Visual Analytics for Transportation Incident Data Sets

Krist Wongsuphasawat, Michael L. Pack, Darya Filippova, Michael VanDaniker, and Andreea Olea

Transportation systems are being monitored at an unprecedented scope, which is resulting in tremendously detailed traffic and incident databases. Although the transportation community emphasizes developing standards for storing these incident data, little effort has been made to design appropriate visual analytics tools to explore the data, extract meaningful knowledge, and represent results. Analyzing these large multivariate geospatial data sets is a nontrivial task. A novel, web-based, visual analytics tool called Fervor is proposed as an application that affords sophisticated, yet user-friendly, analysis of transportation incident data sets. Interactive maps, histograms, two-dimensional plots, and parallel coordinates plots are four featured visualizations that are integrated to allow users to interact simultaneously with and see relationships among multiple visualizations. Using a rich set of filters, users can create custom conditions to filter data and focus on a smaller data set. However, because of the multivariate nature of the data, finding interesting relationships can be a time-consuming task. Therefore, a rank-by-feature framework has been adopted and further expanded to quantify the strength of relationships among the different fields describing the data. In this paper, transportation incident data collected by the Maryland State Highway Administration's CHART program are used; however, the tool can be easily modified to accept other transportation data sets.

Traffic management centers (TMCs) throughout the country generate detailed logs of hundreds of traffic incidents each day. These logs include data about the time and location of an incident, number of vehicles involved, lane closures, weather and road conditions, incident severity, and the like. Although and the transportation community and the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users emphasize developing standards for storing incident data, little effort has been made to design appropriate visual analytics tools to analyze the data, extract meaningful knowledge, and represent results.

Exploring these data-rich logs to glean any significant meaning can be an overwhelming task. Incident data is inherently difficult to explore because of its multivariate nature. Inferring the causal relationships of and trends within incident data can be problematic without using statistical methods. Many current tools are geared toward either geospatial analysis only or nongeospatial analysis only, and do not fully encompass all aspects of incident data. Furthermore, these tools are often either too complex for an untrained user or are rigid in their design and allow only limited functionality.

The challenge is so daunting that many state departments of transportation (DOTs) are either forced to hire dedicated information technology staff to facilitate data mining requests, or worse, simply do not attempt to perform any meaningful exploration of the data. For county and city DOTs, the problem is compounded by lower budgets and a lack of resources. Even when budgets allow for analysis, experts must often make detailed data requests, wait for staff to generate queries, return data, review the results, and then revise the request and start again. This slow process is often frustrating and does not allow one to truly explore data and ask questions, and it often fails to yield meaningful knowledge. Clearly, transportation professionals at the national, state, and local levels need better tools to empower them to accomplish sophisticated analytical data mining in an efficient and effective manner. These tools need to allow a user to more readily discover trends and patterns that would not normally be obvious.

A novel, web-based, visual analytics tool called Fervor is proposed as an application that affords sophisticated yet user-friendly analysis of transportation incident data sets. The tool provides the user with an intuitive suite of functionality that includes data filtering, geospatial visualizations, statistical ranking functions, and multidimensional data exploration capabilities.

Although each function is powerful by itself, Fervor integrates them, allowing users to simultaneously interact with and see relationships among multiple visualizations.

The Fervor application's rank-by-feature framework introduces interesting distributions and relationships according to many criteria: correlation coefficients, uniformity of distribution, number of outliers, and the like. This framework is particularly unique in that most of the ranking criteria in the current literature are suitable only for numerical variables; however, transportation incident data sets consist mainly of categorical variables. This paper also proposes several novel ranking criteria for categorical data.

RELATED WORK

Most states provide access to summaries of their transportation incident data in various forms online (*1–5*). However, these summaries are typically pregenerated reports that do not allow for any interactivity or individual analysis. A few national and state-sponsored transportation incident data analysis tools have been created. For example, the USDOT Volpe Center has published a web page that

Center for Advanced Transportation Technology Laboratory, University of Maryland, College Park, College Park, MD 20742. Corresponding author: M. L. Pack, PackML@umd.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2138, Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 135–145.
DOI: 10.3141/2138-18

allows individuals to access basic national safety data lookup functions (*6*). Like the previous examples, however, these reports are pregenerated, and there is no ability to customize or filter results past that of the state overview level.

The Fatality Analysis Reporting System (FARS) website is sponsored by NHTSA and allows users to create highly customized data queries (*7*). Although the FARS website represents a substantial leap forward in online data retrieval, the FARS tool leaves the burden of visualization and analysis up to the individual user. That is, the user must download raw data, and then perform graphing and statistical functions independent of the website application.

In a 2007 presentation given to the AASHTO Safety Data Systems Task Group, it was noted that new data analysis tools were needed to evaluate transportation incident data. These tools needed to be based on geographic information systems (GISs), be simple, be straightforward, provide for analysis of data in existing state data formats, be available to a variety of user levels, and be expandable (*8*).

To best allow a user to glean the most information from a data set, one must first present the user with an overview (table, map, or picture) of the entire data set. Next the user must be able to zoom and filter the data. As this occurs, the user must be able to access details about the data that remain on demand. Only from this top-down approach can the user begin to recognize patterns, realize what questions should be further probed, and notice trends that would otherwise go unnoticed (*9*). However, if any of these tools are to be successfully implemented, they must be designed with simplicity, speed, and ease-of-use in mind.

Many local and state DOTs have sought such functionality from GISs, such as ESRI's ArcGIS suite of products (*10*). These highly specialized geospatial tools allow for extremely complex spatial data analysis, management, and cartography; however, these powerful products have complex interfaces that can require months and even years of extensive training to fully master. Other tools, such as CommonGIS (*11*) and GeoVista (*12*), attempt to make data visualization simpler, yet in doing so lose the granularity and power that is needed to perform true transportation incident data analysis. Similarly, commercial data visualization products such as Spotfire (*13*) and Tableau (*14*) include mapping subcomponents. However, both tools show data points as single icons which results in occlusion and overcrowding for large data sets such as transportation incident data. Dykes et al. explore population density data using Google Earth in conjunction with other open-source products (*15*). In their example, the Google Earth API (*16*) represents a single entry in the data as a pushpin on the map. As with Spotfire and Tableau, the use of pushpins does not work well with clustered data because of problems with occlusion.

Heat maps are a method for representing spatial data that reveal the high-occurrence areas without obscuring the general view. Other GIS tools (*10, 17, 18*) have used heat maps successfully for a number of purposes, although the difficulties in analyzing transportation incident data with these types of geospatially oriented tools has already been addressed. The geospatial and heat map functions of these applications need to be merged with more user-friendly statistical analysis tools.

In addition to problems relating to ease of use and occlusion, transportation incident data sets are composed of multidimensional data. Dealing with multidimensionality has been a challenge to researchers in many disciplines because of difficulties comprehending data in more than three dimensions while searching for relationships, outliers, clusters, and gaps. This challenge is well recognized and has been dubbed "the curse of high dimensionality." Seo and Shneiderman (*19*) present a conceptual framework for relationship detection called the rank-by-feature framework and demonstrate its capabilities in the Hierarchical Clustering Explorer (HCE). HCE ranks all possible axis-parallel projections against the selected criterion and presents the result in a color-coded grid; however, the ranking criteria in HCE are only applicable to numerical variables. Many variables in transportation incident data sets are categorical (e.g., weather could have values, such as dry, rainy, or cloudy), which indicates the need for a categorical variable ranking methodology. Continued work on HCE by Seo and Gordish-Dressman suggests estimating the relative significance between two categorical variables based on a chi-square test (*20*). Later sections of this paper further explore relationships among categorical variables and discuss several new ranking criteria.

DESCRIPTION OF THE INTERFACE

Fervor follows the general design guidelines of "overview first, zoom and filter, then details-on-demand" as described by Shneiderman (*9*). The screen space is distributed into three sections, as shown in Figure 1. On the left is the control panel, which contains filter and ranking panels. The filter panel contains a rich set of filters that allows users to narrow the data set. The ranking panel adopts the idea of rank-by-feature framework, allowing users to rank the variables or relationships among variables by various criteria. The main area on the right is horizontally divided into two resizable sections. By default, the top right section is dedicated to the map, featuring two displaying modes. Users can select between icon mode, plotting every incident as a circle on the map, or heat mode, using the heat map algorithm to draw color regions on the map. The bottom right section locates a traditional tabular view of data and three visualization components: histogram, two-dimensional (2-D) plot, and parallel coordinates plot. The histogram shows distributions of the data set for a selected variable. The 2-D plot visualizes relationships between a selected pair of variables in scatter plot mode and grid mode. The parallel coordinates plot provides an overview of distributions and relationships of multiple variables. To provide flexibility, users can also drag and drop to move components between the top right and bottom right sections.

A significant feature of Fervor is that all of these seemingly individual components are tied together. Filtering the data set using the left side of the screen will immediately affect all other components. Selecting a particular incident or set of incidents in one component will also highlight identical selections in other components. Figure 2*a* shows how while viewing a histogram, one can select a cluster of points on the map, which immediately highlights those data points on the histogram. Similarly, clicking on an area of the histogram highlights those corresponding data points on the map (Figure 2*b*).

Sample Workflow Example

Fervor starts by loading the entire data set. The interactive map immediately shows a heat map of the incident locations, allowing the user to identify hot spots, or high incident locations. On the filter panel, the user deselects a checkbox for disabled vehicles, removing those incidents from the data set. The user visually identifies a hot spot on the highway and zooms in on it. The user selects this cluster using the selection tool. Clicking the "Delete inverse" button, the user dismisses all data points except those corresponding to the selected incidents. The user further explores the data points in

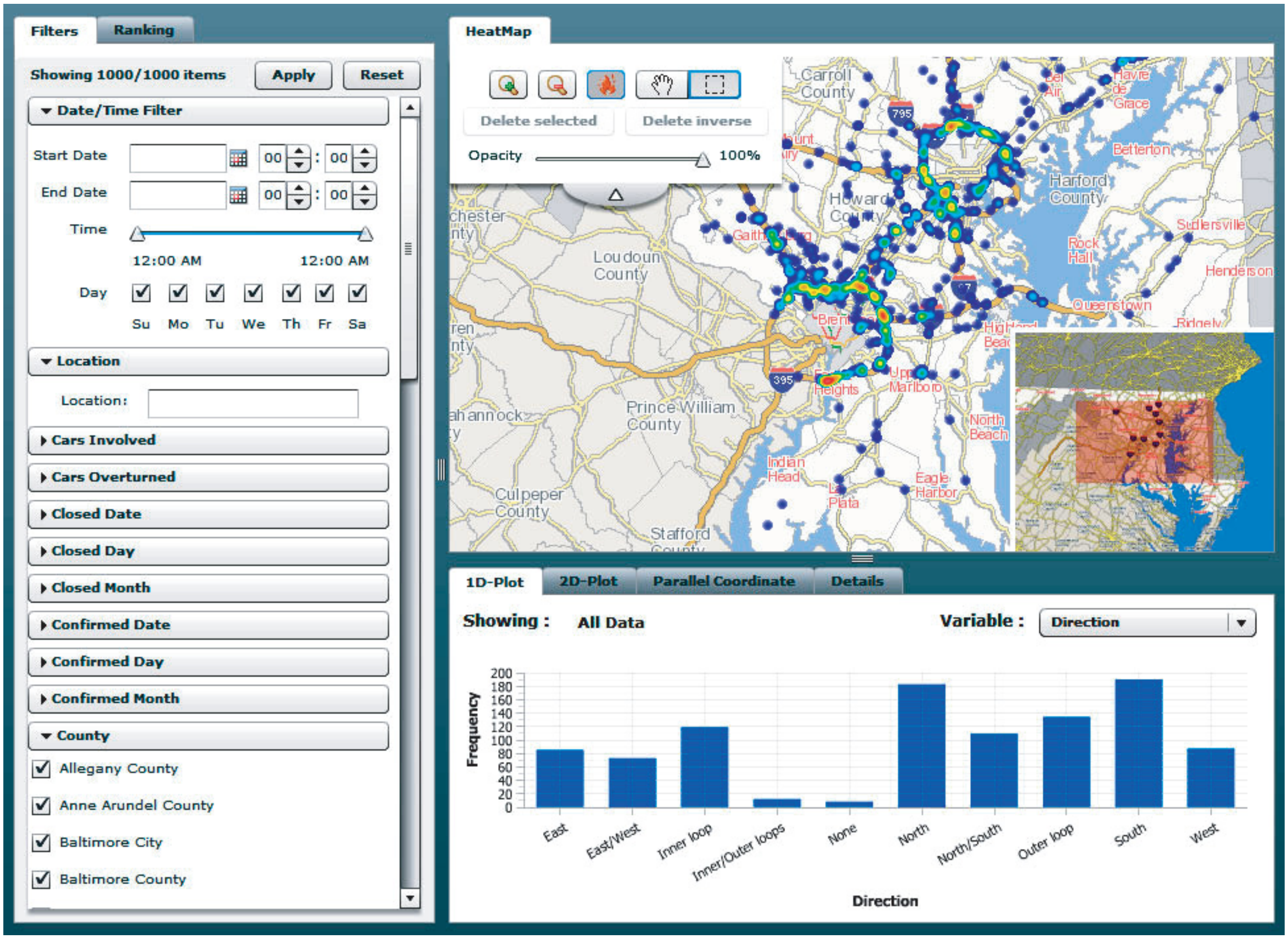


FIGURE 1 Fervor interface.

the cluster by way of the parallel coordinates plot to get an overview of the data distribution. One might discover that there were more accidents during sunny conditions. Next, the user can use the 2-D plot to find out whether there are other confounded variables correlated with the weather condition. However, there are many possible relationships in the data set. Examining every relationship one at a time can be time-consuming. Using the ranking panel, the user can rank the relationships by “Maximum Frequency.” This helps the user distinguish relationships that contain big clusters from other relationships, thus reducing the number of relationships to examine. Relationship between weather conditions and direction appears to have the highest ranking score. The user can click on the relationship to show it on the 2-D plot and discover that the majority of these events occurred in the westbound direction. Further analysis of the temporal information using the histogram reveals that 90% of the incidents occurred at 6 p.m. The user could logically infer that this particular hot spot is likely caused by the hindered vision of drivers heading directly toward the setting sun.

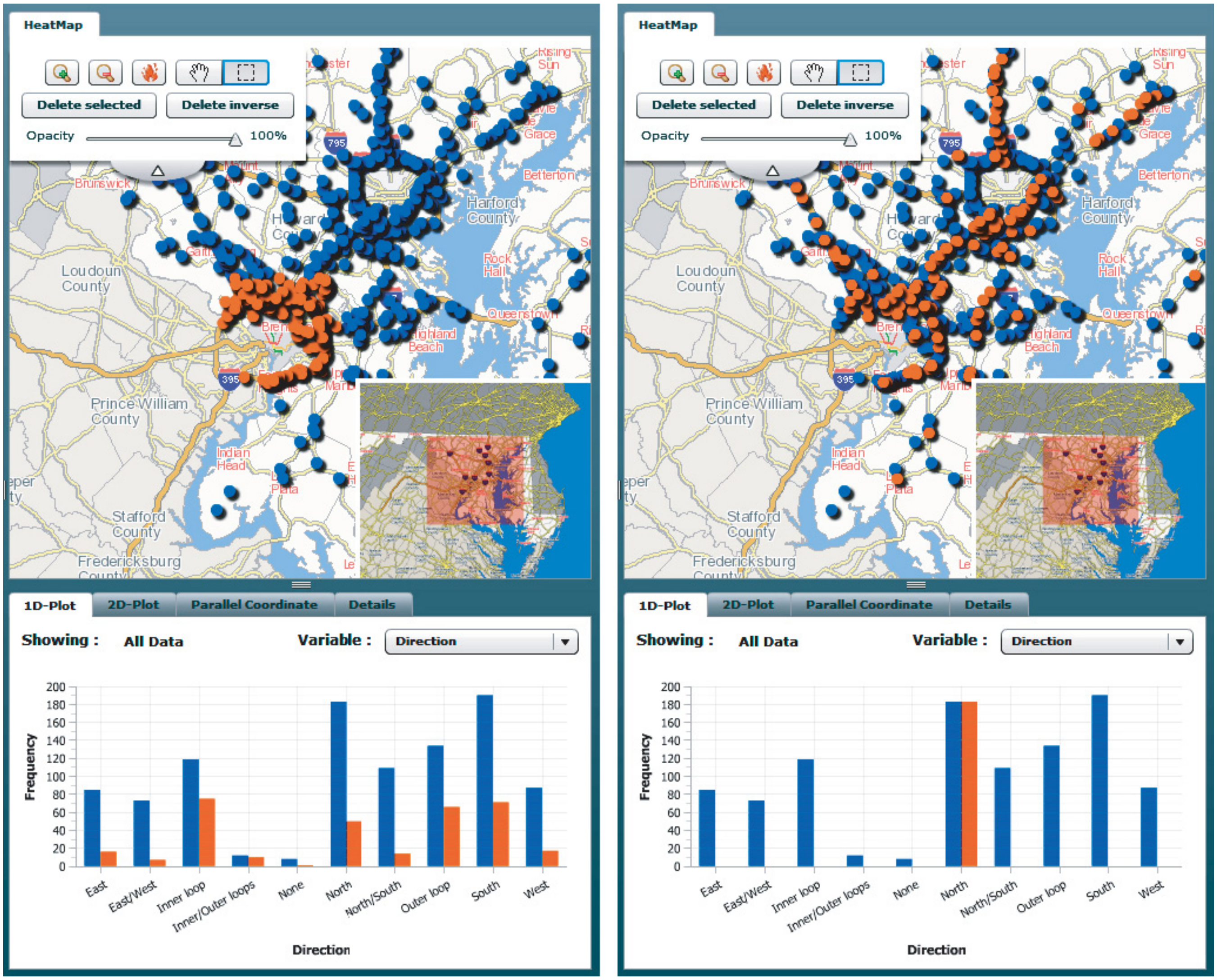
Component Descriptions

The following sections describe in greater detail some of the individual components of Fervor. Although each component is described separately, any interaction with one component ultimately affects the other component windows, as shown in Figure 2.

Mapping

Each accident is plotted on a built-in map, taking advantage of the significance of the geospatial nature of the data. The primary method for exploring regions of interest is through zooming and panning. Figure 3 depicts the two modes available: an icon mode and a heat mode. Users can switch between the modes using the controls along the top of the map. In icon mode, points are rendered as colored dots. Clicking on a dot brings up a details window displaying specific details for that particular accident.

Although rendering each incident as an icon is reasonable in certain domains, it introduces several problems when dense regions need to be analyzed. Occlusion is a major issue when dealing with nearby icons, especially when the map is at the farthest zoom levels. A region with only a few points can look like it has the same number of points as a dense region because it is difficult—if not impossible—to count overlapping points. The heat mode solves this problem by assigning each point a sphere of influence that dissipates when moving away from the point. Spheres of influence have an additive effect on each other, and each pixel is colored based on the



(a) (b)

FIGURE 2 Interactivity of components: (*a*) selecting a cluster of points on map highlights those data points on the histogram and (*b*) highlighting an area on histogram highlights corresponding data points on map.

total influence it receives. Regions with the most influence will be brightly colored in reds while less interesting areas will be assigned lower wavelength colors such as blue. Figure 3 shows two maps with identical data sets rendered in icon mode (left) and heat mode (right). Because of occlusion in icon mode, it is difficult to tell which of the four circled regions has the most points. With heat mapping enabled, it is clear that the third region from the left has more points than the other three regions.

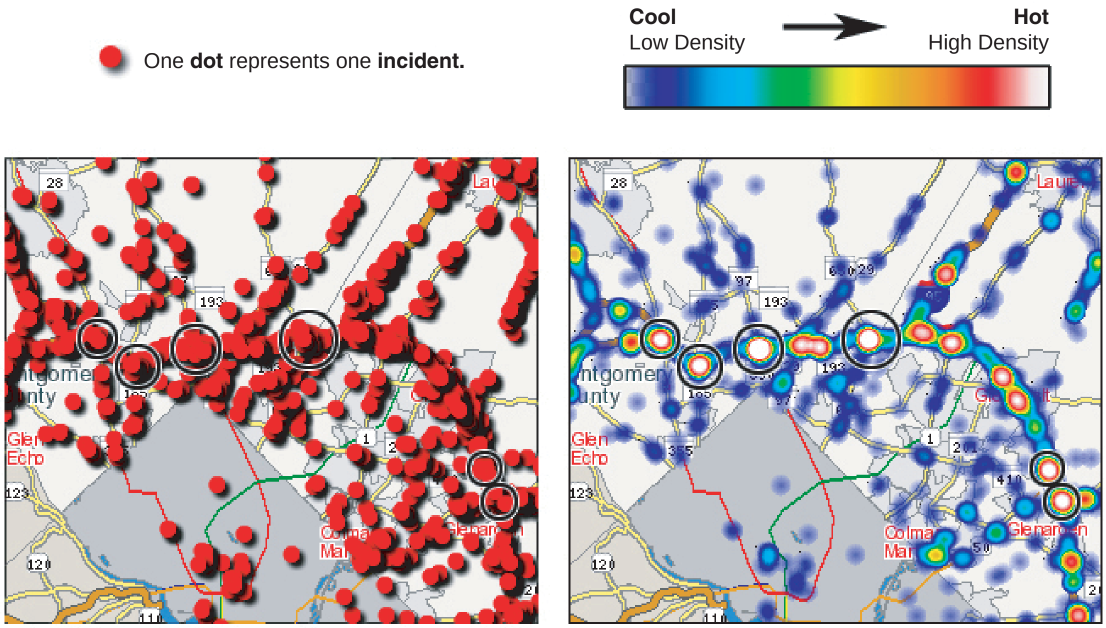
The heat map algorithm used in Fervor is based on open-source work by Corunet (*21*), which uses heat maps to analyze web page clicks. The heat map gradient used by Fervor transitions from low density to high through blue, cyan, green, yellow, orange, red, and white. These colors were selected because their order is associated with increasing temperature, and the relatively large number of distinct colors allows for the user to differentiate slight variations in density without much effort. The heat map uses highly saturated colors because they contrast well with the relatively pale background colors of the underlying map.

Ranking

Because of the high dimensionality of transportation incident data, Fervor adopts the idea of a rank-by-feature framework from the HCE (*19*, *20*) in Fervor's ranking panel to help users find interesting one-dimensional (1-D) and two-dimensional (2-D) distributions. Users are allowed to select a ranking criterion and then sort the results (Figure 4). Selecting a variable or a pair of variables from the ranking panel will show the corresponding histogram or 2-D plot in the lower right window of the application, as shown in Figure 1.

Although HCE illustrates many ranking criteria for 1-D and 2-D distributions, it is primarily focused on numerical variables. Transportation incident data sets consist of three types of variables:

1. Numerical (N) variables [e.g., latitude (–77.124323), number of vehicles involved (0, 1, 2, . . . , *n*)],
2. Date–time (D) variables [e.g., created date (12 Jan 2005), created time (10:45 a.m.)], and



(a) (b)

FIGURE 3 Two Fervor mapping modes: (*a*) icon and (*b*) heat.

3. Categorical (C) variables [e.g., road condition (dry, rain, snow, etc.), direction (north, south, east, west, inner loop, etc.)].

Most variables are of Type C, which presents a challenge because the existing HCE ranking criteria do not provide meaningful results when applied to variables of Type C for both 1-D and 2-D rankings.

One approach to deal with this challenge is to convert categorical variables into ordinal variables. Ordinal variables are categorical variables that can be ordered by some criteria. By ordering the possible values of a categorical variable, a numerical value can be obtained—the rank of the value. Consider an example in which road condition (dry, rain, snow) can be ranked by the relative risk of driving in that condition. In this example, dry would be safest, whereas a rainy road would be more slippery, but not as much as when it snows. Thus, one may assign increasing numerical values 1, 2, 3 to represent each road condition. Fervor developers decided to give the user the ability to order categorical data numerically, which ultimately allows for standard HCE methods to be used.

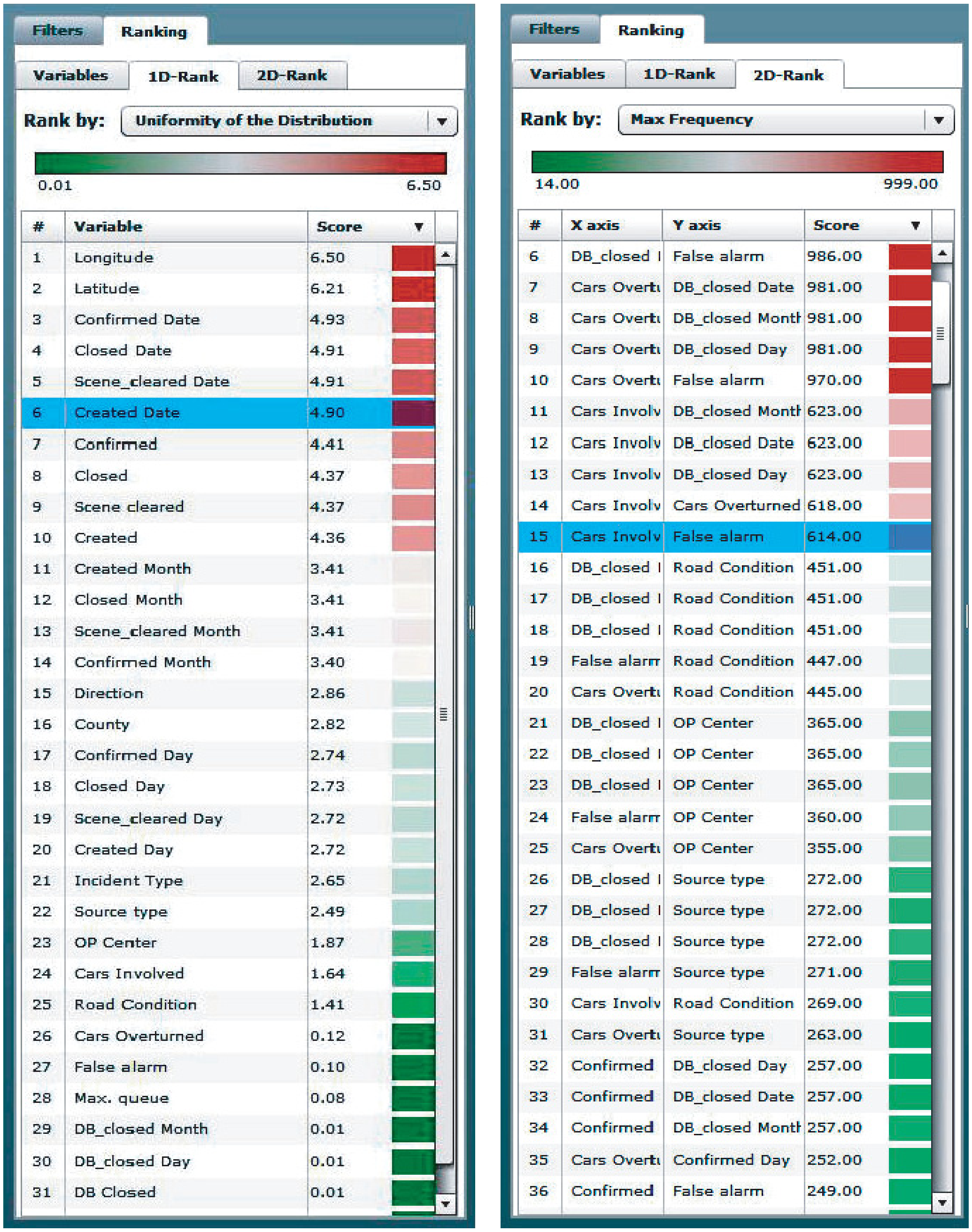
Furthermore, in addition to the five existing 1-D ranking criteria from HCE, Fervor adds the five new criteria presented in Table 1 to rank 1-D distributions for categorical variables. These criteria are based on frequency (F_k), the number of incidents in each category in each variable. Let X be a categorical variable. X has n possible values, X_1 to X_n. F_k is the number of incidents in which $X = X_k$. Using F_k, five additional ranking criteria (Table 1) are used to rank 1-D distributions for categorical variables. Once a 1-D ranking criterion is chosen, the user can click on any of the resulting variables to automatically generate a histogram that then appears in the lower right section of the application.

The existing HCE work presents seven 2-D ranking criteria, but only one ranking criterion for 2-D C-C relationships—a contingency coefficient based on the chi-square test. However, the five 1-D ranking criteria proposed earlier can also be adapted to rank 2-D C-C relationships. These additions improve the flexibility of the rank-by-feature framework, providing more options to users. Instead of using F_k, the frequency now become F_{ij}, the number of incidents in each value pair in each relationship. Let X and Y be categorical variables. X has n possible values, X_1 to X_n, and Y has m possible values, Y_1 to Y_m. F_{ij} is the number of incidents in which $X = X_i$ and $Y = Y_j$. Using F_{ij}, the five ranking criteria in Table 1 can now be used to rank 2-D C-C relationships. Once a 2-D ranking criterion is chosen, the user can then click on any of the resulting variable pairs to automatically generate a number of interactive plots that appear in the lower right section of the application.

Histograms

The histogram panel can be accessed from the 1-D distribution window and shows variations and clusters in the number of incidents for each variable. For example, one can examine the incident frequencies in a monthly, weekly, daily, or hourly fashion. The histogram panel is highly interactive in several ways. First, clicking on certain types of histograms highlights incidents associated with those values in the map. Additionally, temporal data are represented as a hierarchy of values (minutes, hours, days, months, years), and temporal histograms in Fervor support a logical way for interacting with this hierarchy. The initial view displays the correlation between incident numbers and the month of the year in which those incidents occurred (Figure 5*a*). Control-clicking on a particular month, January 2004 in this case, zooms in via a drill-down animation effect (Figure 5*b*) and displays the information by day of the month (Figure 5*c*). Users can further zoom in to view the number of incidents spread over a 24-h period. As seen in Figure 5, the zooming process is animated to prevent users from becoming disoriented by rapid changes in the histograms.

Users also have the capability to zoom back out at any time and select a different temporal category on which to focus. Because temporal data can support multiple levels of zooming, depending on the varying granularity of temporal categorization, users need a simple



(a) (b)

FIGURE 4 Ranking: (*a*) 1-D and (*b*) 2-D.

way to keep track of what zoom level they are at as well as an easy way to navigate back and forth through their trail of zooming. For this purpose, Fervor's histograms use a "breadcrumb trail," which is a horizontal list of labels used for keeping track of a location within a series of views. Clicking on any labeled category in the breadcrumb trail will zoom back directly to that particular zoom level.

2-D Plot

Traditional scatter plots can visualize the relationship between two variables by drawing elements on (x, y) coordinates. Fervor adopts this idea and combines it with the idea of using a color grid—colored tiles where color shades represent values. Placing the mouse over a colored tile or circle will show more information about the item. For every 2-D plot, users can choose between two modes: scatter plot or grid mode. Figure 6 shows the exact same 2-D plot in both scatter plot mode and grid mode.

In the scatter plot mode, incidents are represented by circles. Circles are drawn on (x, y) coordinates according to selected variables. The size of each circle represents frequency of occurrence. Big circles will have higher frequencies than smaller circles. Placing the mouse over a circle brings up additional details about the data and highlights those specific incident records on the map. Figures 7*a* and 7*b* show two examples of 2-D scatter plots.

TABLE 1 1-D and 2-D Categorical Ranking Criteria

Ranking	Description	Formula
Maximum frequency (0 to n)	Ranking by maximum frequency (F) gives priorities to histograms or scatter plots that contain at least one spot that has high frequency. This allows users to easily identify dense areas in the dataset, a key theme for Fervor.	Maximum(F)
Number of potential outliers (0 to n)	Fervor's outlier detection is based on Z-scores. Z-score is calculated from frequency (F). This potential outlier detection defines a value of a variable as an outlier if the calculated Z-score is lower than −1.5 or higher than 1.5.	Count($Z(F) > 1.5$ or $Z(F) < -1.5$)
Percentage of empty area (0 to 100)	The criterion is to sort scatter plots or histograms in terms of percentage which frequencies (F) are zero. A higher percentage suggests that the distributions are more pruned. Users can use this criterion to easily isolate sparse distributions from dense distributions.	Count($F > 0$)/Count(all F) $*$ 100
Number of existing values (0 to n)	One distribution with four out of 10 possible values and another one with 25 out of 100 will have the same percentage of empty area. Hence, this criterion is presented to distinguish them. It is the number of frequencies (F) that are not zero. A lower number suggests that there are fewer existing values in the distribution. Distributions that have small number of existing values might spot some interesting points.	Count($F > 0$)
Standard deviation of distinct frequencies (0 to n)	This criterion is calculated from standard deviation of frequencies (F) which are more than zero. It represents how widely spread the frequencies in the distributions are. A high value represents widely spread frequencies while a low value represents more fairly distributed frequencies. Users can easily identify distributions where incidents are not fairly distributed.	Standard deviation ($F > 0$)

NOTE: Symbol F in this table refers to F_k for 1-D ranking and F_{ij} for 2-D ranking.

To provide an alternative view, the grid mode is introduced. Both x and y axes are divided into rectangular bins. Each section of the grid, or bin, is then colored according to the number of incidents that fall into the given area. Dark colors represent high values and light colors represent low values. This technique can prevent problems with occlusion that might arise in the scatter plot mode, as seen in Figure 6. Figures 7*c* and 7*d* show two examples of 2-D grid plots.

Parallel Coordinates Plot

The parallel coordinates plot (Figure 8) is a way to visualize multivariate data on a plane. Each plot can have as many vertical axes as the data set has variables. Fervor allows the user to specify which variables to include within the plot. A single incident record is represented as a line that goes across all axes from left to right; on each axis, the line goes through the point corresponding to the record's value for that field.

Using the parallel coordinates plot, it is easy to see clusters in the data. If many lines converge at one point on any given axis, then many incident records have that particular value in common for that particular variable. However, if there are only a few lines going through other points on the axis, the values corresponding to these lines are outliers in the original data set. Examining both outliers and clusters may open new insights into the data.

As with the icon mode in the Fervor map, overcrowding may become an issue with large data sets. Therefore, Fervor gives users control over both line thickness and transparency. Additionally, heat maps have been added along the axes to identify the clusters of lines. The heat map overlays can give a clear idea about the distribution of the values for any particular axis as seen in Figure 8*b*. The radius of the heat maps may also be adjusted by the user to prevent occlusion resulting from close data points (Figure 8*c*).

The parallel coordinates plot in Fervor can display both numerical and categorical fields as well as dates and time stamps. As with the other components in Fervor, "brushing and linking" is used to highlight selected items on the map when they are selected in the parallel coordinates window and vice versa. For example, if the parallel coordinates plot shows a high concentration of converging lines in a particular axis, it would be reasonable to assume those accidents are somehow correlated with that property. Selecting those accidents on the plot will highlight the same accidents in all the other views including the map, histogram, and 2-D plot.

EVALUATION

Although a formal usability analysis study has yet to be conducted on the Fervor application, several transportation professionals have experimented with the tool and have provided valuable feedback. Each user was given a short but detailed overview of the tool that explained the proposed workflow and introduced the features. Each evaluator was given time to experiment with the tool and comment on performance, usability, and applicability. One reviewer commented that although heat maps represent the overall distribution of incidents in the region quite well, the heat maps themselves could be misleading because (*a*) they only represent straight counts of incidents rather than rates of incidents based on road size and volumes, and (*b*) they aggregate accidents on adjoining roads together because of their proximity.

All evaluators were impressed with the supporting analytical components that the tool provided and were particularly intrigued by the 2-D and parallel coordinates plots because they provided a systematic way to explore relationships within data. However, users had a tendency to spend more time using features they were more familiar with or were more commonly found in their existing workflows, such as the map and histogram. Positive comments were also received on the many "brushing and linking" features that tied each of the three main panels of the application together, making it easier to explore relationships and find patterns. Users highly appreciated

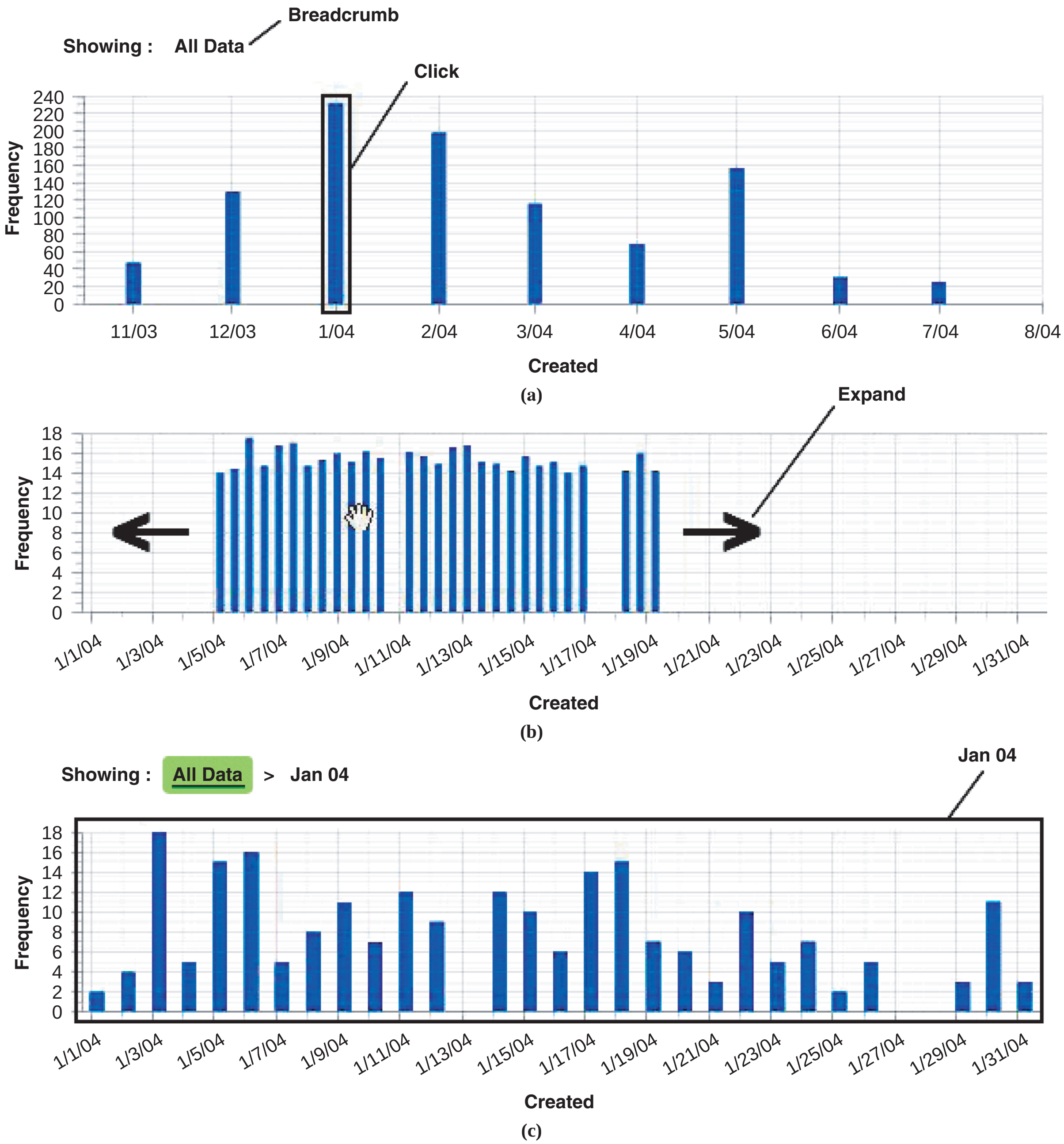


FIGURE 5 Semantic zooming on histograms.

that the application was extremely user friendly and that the interface was well thought-out, required almost no training, and was web-accessible.

FUTURE WORK

Given the feedback on heat maps, developers are working to improve the heat map algorithms so that the user can have control of the heat map granularity and can specify color schemes and breakpoints. Although the application is currently running on incident records from the Maryland CHART program, it has been developed to be easily extended to other data sets with minimal programming. However, a simple data import tool allowing other data sets to be automatically integrated into the application is desired. Fervor could also be adapted to work on multivariate data sets that are not transportation related.

CONCLUSIONS

This paper presented Fervor—a lightweight, web-based application that aims to facilitate visual analytics for transportation incident data sets. Fervor affords sophisticated yet user-friendly analysis of transportation incident data sets. Interactive map, histogram, two-dimensional plot, and parallel coordinates plot are featured

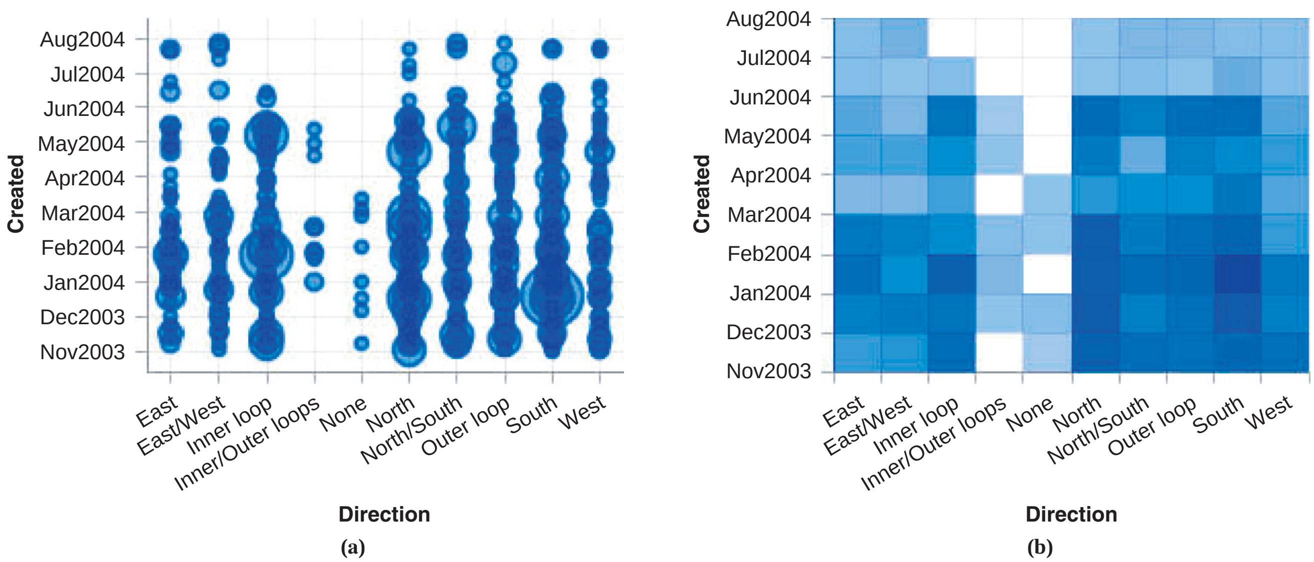


FIGURE 6 Two-dimensional plot features two modes: (*a*) scatter plot and (*b*) grid.

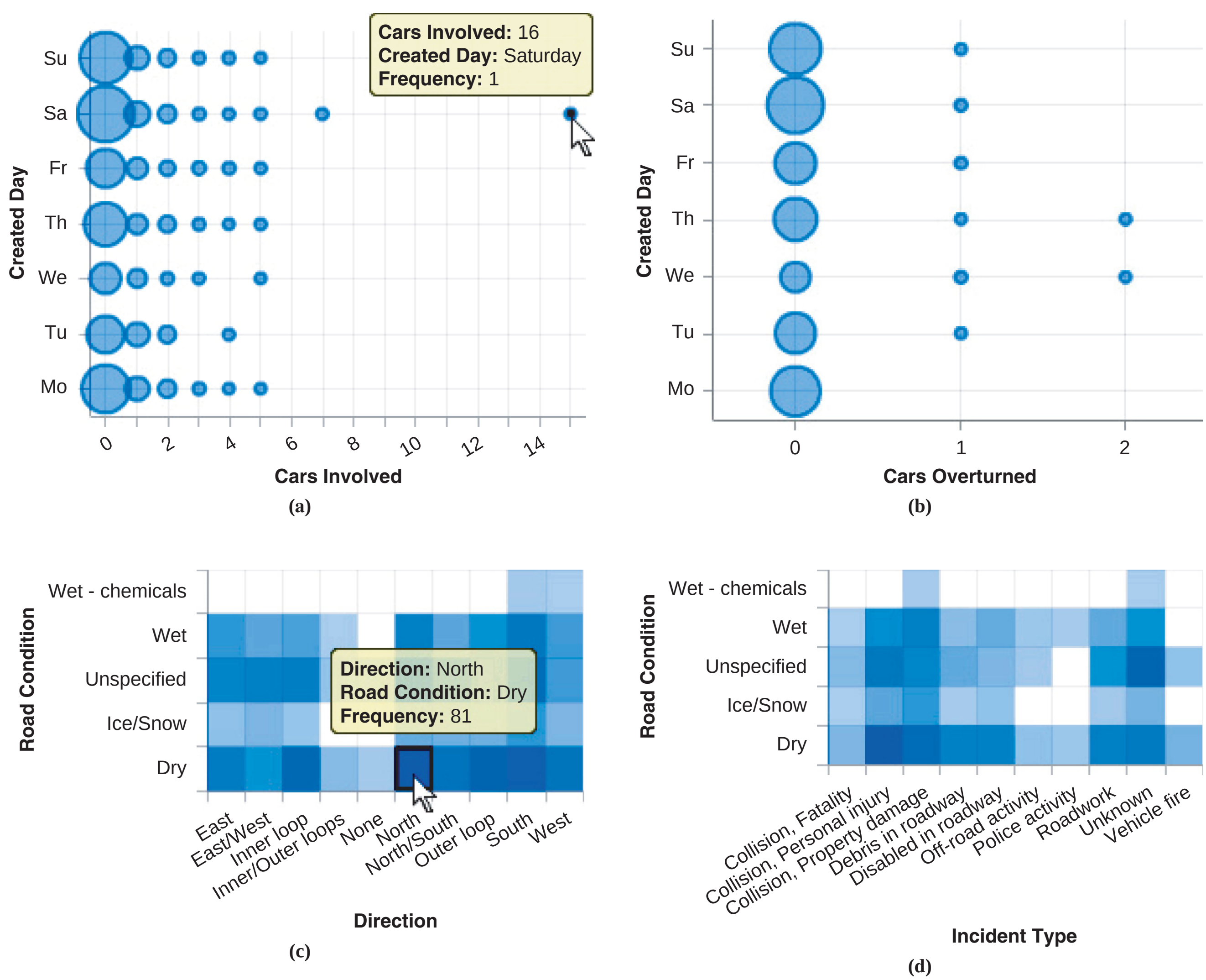


FIGURE 7 Examples of 2-D scatter and grid plots.

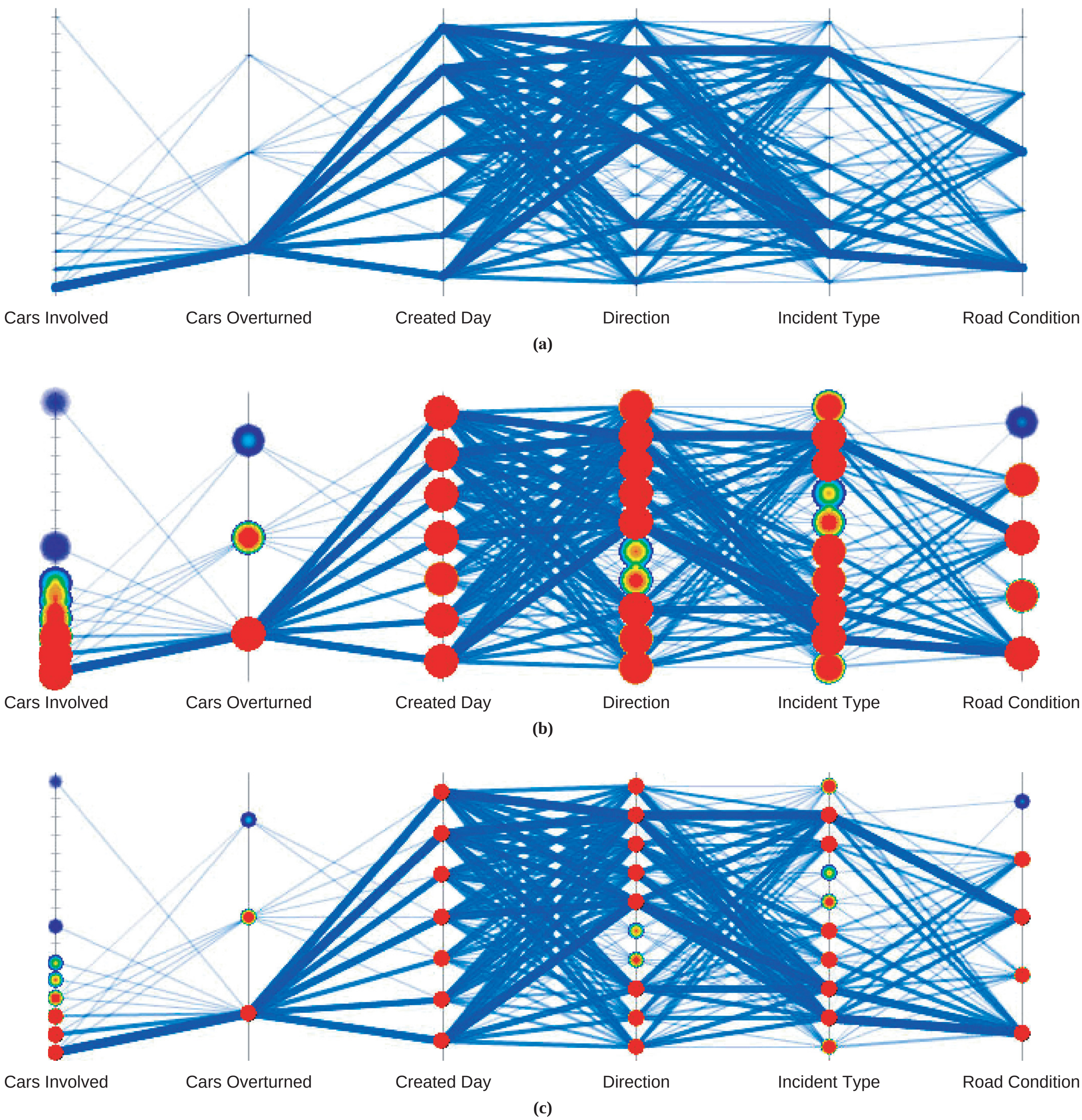


FIGURE 8 Parallel coordinates plot examples.

visualizations that are integrated together to allow users to simultaneously interact with and see relationships among multiple visualizations. Using a rich set of filters, users can create custom conditions to filter data and focus on a smaller data set. Because of the multivariate nature of the data, a rank-by-feature framework has been adopted and further expanded to quantify the strength of relationships among the different fields describing the data. It is hoped that the Fervor application will allow transportation professionals to spend less time and energy worrying about the pure mechanics and economics of asking questions of the data, and afford them the opportunities to determine which questions to ask, seek out new answers, and derive knowledge from these vast data sources, which will ultimately help to make the transportation system safer.

ACKNOWLEDGMENTS

The authors thank the Maryland State Highway Administration and the office of CHART for their continued support and for allowing use of their data. Special thanks to Joonghoon Lee and Ken Knudsen for their contributions to Fervor and to Ben Shneiderman and the students in his information visualization class at the University of Maryland for their guidance and feedback throughout the design and development process.

REFERENCES

1. Georgia Department of Transportation. *Crash Analysis, Statistics and Information Notebook.* www.dot.state.ga.us/statistics/CrashData/Pages/casi.aspx. Accessed July 27, 2008.
2. Hammond, P. J. *Measures, Markers and Mileposts.* Quarterly report to governor, legislature, and Washington State Transportation Commission on transportation programs and department management. Washington State Department of Transportation, Olympia, 2008.
3. University of Maryland's National Study Center for Trauma and EMS. *Maryland Traffic Safety Facts 2006.* http://medschool.umaryland.edu/nscfortrauma/traffic.asp. Accessed July 27, 2008.
4. NHTSA, U.S. Department of Transportation. *Maryland—Toll of Motor Vehicle Crashes.* www-nrd.nhtsa.dot.gov/departments/nrd-30/ncsa/STSI/24_MD/2006/24_MD_2006.htm. Accessed July 27, 2008.
5. Office of Highway Safety, Kentucky Transportation Cabinet. KOHS—Traffic Records. http://highwaysafety.ky.gov/trafficrecords_query_county.asp. Accessed July 27, 2008.
6. FMCSA, U.S. Department of Transportation. Analysis and Information Online. http://ai.volpe.dot.gov/mcspa.asp. Accessed July 27, 2008.
7. NHTSA, U.S. Department of Transportation. Fatality Analysis Reporting System Encyclopedia. www-fars.nhtsa.dot.gov/Main/index.aspx. Accessed July 27, 2008.
8. Welch, T. AASHTO Safety Data Systems Task Group. Presented at AASHTO Standing Committee on Highway Traffic Safety Spring Meeting, Branson, Missouri, 2007.
9. Shneiderman, B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *Proc., IEEE Symposium on Visual Languages,* 1996.
10. ArcGIS. www.esri.com/software/arcgis/. Accessed July 18, 2008.
11. Andrienko, N., and G. Andrienko. Interactive Visual Tools to Explore Spatio-Temporal Variation. *AVI 2004: Proc., Working Conference on Advanced Visual Interfaces,* Association for Computing Machinery, New York, 2004.
12. MacEachren, A., X. Dai, F. Hardisty, D. Guo, and G. Lengerich. Exploring High-D Spaces with Multiform Matrices and Small Multiples. *IEEE Proc., International Symposium on Information Visualization,* 2003.
13. Spotfire. www.spotfire.com. Accessed April 20, 2008.
14. Tableau. www.tableausoftware.com. Accessed April 20, 2008.
15. Dykes J., A. Slingsby, and K. Clarke. Interactive Visual Exploration of a Large Spatio-Temporal Dataset: Reflections on a Geovisualization Mashup. *IEEE Transactions on Visualization and Computer Graphics,* Vol. 13, No. 6, 2007.
16. Google. Google Earth API. http://code.google.com/apis/earth/. Accessed July 18, 2008.
17. Fisher, D. Hotmap: Looking at Geographic Attention. *IEEE Transactions on Visualization and Computer Graphics,* Vol. 13, No. 6, 2007.
18. Universal Mind. *Universal Mind: Demos.* www.universalmind.com/demo/launchpad.cfm. Accessed May 5, 2008.
19. Seo, J., and B. Shneiderman. A Rank-by-Feature Framework for Interactive Exploration of Multidimensional Data. *Information Visualization,* Vol. 4, No. 2, 2005.
20. Seo, J., and H. Gordish-Dressman. Exploratory Data Analysis with Categorical Variables: An Improved Rank-by-Feature Framework and a Case Study. *International Journal of Human–Computer Interaction,* Vol. 23, No. 3, 2007.
21. Corunet. How to Make Heat Maps. http://blog.corunet.com/english/how-to-make-heat-maps. Accessed April 20, 2008.

The Visualization in Transportation Committee sponsored publication of this paper.