Global terrorism visualization

Caroline Ziemkiewicz, Xiaoyu Wang, Alex Godwin, Wenwen Dou, Remco Chang, Robert Kosara, William Ribarsky
Southeast Visualization and Analytics Center (SRVAC), University of North Carolina at Charlotte

Introduction

We present a system for the visual analysis of terrorist activity based on the Global Terrorism Database (GTD), developed by the National Consortium for the Study of Terrorism and Responses to Terrorism (START). This dataset is based on records obtained from Pinkerton Global Intelligence Services and from START's own analysis of news stories and government reports. It contains information on over 65,000 incidents taking place from 1970 to 1997, and will be updated over the next year to include incidents up to 2007.

The goal of the GTD project is to make data on terrorism open-source and available to outside researchers. However, given the size and complexity of the data, this openness is of limited use without tools for accessing and understanding the information within. To assist in this urgent effort, we have developed an interconnected visual analysis system to provide meaningful views on the data and support the investigative process.

Our goal is to allow investigators to take a global, strategic view of patterns in terrorist activity. Ultimately, users must be able to form hypotheses through exploration, investigate the likelihood of those hypotheses in the larger context of global behavior, and make actionable projections of future behavior based on this information.

Approach

As stated, the three major goals of our system are exploration, investigation, and projection. Our technique for addressing this problem is a three-part system, with each part emphasizing one of the three goals (**Fig. 1**). Abstract exploration of the dimensions in the dataset can be performed in the Parallel Sets view; investigation of hypotheses in a global context is possible in the Investigative view; and projection based on temporal patterns is supported by the Entity Relationships view.

While each view is specialized for a key process of investigative analysis, it is only through their interactions that the system as a whole can support the reasoning process at a strategic rather than a tactical level.

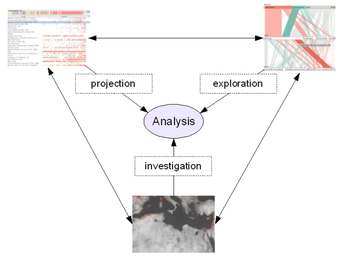

Fig. 1. The global terrorism visualization system design. Exploration of dimensional relationships in the Parallel Sets view suggests hypotheses and interesting outliers. Examining temporal patterns in the Entity Relationships view leads to projections of future behavior. Both hypotheses and predictions can be investigated in a global context in the Investigative view.

System Description

Our system consists of three closely interlinked views that offer complementary perspectives on the data. The Investigative view (**Fig. 2**) presents the data in geotemporal context, organized by group name so that patterns can be viewed on an entity-by-entity basis. This view provides the primary context for investigation by grounding events and entities in time and space, and by providing tools to organize analysis and drill down to the original data. These tools are built around the investigative questions of *who*, *what*, *when*, and *where*, allowing the user to study their hypotheses in relation to these essential questions.

The Parallel Sets view (**Fig. 3**) supports the exploration of relationships among dimensions in a more abstract manner. Parallel sets is a method for showing degree of correlation between categorical dimensions using "ribbons" that are sized based on the number of cases two categories share. By extending this method to include numerical and temporal data as well, we can show trends of correlation across an arbitrary set of dimensions, making this a generalized view in which a user can find interesting trends and outliers through highly flexible interaction.

To further investigate temporal patterns, we have developed a novel view of entity relationships over time based on longest common subsequence (LCS) analysis (**Fig. 4**), which finds patterns of characters across two strings while disregarding irrelevant characters in between. We extend this idea to the analysis of temporal patterns of behavior. Our Entity Relationships view also provides a search-by-example function, which makes it possible to quickly find entities with similar patterns of behavior to an entity of interest. This makes it possible for the user to project a group's behavior onto the activity of similar groups, suggesting possible hypotheses about future changes in behavior.

Communication among the three views allows the investigator to examine a question or hypothesis from several different angles. Changes made in any one view can be reflected in the others, so that incidents are filtered throughout and relationships are recalculated in the Parallel Sets and Entity Relationships views. This creates a high degree of flexibility in supporting the investigative process, and allows the user to explore, investigate, and project hypotheses in any sequence while maintaining a sense of the overall state of their analysis.

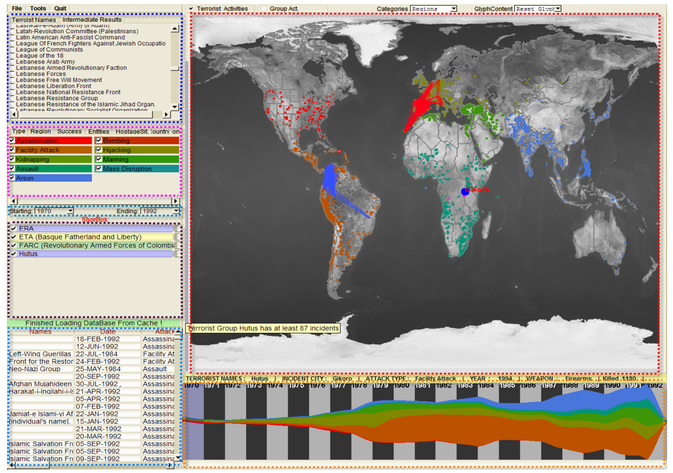

Fig. 2. The Investigative view lets the user analyze the *who*, *what*, *when*, and *where* of incident patterns. In this view, terrorist incidents are organized by the responsible group (*who*) in geographical context (*where*) and along a timeline (*when*), with points colored or filtered by details such as incident type, weapons used, and type of target (*what*). Points on the map represent individual incidents, and lines between points indicate the movements of a single terrorist group. Through interactive exploration of these aspects of incident groups, coupled with the added information from the other views, the user can investigate hypotheses about *why* incidents take place.

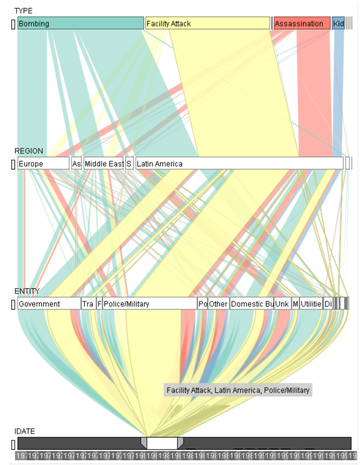

Fig. 3. In the Parallel Sets view, an investigator can explore the data in an abstract way by examining correlations across multiple dimensions. The ribbons between two categorical dimensions show the proportion of cases from one category (e.g., facility attacks) that fall under a category from the second dimension (e.g., Latin America). This view can also show how cases are distributed across numerical and temporal dimensions. Highly interactive filtering and view-building tools support freeform exploration of data. Here we see the high rate of incidents in Latin America during the early 1980s, and how those incidents break down in terms of type and target.

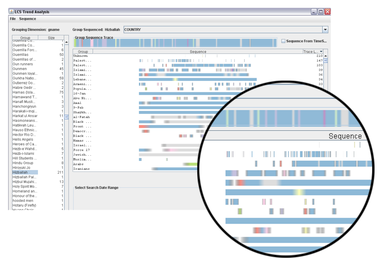

Fig. 4. The Entity Relationships view shows common sequences of event types between two groups. Each long colored bar represents a sequence of events associated with a given group, and each color represents a given value in a categorical variable, such as a country where the attack took place. Groups are ranked by the size of their longest common subsequence with the group of interest (in this case, Hizballah), and the visual depiction of event sequences allows the investigator to see exactly what values the two groups share and what segment of time is represented in the LCS analysis. This makes the LCS process more transparent and more useful for hypothesis building than a mere output of sequence length.

Conclusions and Future Work

By providing a range of visualization tools that address three major facets of the analysis process, we hope to break open the large amounts of data on international terrorism in a way that aids and does not overwhelm investigators. In the near future, we will be deploying our system to several teams of investigators, including social scientists, criminologists, and fellow researchers. This will offer the opportunity to evaluate our approach and further fine-tune the design process.

We have already performed an informal evaluation of the investigative view alone with experts from law enforcement and criminology, and have begun to integrate the feedback from this evaluation into further development of our system. This expert feedback has already indicated the possibility of extending our tools beyond visualization of terrorism data into exploration of other kinds of criminal activity.

We also plan to deploy a version of these tools to the website of the GTD project. This will greatly assist in the START Center's goal of making these data open source and accessible, facilitating their use by investigative analysts, academic researchers, and members of the public. Although the GTD is an invaluable source, the scope of this project is not limited to visualizing a single dataset. We are also collaborating with researchers at the Pacific Northwest National Laboratory (PNNL) to apply these tools to the content analysis of news reports and other unstructured text.

By focusing on the basic properties of investigation and focusing our tools in support of different subgoals, we have produced a generalizable approach to visually supporting the investigative process.

Literature cited

Godwin, A., Chang, R., Kosara, R., and Ribarsky, W. 2008. Visual analysis of entity relationships in Global Terrorism Database. In *SPIE Defense and Security Symposium*. To Appear.

Jones, J., Chang, R., Butkiewicz, T., and Ribarsky, W. 2008. Visualizing uncertainty for geographical information in Global Terrorism Database. In *SPIE Defense and Security Symposium*. To Appear.

Kosara, R., Bendix, F., and Hauser, H. 2006. Parallel sets: visual analysis of categorical data. *IEEE Transactions on Visualization and Computer Graphics* 12(4):558-564.

Wang, X., Miller, E., Smarick, K., Ribarsky, W., and Chang, R. 2008. Investigative visual analysis of Global Terrorism Database. In *Journal of Computer Graphics Forum Eurovis*. Conditionally accepted.

Acknowledgments

This project was funded through the Southeast Regional Visualization and Analytics Center program from the Department of Homeland Security, Science and Technology Directorate, Office of University Programs. We would like to thank our collaborators at the START Center and PNNL, as well as attendees of the Financial Visual Analytics Workshop, for their valuable feedback.

For further information

Please contact *caziemki@uncc.edu*. More information on this and related projects can be obtained at *srvac.uncc.edu*.