. We use this example to demonstrate features of our tool and show that it is adaptable to multiple models.

4.1.2 Model Exploration

In this example, the user has created four different paths for exploration as shown in Figure 5. We have included a variety of decision measures along each path, including combinations of all three mitigative response. The maps surrounding the decision tree structure in Figure 5 represent day 45 of the simulation with respect to a given decision path as indicated by the labels. The decision measures and the times at which they are implemented are provided in Table 1. Note that each decision made in the table generates a branch; however, in order to evaluate the total path, we have collapsed the intermediate paths. In this manner, we only show the decision path resulting from the combination of the decisions shown in Table 1.

Here, the user can quickly see that Path D1 of Figure 5 is the optimal choice in terms of mitigating the outbreak based on the available decision metrics. It is clear that the earlier a decision is made, the more impact it can have on reducing the spread. As each decision point only has a positive effect on the disease reduction, the exploration task in this simulation is relatively trivial. However, this example is included to illustrate that this tool is easily adaptable to multiple models.

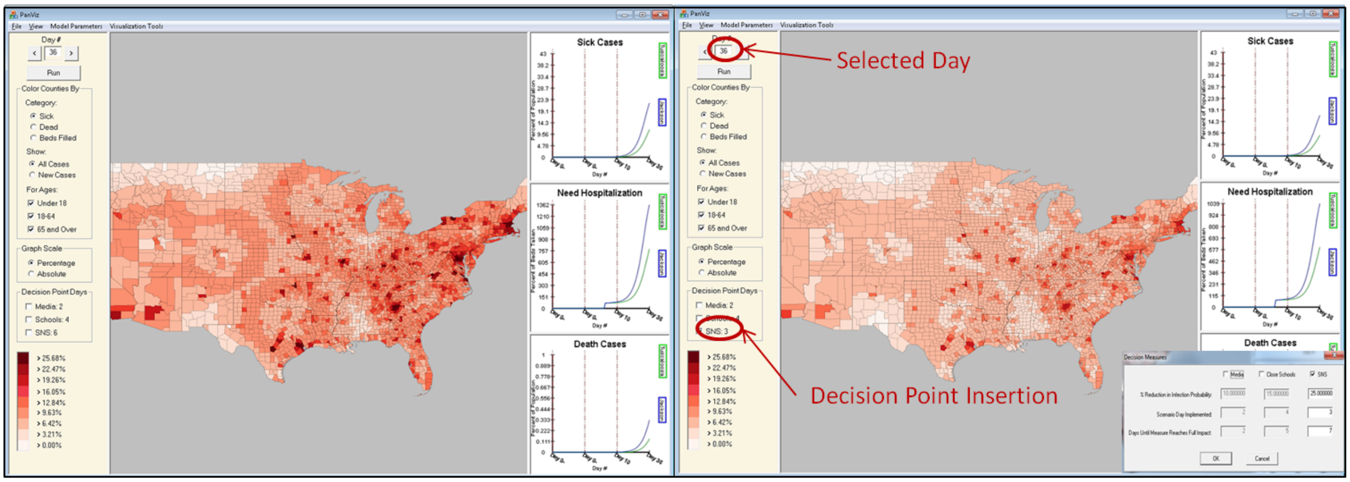

Figure 4: Here we illustrate the effects of utilizing decision measures within the confines of a pandemic influenza simulation. In the left image, the analyst has used no decision measures and is visualizing the spread of the pandemic on day 36 of the simulation. In the right image, the analyst has decided to see what effects (on day 36) deploying the strategic national stockpile on day 3 would have had on the pandemic.

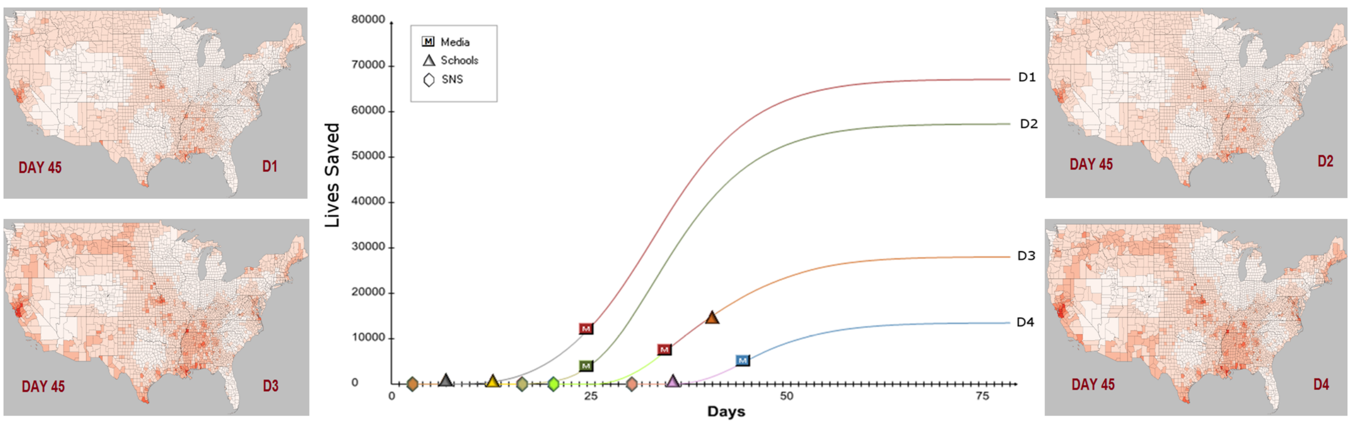

Figure 5: Pandemic Influenza Case Study. Here the user has introduced a variety of different decision measures at various points in time and in different combinatorial order. We explore the resultant simulation spaces in the geographical space with the maps surrounding the central image. Each map corresponds to a different decision tree branch as denoted by the corresponding label.

4.2 Rift Valley Fever (RVF)

Our second epidemiological spread model utilizes a differential equation model that simulates the spread of RVF through a simulated mosquito and cattle population in Texas [9]. As input, it requires the underlying county populations of both the cattle and two types of mosquitoes (Aedes and Culex) in the area. The differential equation model then accounts for the transmission of the disease both through mosquito to cattle infections and cattle to cattle infections. Mosquito larvae are also considered in this model as a means of the disease being prevalent at the mosquitoes' birth. Lives lost/saved in this case are referring to the underlying cattle population. Note the simulation stops at the state border.

4.2.1 Mitigative Response Measures

Within our modeling tool, users are able to choose from two different local decision measures: (1) pesticides; and (2) quarantine. Users are able to interactively apply a quarantine or pesticide spray to any individual county or multiple counties at once during the simulation. This is done by mouse clicking on counties and then selecting to spray or enforce a quarantine on those counties. Analysts can combine Aedes and Culex pesticides

Table 1: Summary of decision paths generated for the pandemic influenza simulation. Each entry represents the day a type of decision was employed in the path.

Decision Path	School Closure	Media	SNS
D1	6	25	2
D2	12	25	16
D3	40	35	20
D4	35	45	30

for a combined spray. This is represented by a legend item('Both Pesticides') in Figure 7. Our system can easily be modified to allow additional combinations without any performance penalty. In the case of multiple simultaneous decision measures, the parameters associated with the combined mitigative response measures will be updated and passed to the model simulator that will recalculate the simulation results.

Pesticides:
In order to apply pesticides, an analyst selects the set of geographical regions for this operation and also the type of pesticide to kill a particular type of mosquito species or both. Figure 6 shows that

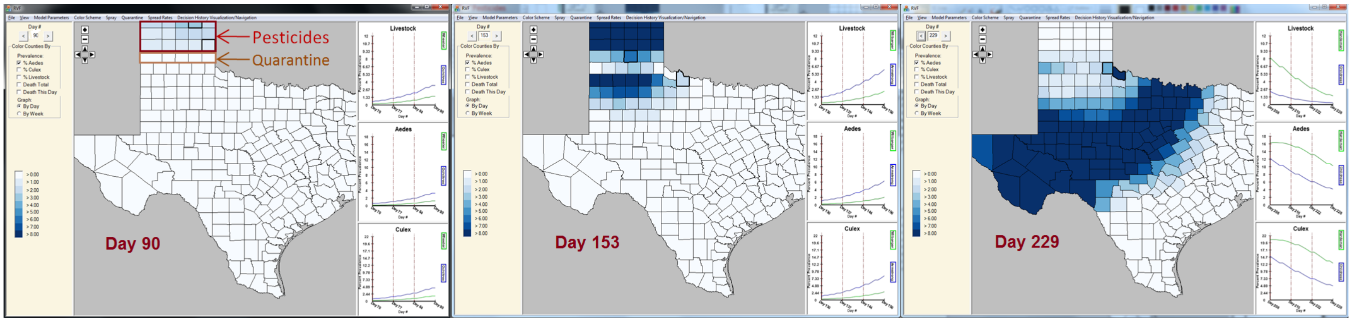

Figure 6: Here we illustrate the effects of utilizing decision measures within the confines of the Rift Valley fever simulation. In the left image, the analyst has employed both the use of quarantine and pesticide spray to try and reduce the disease spread. However, as infected mosquito eggs have already propagated to neighboring counties, they find that the decision measures taken have less impact then expected. The spread of the disease after the application of pesticides and quarantine is seen in the middle and right figures. The graphs in the right panel represent the population of livestock and population of Aedes & Culex mosquito types for selected counties within the spatiotemporal view. Selected counties are highlighted by drawing bold black boundaries.

Table 2: Summary of decision paths generated for the Rift Valley fever simulation. Each entry represents the day (or days) a type of decision was employed in the path.

	Pesticide Aedes	Pesticide Culex	Quarantine
A	110,262	262	-
B	143	143,263	85

after applying the pesticides even if all the mosquito population is killed along with the eggs, mosquitoes from neighboring counties may still migrate to the area. A portion of this migration is caused by the transfer of livestock from the neighboring regions or existing infected livestock may infect new uninfected mosquitoes. Analyst can try different combination of regions for pesticides and based on the simulated spread results and analyze which combination may work best within a given scenario.

Quarantine:
The second mitigative measure supported is the quarantine operation. In this operation, a user selects a set of geographical regions (as seen in Figure 6) which will no longer allow transport of cattle into or out of the region. Quarantine results in setting the travel rates for the livestock across selected regions to zero; however, mosquito travel is still unrestricted.

4.2.2 Model Exploration

In this model, the space of possible mitigative responses is limitless since the user can select any subset of geographical regions for pesticides or quarantine and the order/frequency of these mitigative responses may also vary. With 254 counties in Texas, the user can choose from a variety of decision measures. As previously stated, the goal of this tool is not necessarily to find the optimal solution; instead, the goal is to allow users to play out various scenarios based on their underlying knowledge of the resources available. Such a tool can then potentially alert decision makers to shortcomings in their plans or resources.

In this example, the user has created two main paths for exploration as shown in Figure 7. The decision measures and times at which they are implemented are provided in Table 2. The maps surrounding the decision tree structure Figure 7 represent the days in which responses were taken and the highlighted counties are those in which a response was implemented. Note that at no point during the exploration do the simulation parameters change. The only things that can impact the result of the simulation are the mitigative responses.

In the initial exploration, the goal was to explore the effects of not using a quarantine. Depending on the time of year the outbreak may occur, a quarantine could have significant impacts on beef sales. Thus, decision makers may wish to only spray for mosquitoes near those affected counties in an attempt to prevent the spread. The first decision in path A, is thus to spray. However, a pesticide that is only effective against one type of mosquito (Culex) was used, representing a deficiency in local supplies. Path A initially saved some lives but ended up slightly below the horizontal axis.

After seeing that Path A appears to have a mostly positive effect (with regards to reducing the disease impact on the cattle population), the user then decides to spray for mosquitoes at the apex of the current Path A. Counties near two major rivers are selected, and sprayed for both major types of mosquitoes. However, the resultant impact is actually worsened.

Without the decision support history tree to observe the response of A, it would be difficult to visually tell the resulting difference between the original simulation and the initial decision branch as the two result in a nearly identical loss of cattle. However, if exploring this initially, the user would see a gain and may have concluded that the decision path chosen would result in saving lives. In fact, the decisions taken here would only waste resources and result in approximately the same (or an even worse) outcome.

In path A, the user wanted to explore the effects of enforcing an early quarantine. Initially, this path appeared to provide little impact with respect to the baseline, following closely to the horizontal axis; however, by the end of the simulation, we see that the initial quarantine (the blue line labeled as a branch of path B) actually results in a fair number of lives save.

From this analysis, the user wants to see if they can do better by inserting a decision measure near the first uptick of the blue path at day 143. Here, the user preemptively sprays counties near the two major rivers and at the edge of the oncoming spread. The green decision path (again labeled as a branch of path B) is generated. Here, we see that the green path actually outperforms the blue path, resulting in an even higher number of lives saved.

Finally, the user inserts another decision measure at the apex of the green path (day 263), this time spraying for only Aedes mosquitoes. The spray is done over the same major rivers as the previous injection, and initially a large upswing is seen. Unfortunately, by the end of the simulation, the end result actually is worse than the quarantine alone and the quarantine combined with a secondary spray. This is labeled as Path B. From this, we see that early preventative measures work better (as expected); furthermore, late term measures can actually negatively impact the disease spread.

Based on the decisions taken, users can clearly see that the green branch of path B in Figure 7 is the optimal choice as compared to

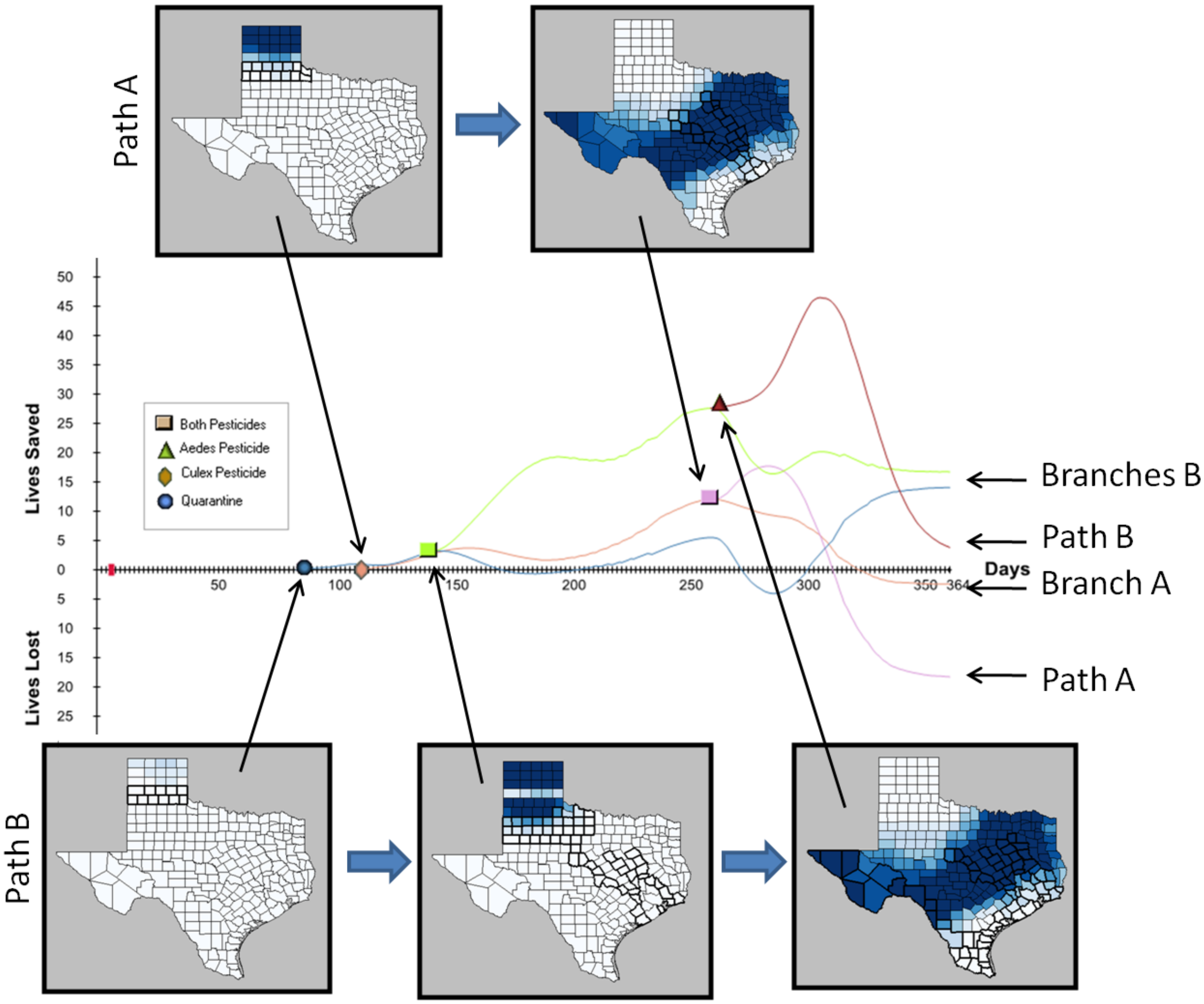

Figure 7: Rift Valley fever Case Study. Here the user has introduced a variety of different decision measures at various points in time and in different combinatorial order. We explore the resultant simulation spaces in the geographical space with the maps surrounding the central image. Each map corresponds to a different decision tree branch as denoted by the corresponding label.

other explored paths. It can be observed that by adding in the secondary decision measure, we reduced the effect of the downswing seen in the blue path near day 275. Users may also choose to go back and explore what went wrong in path A as this resulted in the largest loss of life.

Clearly, as the decision state space becomes more complex, the resultant decision history tree also grows in complexity. However, the current system allows users to interactively hide branches, thus allowing them to quickly remove suboptimal decision paths from the visualization. Future work will also focus on suggesting potential solutions based on the current simulation day and other potential user defined constraints. We also plan to allow for local history decision visualization (as opposed to the current global history option). Future work will explore the inclusion of small multiples embedded in mouse overs of the map viewing window as a means of displaying local history graphs for a given county.

4.3 Memory Requirements and Performance

Whenever a mitigative response measure is performed, the model recalculates the epidemic spread data beginning from the selected day through the end of the simulation. This newly calculated spread data, along with the mitigative measure description, is saved in a system structure that keeps track of the complete decision history. The last decision path added is automatically marked as active and any further decision will branch from the active decision path. Among all decision paths, only one path can remain active at a time. Users may reactivate any decision path by mousing over the desired decision path and selecting the activate path option. This mechanism enables the user to recover any intermediate state and obtain a frame-by-frame simulation. A frame-by-frame simulation shows the epidemic spread for the active path. The memory requirements to support such a mechanism depend on the number of geographical regions (e.g., counties/cities), nature of the model, and the number of days in the simulation. In our current implementation, there is a system constant that allows us to define the upper limit for the maximum number of decision paths that can be created. This constant can be adjusted by the user. In future extensions of this work, we plan to include *save* and *reload* option, that would support saving the least frequently used decision history paths to disk. An alternative approach would be to discard the simulation data for previous decision paths and save only the minimum possible information required to recalculate the spread data if a user reactivates the decision path. Unfortunately, this approach would be very costly if the simulation is expensive. This approach could be useful in scenarios where the memory requirements for keeping decision history paths are large and the simulation is relatively inexpensive.

We have tested the performance of the system for pandemic influenza and RVF epidemiological models. The system was tested on a machine with an Intel Xeon E5504 2 GHz Quad-Core processor with 6 GB physical memory and Windows 7 (64 bit) operating system. For pandemic influenza, the response was generated at interactive rates for 80 days of simulation for the entire United States. For the Rift Valley fever simulation, the system took on average 48.768719 milliseconds to calculate epidemic spread data for one day of simulation for the entire state of Texas at the county level (254 counties). The system took approximately 17.8 seconds to calculate the spread data for a complete year. The average response time was similar for all possible combinations of mitigative response measures. Most of the execution time was spent on the calculations by the RVF ordinary differential equation (ODE) solver.

5 Conclusions and Future work

The proposed decision support environment facilitates researchers, analysts and public health officials in their study of epidemic spreads under varying scenarios and decision measures. Our decision history visualization and navigational support helps the users analyze the consequences of their decisions over time and understand both the short term and long term impact of their mitigative responses. Users can also drill down into a given scenario using the spatiotemporal model view in order to better assess the effects of individual paths. The architecture proposed in this paper provides flexibility in terms of incorporating different epidemiological models and applying the same set of tools to identify suitable mitigative strategies in case of an outbreak. In our future work, we plan to include an economic model into the system that will help us visualize the impact of the epidemic spread and corresponding responses on local economy. We also plan to include additional mitigative tools in the set of available mitigative measures.

We also plan to investigate if the model could be extended with information such as weather conditions, population characteristics etc. This would enable us to incorporate additional factors while creating hypothetical spread scenarios and get more accurate spread modeling. We also plan to add new features to the visual interface including play/stop buttons to automatically animate the simulation. We also plan to add decision path *save* and *reload* features that would allow users to save the least frequently used decision path to disk and later restore them. This feature will make the system more scalable.

Furthermore, decision history trees can be integrated with any system in which the user wants to keep track of multiple simulation paths and compare the end results. For example, if we have a financial model capable of generating spatiotemporal data and an associated model simulator, we can attach a user interface to the model simulator. This interface would allow the insertion of decisions paths at different points in time. Each decision initiates a new simulation path and the corresponding end results. In addition, our system can also be used to validate how effectively the model captures a real life epidemic spread. For this case, we must have the past data available for validation. By initializing the model simulator with a given scenario and comparing the results of model with actual data, we can gain insight into the accuracy of the model with respect to real life behavior.

Acknowledgements

The authors wish to thank David Hartley and Tianchan Niu for providing details on the Rift Valley fever simulation and Min Chen, Harold Bosch and Denis Thom for their helpful discussions in editing this paper. This work was supported in part by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0002, the Foreign Animal and Zoonotic Disease Center, and the Defense Threat Reduction Agency under Award Number HDTRA 1-10-0083.

References

[1] Atlanta: Centers for Disease Control and Prevention. Community strategy for pandemic influenza mitigation in the United States - early, targeted, layered use of nonpharmaceutical interventions. Atlanta: Centers for Disease Control and Prevention, 2007.

[2] D. Baur, F. Seiffert, M. Sedlmair, and S. Boring. The streams of our lives: Visualizing listening histories in context. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1119 –1128, 2010.

[3] M. Billings. The influenza pandemic of 1918, 1997.

[4] K. Brodlie, A. Poon, H. Wright, L. Brankin, G. Banecki, and A. Gay. Grasparc: A problem solving environment integrating computation and visualization. In *Proceedings of the 4th conference on Visualization*, pages 102–109, 1993.

[5] S. Bruckner and T. Möller. Result-driven exploration of simulation parameter spaces for visual effects design. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1468 –1476, 2010.

[6] S. K. Card, B. Sun, B. A. Pendleton, J. Heer, and J. W. Bodnar. Time tree: Exploring time changing hierarchies. *IEEE Symposium On Visual Analytics Science And Technology*, pages 3–10, 2006.

[7] D. Dausey, J. Aledort, and N. Lurie. Tabletop exercises for pandemic influenza preparedness in local public health agencies. TR-319-DHHS, prepared for the U. S. Department of Health and Human Services Office of the Assistant Secretary for Public Health Emergency Preparedness, 2005.

[8] L. R. Elveback, J. P. Fox, E. Ackerman, A. Langworthy, M. Boyd, and L. Gatewood. An influenza simulation model for immunization studies. *American Journal of Epidemiology*, 103(2):152–165, 1976.

[9] H. D. Gaff, D. M. Hartley, and N. P. Leahy. An epidemiological model of Rift Valley Fever. *Electronic Journal of Differential Equations*, 115, 2007.

[10] T. C. Germann, K. Kadau, I. M. Longini, and C. A. Macken. Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences*, 103(15):5935–5940, Apr 2006.

[11] D. Guo. Visual analytics of spatial interaction patterns for pandemic decision support. *International Journal of Geographic Information Science*, 21:859–877, January 2007.

[12] J. Heer, J. Mackinlay, C. Stolte, and M. Agrawala. Graphical histories for visualization: Supporting analysis, communication, and evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1189 –1196, 2008.

[13] Homeland Security Council. National strategy for pandemic influenza. The White House website, November 2005.

[14] T. J. Jankun-Kelly, K.-L. Ma, and M. Gertz. A model and framework for visualization exploration. *IEEE Transactions on Visualization and Computer Graphics*, 13(2):357–369, 2007.

[15] N. Kadivar, V. Chen, D. Dunsmuir, E. Lee, C. Qian, J. Dill, C. Shaw, and R. Woodbury. Capturing and supporting the analysis process. In *IEEE Symposium On Visual Analytics Science And Technology*, pages 131 –138, 2009.

[16] J. Kohlhammer, T. May, and M. Hoffmann. Visual analytics for the strategic decision making process. In R. D. Amicis, R. Stojanovic, and G. Conti, editors, *GeoSpatial Visual Analytics*, NATO Science for Peace and Security Series C: Environmental Security, pages 299–310. Springer Netherlands, 2009.

[17] R. Kosara and S. Miksch. Metaphors of movement: A visualization and user interface for time-oriented, skeletal plans. In *Artificial Intelligence in Medicine, Special Issue: Information visualiztion in medicine*, pages 111–131, 2001.

[18] R. Maciejewski, P. Livengood, S. Rudolph, T. F. Collins, D. S. Ebert, R. T. Brigantic, C. D. Corley, G. A. Muller, and S. W. Sanders. A pandemic influenza modeling and visualization tool. *Journal of Visual Languages and Computing*, To appear 2011.

[19] J. Mackinlay, P. Hanrahan, and C. Stolte. Show me: Automatic presentation for visual analysis. *IEEE Transactions on Visualization and Computer Graphics*, 13:1137–1144, 2007.

[20] 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.

[21] G. Miller, S. Randolph, and J. Patterson. Responding to Simulated Pandemic Influenza in San Antonio, Texas. *Infection Control and Hospital Epidemiology*, 29(4):320–326, April 2008.

[22] C. Plaisant, B. Milash, A. Rose, S. Widoff, and B. Shneiderman. Lifelines: visualizing personal histories. In *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground*, CHI '96, pages 221–227, 1996.

[23] Y. B. Shrinivasan and J. J. van Wijk. Supporting the analytical reasoning process in information visualization. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1237–1246, New York, NY, USA, 2008. ACM.

[24] C. Silva, J. Freire, and S. Callahan. Provenance for visualizations: Reproducibility and beyond. *Computing in Science Engineering*, 9(5):82 –89, 2007.

[25] C. Stolte, D. Tang, and P. Hanrahan. Polaris: A system for query, analysis, and visualization of multidimensional relational databases. *IEEE Transactions on Visualization and Computer Graphics*, 8:52–65, 2002.

[26] M. Suntinger, H. Obweger, J. Schiefer, and M. Gröller. Event tunnel: Exploring event-driven business processes. *IEEE Computer Graphics and Applications*, 28(5):46 –55, 2008.

[27] United States Census Bureau. Population demographics, 2000.

[28] T. D. Wang, C. Plaisant, B. Shneiderman, N. Spring, D. Roseman, G. Marchand, V. Mukherjee, and M. Smith. Temporal summaries: Supporting temporal categorical searching, aggregation and comparison. *IEEE Transactions on Visualization and Computer Graphics*, 15:1049–1056, 2009.

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

[30] H. Wright and J. P. R. B. Walton. Hyperscribe: A data management facility for the dataflow visualization pipeline. In *IRIS Explorer Technical Report IETR/4, NAG Ltd*, 1996.