Dynamic Wide-area Congestion & Incident Monitoring Using Probe Data

Andrew S. Lund
University of Maryland, College Park
Center for Advanced Transportation Technology Laboratory
College Park, MD 20742
Phone: (301) 405-2874
Email: alund1@umd.edu

Michael L. Pack
University of Maryland, College Park
Center for Advanced Transportation Technology Laboratory
College Park, MD 20742
Phone: 301-405-0722
Email: PackML@umd.edu

Corresponding Author: Andrew S. Lund

Word Count: 4,999 Words + 10 Figures = 7,499 Words
Submitted in response to ABJ95-Visualization in Transportation and ABJ20 Statewide Transportation Data and Information Systems call for papers for presentation at the 2010 Transportation Research Board Annual Meeting and publication in the Transportation Research Record.

November, 2009
ABSTRACT

Traffic congestion and overall performance monitoring of roadways continues to be a major initiative of departments of transportation and planning boards nationwide. The information that can be gleaned from traffic monitoring programs has direct relevance to both analysts and the general public alike. Traditional data analysis tools and methods in this area fail to connect congestion data with incident and event data, making the task of determining the causes of congestion difficult in many cases. As more traffic data is collected, the need for tools that can facilitate effective visualization of this data, both archived and real-time, is becoming increasingly apparent.

A Congestion and Incident Scanner tool is proposed as a web-based application that affords dynamic and interactive analysis of traffic congestion conditions. The tool provides an intuitive, minimalistic interface for generating congestion performance visualizations for specific date ranges and locations. These visualizations improve upon existing tools by allowing users to interact and manipulate the visualizations to better highlight specific areas of interest. The integration of incident and traffic event data into the visualization allows users to easily correlate congestion abnormalities to possible causes, as well as evaluate the full effects of events, such as roadwork and major incidents, on traffic conditions.
INTRODUCTION

Congestion and overall performance monitoring of the nation’s roadways continues to be a major initiative of federal, state, and local departments of transportation and metropolitan planning boards. While congestion and bottleneck reports are used routinely for planning and operations, programs like the Highway Performance Monitoring System (HPMS) provide regular congestion measures to Congress so that they may establish both authorization and appropriation legislation which impacts program funding and highway taxation. The traveling public and shipping companies are also concerned with congestion conditions, especially when measures like travel-times can be delivered in real-time, allowing individuals to make more informed travel route decisions, which should ultimately prevent additional congestion and delays.

Traditional transportation operations and planning groups have relied on a mix of point-sensors, simulations, modeling, and statistical analysis to derive appropriate measures. Several problems remain with this approach including:

- the cost of installing and maintaining an adequate number of sensors on the roadways,
- the inability to collect, analyze, and redistribute the data in a timely manner,
- the reliance on modeling and simulation to “fill the collection gaps,”
- the lack of a dynamic way to allow data analysts and travelers to select start and end points on a road and immediately see congestion or travel time delays, and
- the appropriate visualizations that merge speed and congestion data with event data like weather, incidents, and construction data in a meaningful way.

The need for more ubiquitous data coverage remains. In an effort to solve some of the data collection and coverage issues mentioned above, the I-95 Corridor Coalition launched a new probe-based data collection initiative this year. On July 1, 2008 the I-95 Corridor Coalition began distributing continuous real-time travel information along the I-95 Corridor and key arterials from New Jersey through North Carolina. The data from this project, which covers well over 4000 centerline miles of freeways and arterials, comes from an amalgamation of probe technology (GPS-enabled vehicles) and road sensors. This data, discussed more later, is continuously delivered to stakeholders at 1-minute intervals, and is inherently different from traditional point-sensor data, presenting a wealth of presentation and analysis challenges as well as opportunities.

With this newfound wealth of data, the need for tools to analyze and present both real-time and archived data has become more and more apparent. With dynamic visualization, analysts could more easily identify and analyze recurring problem areas or review the full effects of incidents and construction events on traffic congestion. This information can then lead to better informed decisions on funding for future road improvement projects, operations, and planning.

In October of 2004, the Federal Highway Administration (FHWA) published a report titled “Lessons Learned: Monitoring Highway Congestion and Reliability Using Archived Traffic Detector Data” (1). In this report, the FHWA explores the potential usefulness of collecting and analyzing congestion performance data, and stresses the need for the use of effective presentation methods. One key
aspect included as an important part of presenting the data is the integration of event and incident information with the congestion performance measures.

This paper proposes a dynamic, web-based Congestion and Incident Scanner tool that can be used to evaluate congestion performance measurements from probe-based data with integrated traffic event information. The tool provides the user with a simple, interactive, and user-friendly interface for generating congestion measure visualizations for specific date ranges and locations. These visualizations improve upon existing tools that produce similar graphics by allowing users to interact and manipulate the visualizations to better highlight specific areas of interest. The integration of incident and construction data into the visualization allows users to evaluate possible causes of particular congestion events or even examine how bottlenecks may actually cause incidents.

RELATED WORK

Most archived traffic data management systems allow users to visualize traffic data as a series of 2-dimensional charts and graphs. Often each chart will only represent one specific detector location or some other measure. The Freeway Performance Measurement System (PeMS) (2), a data archive and performance monitoring system developed at the University of California-Berkeley with the California Department of Transportation, presents dashboard-style traffic monitoring through an impressive series of line graphs, tables, and a regional map with roads colored according to specific metric values. The University of Maryland’s Regional Integrated Transportation Information System (RITIS) provides similar reports and graphics, though less comprehensive, for individual detector locations (Figure 1).

While these visualization methods can be effective for a single location or point in time, they do not provide a way to graph a range of locations and times simultaneously.

Figure 1: The RITIS archived detector data analysis application viewing two detectors simultaneously.
In an effort to visualize corridor traffic conditions in a more meaningful way, the traditional contour-line mapping approach is often used. Both the Washington State Department of Transportation (WSDOT) and the Chicago Metropolitan Agency for Planning (CMAP) use this visualization method for presenting congestion data (3, 4). Using this approach, time and location can be plotted along the axes, and different colors can be used to represent different values of traffic metrics at each time and location on the graph. Visual representations of the roadways being charted are included between either side of the congestion visualizations as a location reference for points on the charts. WSDOT’s more map-accurate visualization of the road, as seen in Figure 2 and (3), is quite effective for roads which generally run in a straight line. However, they run into display issues with roads such as beltways and windy roads that often loop back on themselves. CMAP’s representation, seen in Figure 3 and (4), removes this issue by presenting the road as a straight line, allowing for any stretch of road to be represented, regardless of its shape.

Figure 2: WSDOT temporal and spatial congestion mapping graphics.
In a 2004 report prepared for the Federal Highway Administration on Monitoring Highway Congestion (1), both the WSDOT and CMAP graphics were referred to as “exemplary”. The report stressed that these types of graphics are more easily interpreted by diverse audiences than otherwise complex data tables and lengthy textual descriptions.

In addition to congestion visualizations, CMAP also produces a separate set of charts showing crash frequency at various locations and times of day along a road (4). Comparing these two visualizations side by side can allow correlations to be drawn between congestion and high crash counts, but having the information all integrated together in one chart would be preferred.

The value in creating visualizations of data stems from the innate human ability to quickly and accurately identify patterns, clusters, and abnormalities in graphics, and then rapidly recall and recognize similar patterns in other graphics (5, 6). By creating tools that allow for on-demand retrieval of data visualizations, these abilities can be used to ascertain and correlate information from the data much more rapidly than if viewed in a textual format. In a paper titled “The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations,” a design guideline for dynamic data visualization interfaces was presented that puts emphasis on allowing interaction with the data. Interactions specifically mentioned in this design guideline are zooming in on items of interest, filtering out uninteresting items, and providing “details-on-demand.” Being able to interact with the data in these ways can improve a user’s ability to focus on specific information, allowing for easier pattern recognition and correlations to be made in the data they are interested in.

The Advanced Interactive Traffic Visualization System (AITVS) (7) and the Portland Transportation Archived Listing (PORTAL) (8, 9) are two applications that have been developed to enable on-demand viewing of contour-line congestion visualizations for various roads. These applications present tremendous opportunity for analyzing historical traffic conditions for a given road, but they are limited to roads which have been equipped with stationary traffic detectors, and the only location options
available appear to be restricted to either the entire stretch of a detector-equipped road, or a specific detector location. Both applications present only one side of the road at a time, and do not appear to offer any sort of interaction with the visualizations once they are created, such as zooming in on certain areas of the chart or modifying the value ranges and colors used when visualizing the data. Finally, both of these applications lack the ability to associate event data with observed congestion conditions. These limitations make it difficult for these applications to be used for much more than observing general congestion trends along major freeways.

The following sections of this paper will discuss a new visualization tool that solves many of the problems discussed previously. The web-based tool is highly interactive, dynamic, and it produces meaningful visualizations that allow diverse groups the ability to detect bottlenecks and their association with incidents and construction events. These tools use data from regional integrated real-time incident databases along with probe-based space mean speed traffic data.

DESCRIPTION OF THE DATA

The traffic data used in this visualization comes from the I-95 Corridor Coalition’s Vehicle Probe Project. The primary source of this data, provided by INRIX, is GPS equipped fleets of vehicles. The probe data is supplemented by point-sensor technologies where available. The resulting data is the average speed of a particular road segment, where segments are defined by TMC codes. The data is refreshed at 1-minute intervals, and is then transmitted to the University of Maryland’s transportation data archive for storage. Since TMC segments vary greatly in length from road to road, there is an inherent challenge in how to display this in a meaningful graphic that can account for both space and time.

The incident and construction data used in this visualization comes from the University of Maryland’s Regional Integrated Transportation Information System (RITIS) which fuses data from multiple state and regional traffic management centers (like those in Maryland, Virginia, and the District of Columbia) into a single point of redistribution and analysis. Incident data will include the location and type of the incident, the number of lanes closed, the responders on the scene, the number and type of vehicles involved, and the variable message signs used to distribute information to the public.

THE CONGESTION AND INCIDENT SCANNER

The Congestion and Incident Scanner tool presented in this paper attempts to address the problem of visualizing archived congestion data by using a similar contour-line visualization approach as that of previous work done on the subject while adding significant functionality including enhanced interactivity and features which allow for more dynamic data retrieval and evaluation. The application’s web interface is divided into two main sections. Across the top of the screen, users are presented with an array of controls that are used for specifying what data should be queried, as well as modifying how the data is displayed in the charts. The space below these controls is used for displaying the data visualizations.

Data Query and Display Options

After logging into the system, the user is presented with a few options from which to choose. The first of these options is to select the road along with a start and end point on the route they wish to study. This is accomplished through a simple user interface using drop-down controls as seen in Figure 4. When a road is selected, the “Begin” and “End” drop-down options are automatically updated to reflect
only those cross-streets or reference points found along the selected road. Due to the large number of roads available in this selection, a state filter is provided for each of these drop-downs to simplify the choices available to the user. The drop-down controls also offer an “auto-complete” function whereby a user can type the name of a road in the box, and the list will automatically be searched for the text that is entered.

Figure 4: Controls used to define data query parameters.

Following the road selections, the user selects the desired date or date range. Selecting a date range will aggregate the speed data collected over the full date range selected, grouping the data by time of day over a 24-hour time span. The final selection gives the user data aggregation options which include 1-minute, 5-minute, 10-minute, and 15-minute choices. Once these parameters have been selected, the submit button is pressed, sending the user’s request to the server which queries the database and returns the visual results described below.

The remaining controls along the top section of the interface are used to manipulate the appearance of the congestion visualizations once data is returned from the database. Changes to these options are immediately reflected in the congestion visualizations generated by the application without having to resubmit a query.

The grey box labeled Data Options found in the upper right corner of the interface allows the user to switch between three different display modes (Figure 5). These modes include Percent of Reference Speed, Percent of Average Speed, and Recorded Speed. These options can be switched at any time by the user even after the query has been run. The differences between these three modes and their intended uses will be explored later in this paper. Also included in the Data Options box is an array of controls used for modifying the colors used in the congestion visualizations, as well as the range of values each color represents. By modifying these values, users can highlight specific traffic conditions and customize the visual representation of the results.

Figure 5: Controls for modifying the appearance of congestion visualizations after the data query has been run.

Outside of the Data Options box are controls to toggle the display of incident and event icons, create a link to the currently displayed data, generate an image or PDF file of the current display, and adjust which hours of the day are shown on the visualizations.
Traffic Congestion Visualizations

The lower section of the interface presents the data returned by the user’s request. This section consists of three columns. The center column contains a visual representation of the road, including all cross street locations, while the two outer columns contain the primary congestion visualizations for each side of the road. The center road diagram includes labels placed along its length, representing the various intersections and points of interest along the road, and serves as a visual reference for users to easily identify specific locations when viewing the congestion visualizations.

The congestion visualizations used in this application employ the contour-line mapping approach that has been well established by previous work on the subject. The x-axis of the charts represents the time of day that the data was recorded. By default, this axis shows a full 24-hour time span beginning at 12:00 AM. The y-axis of the charts represents the distance, in miles, along the stretch of road being shown. The speed data for a particular location at a particular time is plotted as a color representing the range of values in which it falls. This color and value range is defined by the user in the Data Options box at the top of the screen. The charts on either side of the center road graphic are labeled with which side of the road they represent, as well as the direction of traffic flow.
Display Modes
As shown in Figure 5, there are three different data display options that the user can switch between:
Percent of Reference Speed, Percent of Average Speed, and Recorded Speed. By default, the Percent of Reference Speed display mode is used. These values are determined by comparing the measured speed to the reference speed of the road at the location the speed was taken. A road’s reference speed is the average free-flow speed of traffic along the road, and is generally comparable to the road’s posted speed limit. Lower values in this mode indicate traffic is moving slower than the intended speed for the road, and represent congested conditions. Figure 7-left shows the Percent of Reference Speed visualization for the Outer Loop of the Capitol Beltway in Maryland and Virginia between the exits for I-95 North and VA-123. The red and orange masses indicate areas of congestion. In this example, there is a long stretch of congestion (dark red by default) that stretches from VA-193 all the way back to MD-650, as well as a couple of additional smaller, less severe areas.

Figure 7: Percent of Reference Speed (left) and Percent of Average Speed (right) display modes.

The second display mode option is Percent of Average Speed. This mode compares the recorded speed to the normal (or average) speed of the road at the location and time the speed was measured. This average speed is determined by averaging archived recorded speed data for that specific location, time of day, and day of week. By plotting these values, normal areas of congestion are effectively filtered out of the visualization, leaving only abnormal congestion areas visible. Figure 7-right shows the same location and time as Figure 7-left, but is using the Percent of Average Speed display mode. Congestion drawn in the upper left and lower right portion of Figure 7-left disappear in Figure 7-right, which indicates that those particular areas of congestion were due to normal rush-hour traffic. The large section of congestion remaining between 8:30AM and 11AM stretching from VA-193 to MD-650 is still visible, which indicates abnormal congestion.
The final display option, *Recorded Speed*, uses only the recorded speed values, in miles per hour (mph), to create the graphics. This display mode is used for easily highlighting specific speed ranges in the data without having to know the reference speed of a road, or convert the desired speeds into a percentage of the reference value. By default, speeds are broken into ranges of 10 mph, starting with 0-10 mph. These ranges roughly correspond to the default *Percent of Reference Speed* values for roads with a reference speed of 60 mph.

**Incident and Event Overlay**

Being able to discover anomalies in traffic patterns can be an important first step in discovering trouble spots along a particular stretch of road, but speed data alone is often not enough to understand the full picture of why congestion occurred. After discovering an area of abnormal congestion, the next logical piece of information to look for is what caused the congestion. This application attempts to solve this problem by displaying interactive icons representing traffic events and incidents on top of the congestion visualizations. By integrating event data into the application, correlations between areas of congestion and events or incidents may be easily drawn.

![Figure 8: A stretch of I-495 in Maryland and Virginia, showing 11 incidents on March 12th, 2009.](image-url)
icons indicate the duration in which the event or incident was active for. Figure 8 shows the same stretch of road as Figure 7, with data recorded on March 12th, 2009. Notice that there are 11 incidents visible. Clicking on an icon will also open a new window with a more in-depth description of the event which might include lane closure information, responder information, response plans, number of vehicles involved, etc.

When moving the mouse cursor over the incident icons or any other portion of the congestion visualization, a tooltip will appear indicating the time of day as well as the recorded speed, percentage of average speed, or the percentage of the reference speed at the current mouse position depending on which display mode the user is in (Figure 9-right).

Figure 9: Example of an incident icon pop-up (left) and a speed data pop-up (right).

Users can click and drag the mouse horizontally on the charts to highlight specific times of day they wish to view closer, which will cause the application to zoom in and display only data from the selected time range. Zooming can also be achieved by using a mouse’s scroll-wheel. Figure 10’s insert shows a user selecting the area between 8AM and 12PM of the long stretch of congestion on I-495 that has been previously discussed. The main image of this figure shows the resultant zoomed-in view. With the event and incident icons enabled, a number of events and incidents can be quickly associated with the congestion. At the front of the congestion area is an icon corresponding to a traffic incident. Moving a mouse cursor over this icon reveals that the incident caused 3 lane closures, as well as a blocked shoulder at its worse point (Figure 9-left). The duration tail for this incident indicates that the event lasted close to an hour and a half, and extends nearly the full width of the congestion area.

The visualization shows that the incident occurred after congested conditions had already begun. This could indicate that the incident was a result of sudden congestion caused by some other factor. However, it is more likely that the incident actually occurred slightly earlier, but its entry into the incident management system was slightly delayed due to communication issues or because of verification problems. Regardless of what caused the incident, the amount of lanes closed and the duration of the incident indicate that this incident was the primary cause of the abnormal congestion observed in this example.
Figure 10: An area of congestion being highlighted by the user (insert), and the resulting zoomed-in view of that area following the selection.

An interesting trend to notice about this particular incident is the gradual back-up of congestion along the road over time, which creates the right-leaning effect seen in the congestion area in Figure 10. As time progresses to the right, the congestion moves further and further upstream of traffic. Along either side of the congestion area, additional event and incident icons can be seen. A “Congestion Event” icon and a “Planned Roadway Closure” icon are displayed on the same side of the road. The congestion icon indicates that an abnormal congestion area was observed and recorded in the incident management system, which is clearly confirmed through the speed data underneath. The planned roadway closure, while not directly related to the incident, most likely contributed additional congestion to the already extended backup. The right side of the congestion area contains two icons indicating disabled vehicles. While there is not enough information associated with these incidents to fully conclude what caused the vehicles to become disabled, their close proximity to the congestion area on the chart could indicate that they are related. Taking this into account, one could conclude that they were the result of a low speed collision that occurred during the congestion, or possibly they could have become disabled as a result of some mechanical failure due to the stop-and-go, low speed travel caused by the congestion or running out of gas. Figure 8 also shows that there are small pockets of congestion on the opposite side of the road where the incident occurred. This congestion is likely the result of rubber-necking.

Of particular interest is the fact that the incident discussed in Figure 9 appears in the incident management system almost 40-minutes after there are noticeable and significant speed reductions on the...
roadway. This implies that there was a significant delay in receiving notice of this incident and/or entering the incident information into the traffic management center system. This lag time can be attributed to several sources. This is only one such example of serious delay in notification and incident detection that is visible from within this tool. Decision makers have seen these visualizations and have used them to support the funding of new programs in the National Capital Region which are targeting quicker incident detection and enhanced information sharing between systems in the region.

**FUTURE WORK**

While the Congestion and Incident Scanner presented in this paper has much of its intended functionality implemented, there are additional features that would enhance the end-user experience. The road selection process currently in place can become cumbersome in some cases due to the potentially large amount of roads and corresponding intersections that can exist in a given area. Users may not be familiar with every road name or intersection presented as options in the drop-downs, potentially causing confusion when attempting to select a road. An interface allowing users to choose a section of road on a map would help by allowing users to view specific data without having to be familiar with nearby intersections.

Another planned improvement is to enable users to view speed, volume, and occupancy data collected from stationary roadside traffic detectors where available, in addition to the already implemented speed data collected from probe detectors. This would allow further data correlations between the new traffic measurement types and traffic events to be made, which may provide a better insight into the nature of congested areas.

There are other intelligent transportation devices, such as Variable Message Signs, that could also be represented on the visualization tool. It has long been argued that lengthy or alarming messages placed on these signs can adversely affect speeds and actually cause congestion. Placing active signs and the content of their messages onto the visualization could help prove or disprove this theory and help to shift operational strategies.

Users must currently “ask” the tool to tell them about conditions on a particular stretch of road over some date and time. Thus, a user must be actively searching for problems. By developing an automatic data analysis component for this application, users could be guided to specific roadway locations that have recurring or abnormal traffic congestion issues. This would allow users to quickly pull up visualizations of trouble areas in order to analyze the causes and overall impacts of specific congestion events.

**CONCLUSIONS**

This paper presented the Congestion and Incident Scanner, a web-based application that is used for visualizing archived congestion performance measurement data with integrated traffic event and incident data. The application allows users to view congestion data visualizations for a wide area of roadways, giving them options as to which specific section of road and from what time and date range to view data. Users can manipulate the visualizations in real time by changing what metric values are shown, the value ranges and colors used to represent the metric values, and zooming in on specific times of day.
The application also addresses the need for integrating traffic event and incident data into congestion performance visualizations by including icons representing these events in the visualizations. With this data, users can explore potential causes of nonrecurring congestion or evaluate the effect of an incident on traffic.

Planners can use the tool to determine if bottlenecks are the result of incidents, which could be occurring at high rates that may need safety countermeasures, or to see if the bottlenecks may be the result of other geometric or volume-related issues. By analyzing these details, planning agencies and operations groups can better focus on improving those areas that pose higher safety risks to the travelling public.

**ACKNOWLEDGEMENTS**

The authors wish to thank the Maryland State Highway Administration office of CHART and ITS Development for their sponsorship of this project.
REFERENCES


