

Investigative Visual Analysis of Global Terrorism

paper ID 1185

Abstract

Recent increases in terrorist activity around the world have made analyzing and understanding such activities more critical than ever. With the help of organizations such as the National Center for the Study of Terrorism and Responses to Terrorism (START), we now have detailed historical information on each terrorist event around the world since 1970. However, due to the size and complexity of the data, identifying terrorists' patterns and trends has been difficult. To better enable investigators in understanding terrorist activities, we propose a visual analytical system that focuses on depicting one of the most fundamental concepts in investigative analysis, the five W's (who, what, where, when, and why). Views in our system are highly correlated, and each represents one of the W's. With this approach, an investigator can interactively explore terrorist activities efficiently and discover reasons of attacks (why) by identifying patterns temporally (when), geo-spatially (where), between multiple terrorist groups (who), and across different methods or modes of attacks (what). By coupling a global perspective with the details gleaned from asking these five questions, the system allows analysts to think both tactically and strategically.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Line and Curve Generation

1. Introduction

The war on terror has taken center stage since the 9/11 attacks. With the establishment of the Department of Homeland Security and corresponding international efforts, more analysts are participating in the investigation of world-wide terrorist activity. These analysts seek to reveal the patterns of activities of different terrorist groups and the relationships among them. To make their investigations complete, analysts need to uncover the facts of terrorist events and incorporate them into a broader context. However, until recently, most reports on terrorism have been scattered across different data sources, making it difficult to build a cohesive picture.

The Global Terrorism Database (GTD) project has consolidated both domestic and international terrorist activities between 1970 and 1997. With the wealth of data represented by the GTD, the challenge now becomes to understand and uncover important patterns and relationships. Unfortunately, analysts until now have had limited tools to help build hypotheses and identify trends, limiting the speed of understanding this important information.

We present an interactive visual analytics system to explore the GTD. Our system is designed around depicting one of the most fundamental concepts in investigative analysis,

the five W's (*who, what, where, when, and why*). Four coordinated views are built to depict each of the first four W's (Figure 1), while the exploratory relation-probing nature of the system allows the analyst to build her own *why* out of the information at hand. We have informally evaluated this approach by presenting our system to three groups of expert investigative analysts with backgrounds in criminology and law enforcement. The results of this evaluation show that the system can assist the analyst in building an integrated understanding of terrorist activities.

The system presented in this paper represents significant improvements over existing practice in the following ways:

- It provides visual analysis of the five W's, which supports existing investigative thinking.
- It supports both strategic analysis of high-level patterns and tactical analysis of individual events.
- It facilitates communication of investigative findings and hypotheses among analysts.

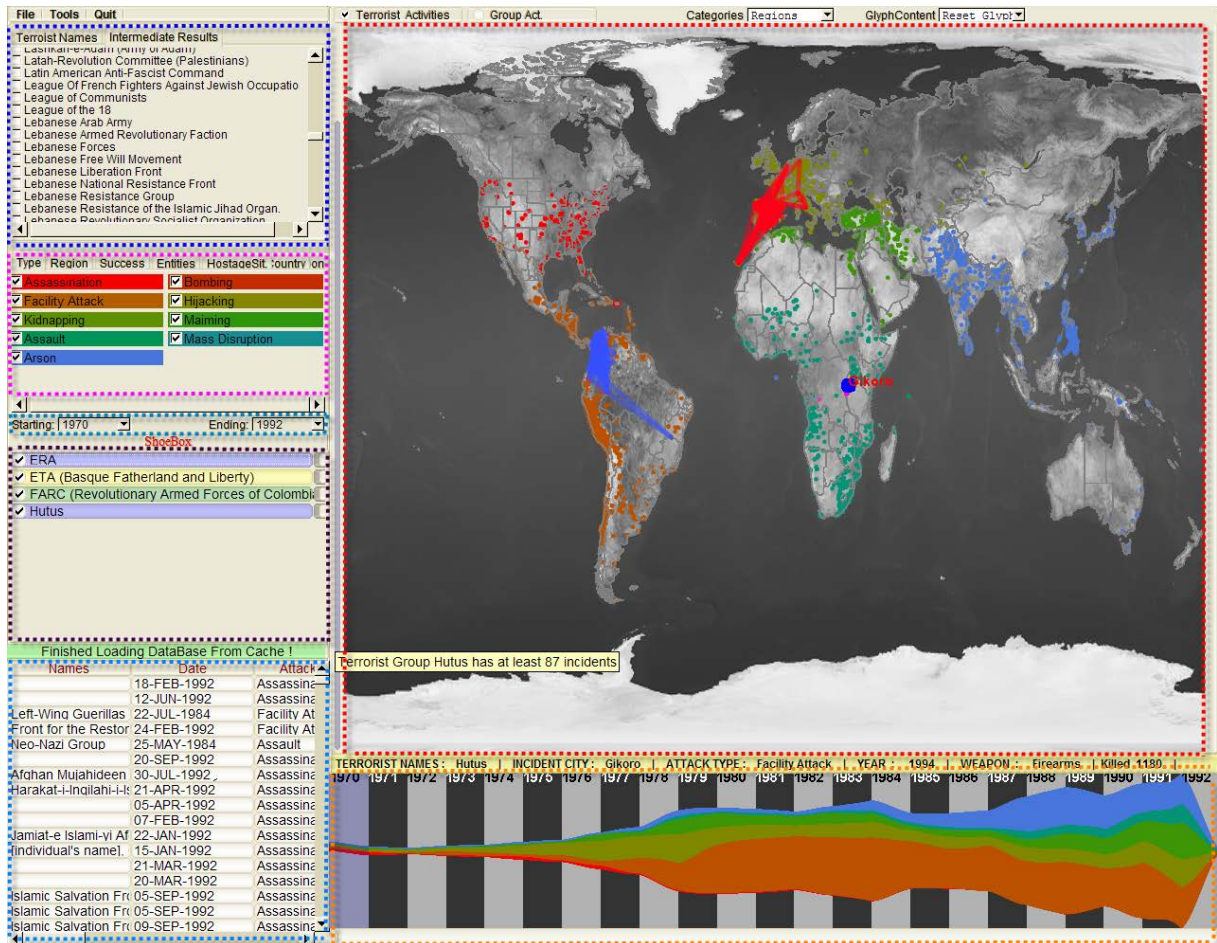


Figure 1: This is the overview of the entire system, including the map view (top right), temporal view (bottom right), entity view (top left) and activities panel (middle left).

2. Related Works

There is an extensive literature devoted to the study and analysis of terrorism. These studies mostly focus on presenting the results of their analyses using qualitative descriptions or simple graphs (such as histograms or line graphs over time) to show trends of one or two variables [GLM07]. In the case of the Global Terrorism Database, which contains more than one hundred dimensions, these descriptions and graphs simply do not sufficiently communicate the complex relationships among all the variables and, in particular, their detailed spatial relations over time. In this situation, it is difficult for the analyst to identify global patterns and trends or to formulate hypotheses and perform high level strategic reasoning.

The use of exploratory visual analytics concepts to understand complex relationships in terrorism activities is still in its infancy. The research up to this point can be divided

into two groups: social network and geo-temporal visualizations. Social network visualization and analysis is a well established area in both the sociology [Fre00] and visualization communities [SMER06], but few systems have been applied to understanding large amounts of terrorist activities. Shen et al. [SMER06] developed OntoVis, which utilized an ontology graph to visualize large heterogeneous networks and applied it to depict relationships between terrorist groups. Perer and Schneiderman [Per06] analyzed the GTD in their SocialAction system and could display relationships between terrorist groups and/or countries. On the other hand, although geo-temporal visualization is also a well established area of research, particularly in the field of geographical information systems, there has been little application of this approach to terrorism data. In fact, the only system that we're aware of that has been used directly in understanding terrorist activities by depicting temporal, geospatial, and

multivariate aspects of terrorism was developed by Guo et al. [GLM07], which does not focus on the relationships between individual entities.

Surprisingly, we have found very few comprehensive systems that attempt to incorporate both the social and the geo-temporal aspects of terrorist activities. A framework proposed by Zhu and Chen in a poster [Zhu07] demonstrated a structure for automatically identifying who, when, what, and where in a salient story regarding terrorist events. While their framework concentrates on extraction of information from unstructured text, our system uses existing terrorism data and focuses on the interactive exploration and analysis of the relationships within the data.

3. Global Terrorism Database

The U.S. Department of Homeland Security (DHS) has as a primary mission preventing terrorist attacks within the United States and reducing the vulnerability of the United States to terrorism. With support from the DHS, a team of researchers at the National Consortium for the Study of Terrorism and Responses to Terrorism (START) has developed a new database designed to help analysts, practitioners, and policy-makers achieve DHS's mission while also providing a resource for the other investigative analysts, such as reporters.

The Global Terrorism Database, or GTD, provides detailed information on terrorist events that have occurred all around the world since 1970, including all domestic cases (e.g., an American attacking a target within the United States, or a Frenchman attacking a target in France) as well as international events (where the perpetrator attacks a target in a foreign country, as with the 9/11 attacks).

Compiled by a team of social scientists, the data have been assembled to allow for systematic analyses of trends in terrorist activity over time, across places, and by different terrorist organizations. While the world may face unique terrorist threats today, analysts can learn a great deal about today's threats by exploring the behaviors of terrorist groups in the distant and recent past and the impacts that these actors, and the events they perpetrated, have had. The GTD is an essential tool to allow for such exploration.

The GTD began in 2001 when Gary LaFree and Laura Dugan at the University of Maryland obtained a large collection of data originally collected by the Pinkerton Global Intelligence Services (PGIS). From 1970 to 1997, Pinkerton trained researchers to identify and record terrorism incidents from wire services, government reports, and major international newspapers. With the launch of START in 2005, funds were made available to conduct validity and reliability checks of the existing PGIS data, and starting in April 2006, in partnership with the Center for Terrorism and Intelligence Studies (CETIS), collection on data for events post-1997 began. By September 2008 the new data will be systematically

integrated with the original data to form a single source of information on terrorist attacks, from 1970 to 2007.

The current database (GTD 1.0) has over 60,000 incidents and 119 dimensions. Over 2000 terrorist groups (*who*) have been recorded in this database, connected with events on a 27-year time line from 1970 to 1997 (*when*). We use a set of well-defined categorical dimensions as the *what* in our system. For instance, we list attack type, target type and weapon type in the *what* view, so that the tactical incident can be analyzed from different aspects. Therefore, by mapping each individual incident to its location (*where*) and showing the rest of the W's, we offer a complete, highly interactive system to assist analysts in researching this critically important dataset (or any other similarly structured data).

4. Investigative Analysts

Since the first public release of the GTD in May 2007, there have been more than 110,000 hits to the web-based GTD interface. Generally speaking there are three types of users: the general public, investigative analysts, and terrorism experts and researchers, including counter-terrorism practitioners. These three groups of users correspond to three levels of use of the GTD in terms of sophistication and depth. The general public represents users who are largely unaware of historic global terrorist activities; investigative analysts are those with some knowledge of terrorist activities; and terrorism experts and researchers have in-depth knowledge of the groups and events, especially in the areas of their expertise. Surprisingly, most queries to START about the GTD comes from investigative analysts who are interested in identifying terrorist events and finding relationships or correlations between terrorist activities that are not previously known. The typical task for an investigative analyst could be to identify patterns of terrorist activities in a specific country over some time, to discover common targets of terrorist groups, or to find out if attacks occur near each other and around the same time, indicating the possibility of collaboration or coordination among the responsible groups. It is with the needs of investigative analysts in mind that we design our system. The aim is to create the system such that it is intuitive for analysts and yet powerful enough to support the investigative process by answering the five W's. In this way and because its analysis can be both general and highly detailed, it can also support the needs of expert users and researchers.

5. System Overview

In order to allow investigative analysts to freely explore terrorist activities and discover new trends and relationships, our system uses a different pane to depict each of the W's in a highly coordinated manner such that interaction with one of the panes will immediately affect the views and results of the others. This multi-view approach is flexible in nature and allows an analyst to inject knowledge of any of the four W's

in any order. As more information is given to the system, it produces more precise and detailed the results become, allowing the analyst to reduce irrelevant information and focus on the desired incidents or relationships.

To support the interaction of *where*, we create a zoomable map view that locates each terrorist incident globally. An interactive ThemeRiver [HHWN02] shows the trend of a user-chosen variable over time, depicting *when*. A list of checkboxes organized by dimensions allows the user to select the attributes of the incidents (*what*), and an updating panel details the terrorist groups that are of interest to the analyst (*who*). We further create a shoebox pane that serves as a container for evidence collection. Terrorist groups in the shoebox can be selected and compared to identify similar trends and patterns. Lastly, a simple tabular pane displays the original incident record in plain text, which allows the analyst to drill down to the details of each incident at any time during the investigation.

Where: Map View

Maps are arguably one of the easiest and most intuitive visualizations to understand [Pla05], so the *where* view is the centerpiece of our system to visualize entities and their relationships. Since the majority of the incidents in the GTD contain geographical information of varying degrees of accuracy, our map is zoomable and pan-able to enable investigation at any granularity. By placing each incident on the map, detailed geographic relationships and patterns immediately become apparent. Figure 3 shows the attacks in Egypt are all along the Nile River, while most attacks in northern Africa take took place along the coast of the Mediterranean Sea. The incidents can be colored based on categorical dimensions of the data such as attack method or weapon type (Figure 3) that can easily reveal patterns based on locations and trends over time. Display of numerical data such as number of casualties is shown using transparent circles (Figure 4), with the size of the circle corresponding to the value for each incident. In this way multiple relations with precise geographic patterns and distributions over time can be built up and understood.

One important feature of the map view is depicting the temporal sequence of activities of the terrorist groups by connecting the incidents with edges (Figure 4). Although the obvious use of this visualization is understanding the sequence in which incidents occur, the greatest benefit of connecting the incidents is to give a *shape* to each terrorist group. By relying on humans' ability to detect and compare different shapes, this visualization makes comparing between different groups intuitive. Furthermore, as shown in the evaluation section (section 7.2), using these kind of shapes makes detecting certain errors and outliers straightforward and obvious (Figure 4).

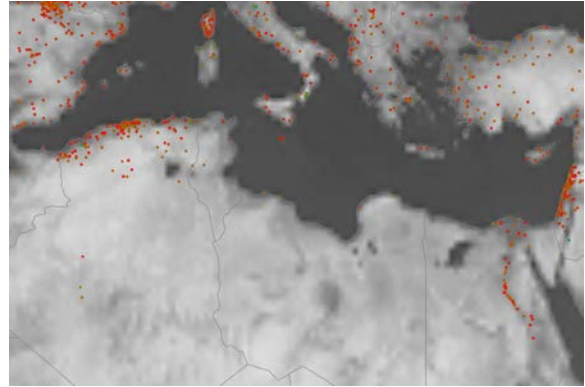


Figure 3: Global distribution of incidents are shown colored by attack type. The locations of the incidents show that attacks in North Africa occur largely along the coast and the Nile River.

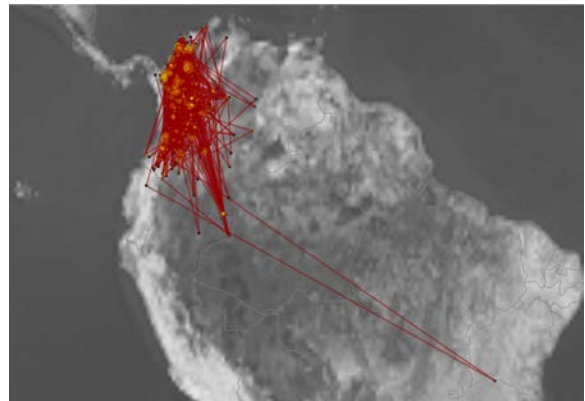


Figure 4: FARC's (Revolutionary Armed Forces of Colombia) activities are connected by lines to indicate the temporal sequence of events. The yellow circles represent the number of casualties per incident. The overall shape indicates that FARC is mainly a domestic terrorist group, but has one significant geographical outlier.

When: Temporal View

Understanding temporal patterns and trends in terrorist activities is the *when* aspect of an investigation. Our method for interacting with time, based on ThemeRiver [HHWN02], shows time in relation to categorical dimensions in the terrorist events. The streams in our ThemeRiver correspond to values in a categorical dimension that is interactively selected by the analyst. The ThemeRiver reveals global temporal trends and patterns, as well as the relative growth and decline among the streams over time, making it much more knowledge-rich than simple time slider. As an example, Fig-

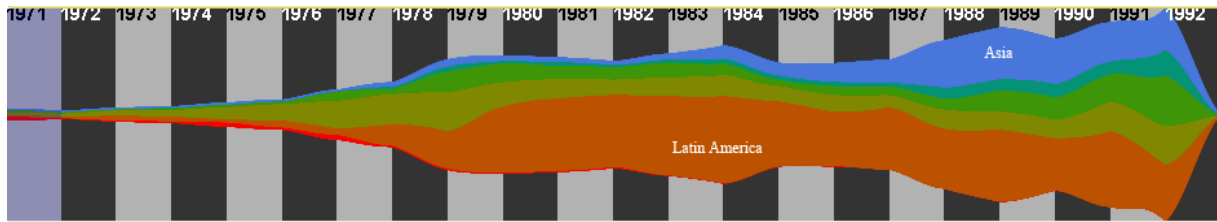


Figure 2: Regional ThemeRiver: this view contains both dimensional information and temporal information, indicating global terrorism trends from 1971 to 1992. Here it shows regional distributions, as labeled.

ure 2 shows the overall increase of terrorism around the world as well as the growth of terrorism in each of the continents. Specifically, the rise of terrorism in Latin America (brown) in the early 80s is evident, as is the rapid growth of activities in Asia in the late 80s (light blue).

What: Activities Panel

To help organize over a hundred dimensions in the GTD dataset, we use a panel with tabs that correspond directly to the categorical dimensions in the data. Each tab represents a dimension, and the possible values for the dimensions are shown in different colors and unique labels in the form of checkboxes (Figure 5). The combination of these checkboxes denotes the *what* aspect of the terrorist incidents. Interacting with these categories by checking/un-checking corresponding checkboxes in the activity panel allows the analyst to filter and identify terrorist events that fulfill any investigative criteria. It is important to note that the colors shown in the activities panel are the same as the colors in the ThemeRiver as well as the map view and can therefore be used as a color legend for the entire view.

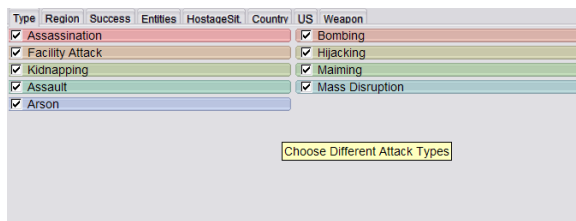


Figure 5: Activities Panel: This is an interactive filtering panel, which uses dimensions in the GTD to filter and color events in the other views.

Who: Entity View

We use a combination of three panels to enable the analyst to find any targeted terrorist groups among more than two thousand in the GTD. An alphabetical list of the groups allows the analyst to start the investigation by searching on

specific names (Figure 6(left)). A correlated view connects to the map view, the ThemeRiver, and the activities panel in a way that filters the terrorist groups by *where*, *when*, and *what*. Lastly, we offer a shoebox as a container for collecting terrorist groups that are of interest to the analyst (Figure 6(right)). An individual terrorist group is inserted or removed from the shoebox manually, allowing the analyst to form and test hypotheses.



Figure 6: Entity view: The Terrorist Name tab (left) lists 2404 terrorist groups extracted from the GTD, while the Intermediate Results tab shows suggestions for possible terrorist groups. The Shoebox (right) is a container for collections of user-selected terrorist groups, providing the user an easy way to compare different terrorist groups.

6. Scenarios

To show the utility of our system, we identified a few scenarios in collaboration with the START center that demonstrate possible ways in which the system can be employed. Our focus in these scenarios is not only on the goals of the investigation, but also on the paths which an analyst might take to arrive at the conclusions. In both scenarios described below, we demonstrate that an analyst can begin to deduce the *why* of the attacks by identifying the other four other W's.

6.1. Linking Tactical Incidents and Global Strategy

In this scenario, an analyst begins by examining one of the most talked-about groups in recent years, Hizballah (sometimes spelled as Hizbollah/Hezbollah/Hizbullah), by selecting it from the entity view (*who*). It is widely known that

Hizballah is an Islamic political and paramilitary organization based in Lebanon that has recently been in conflict with Israel. Based on their reputation, one might reason that the activities carried out by Hizballah would concentrate in the Middle East. However, upon highlighting Hizballah in the entity view, it becomes immediately evident that Hizballah's terrorist activity between 1979 and 1997 is not limited to a specific region (Figure 7). A closer inspection in the map view (*where*) shows that out of the 200 attacks by Hizballah, the majority of the incidents take place within Lebanon, but it is clear that they do indeed have a global strategy of operation in both Europe and South America.

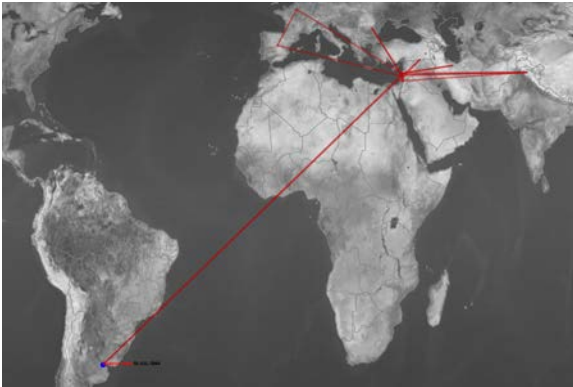


Figure 7: In this example, edges indicate that Hizballah is an active international terrorist group, with a wide attack range. The highlighted city illustrates a significant outlier in Argentina in 1994.

One obvious outlier in the pattern of incidents is the major attack in Argentina in 1994 for which Hizballah claimed responsibility. Double-clicking on that incident in the map view shows that the attack was on a Jewish community center in Buenos Aires that resulted in 96 deaths and more than 250 injuries. Although inconclusive, this incident suggests that Hizballah's attacks are not based on their geographical locations, but influenced instead by the religious beliefs of its targets (*why*).

Examining the timeline (*when*) also reveals a change in Hizballah's attack pattern over time (Figure 8). It appears that between 1984 and 1988, Hizballah's primary attack method (*what*) was kidnapping. Starting in 1989, its strategy changed to a mix of bombing and facility attack. Finally, facility attack became the predominant attack method in the early to mid 90s, while the number of bombings decreased and eventually stopped.

The investigation in this scenario included all four of the W's and ended in a limited but plausible hypothesis of *why* Hizballah conducted terrorism around the world. It uncovers both tactical decisions and methods of the Hizballah's

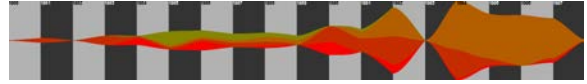


Figure 8: This image depicts changes in Hizballah's attack methods over the last 17 years, which shifted from kidnappings to a mix of bombings and facility attacks, to almost entirely facility attacks. Note that the GTD does not contain data for the year 1993 [LD07].

attacks, but also makes clear the trend and pattern of their global operations.

6.2. Discovering Unexpected Patterns

By examining the overview shown in Figure 1, we can see that a great deal of terrorist attacks took place in the Philippines. Zooming into that specific region and selecting the entire country in the map view (*what*) lists all the terrorist groups active between 1970 and 1997. A quick search in the entity view (*who*) shows that the NPA (New People's Army) is one of the most active groups in the region (Figure 9). Highlighting NPA reveals that although active, NPA is strictly domestic and has never performed activities outside of the Philippines.

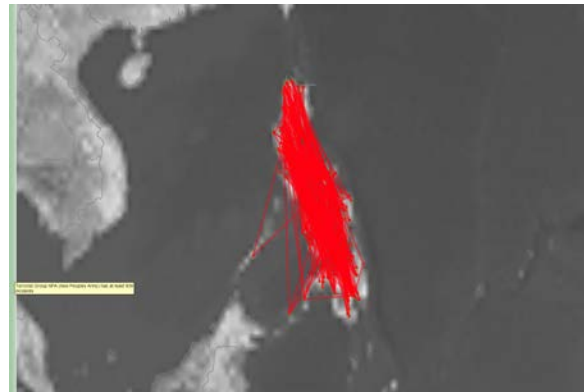


Figure 9: This image indicates that the NPA was an active domestic terrorist group that was responsible for more than a thousand incidents within the Philippines.

An examination of the ThemeRiver (Figure 10), however, shows an interesting temporal pattern in NPA's activities. It appears that while NPA was indeed active during the '80s, it stopped operating entirely in 1992. This unexpected temporal behavior, unfortunately, cannot be explained using only the data within the GTD. In order to find out what happened in 1992 in the Philippines (*why*), we turn to newspaper archives and found out that in 1992 Fidel Ramos became

the twelfth president of the Philippines, and immediately offered peace treaties to various terrorist groups including the NPA [She92]. Although it is unclear whether this event is directly linked to the diminished activities of NPA, it forms the basis for a plausible hypothesis for further investigation. This shows how evidence gained from the GTD visual analysis can provide a specific basis for continued investigations using open source or other data.

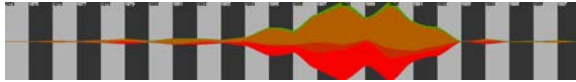


Figure 10: This result shows changes in the NPA's attack frequency during the time period of the dataset. Though it was very active in the late '80s, it suddenly disappeared in 1992.

7. User Evaluation

We asked three groups of experts in law enforcement and criminology to evaluate our system. Our experts are from the criminology department at the University of Maryland, the National Insurance Crime Bureau (NICB), and the Drug Enforcement Administration (DEA). The evaluation was conducted informally. First we presented our system by demonstrating the investigative scenarios described in the previous section. Then the experts were given some time to interactively explore the system and the GTD. Finally, we concluded the evaluation by asking them to give feedback and comments. All these experts gave consent to have their comments and affiliations published in this paper.

7.1. The Five W's (Who, Where, When, What, Why)

At the heart of our system is the integration of the five W's in a comprehensive and cohesive manner. This design is validated by one of our expert evaluators, who serves as a special agent in the NICB. He said, "The five W's you are using here are exactly similar to what we have in police reports. It is critically important for us to understand individual incidents from different aspects, like the *who*, *where*, *when*, *what*, *why*." Also, according to GTD data designers at University of Maryland, "In particular, your system presents users with an efficient means to access multiple dimensions of terrorist event data simultaneously. True to the goals of the project, it allows users to see where specific terrorist groups were operating during defined time periods and to discern the nature of the weapons and attacks that the groups engaged in. As such, it provides a streamlined mechanism for helping users to identify behavioral trends among terrorist groups over time. Traditionally, users working with the numeric version of the GTD would have to sort through tens of thousands of rows of data to come up with information on

cases that meet specific criteria. This tool greatly facilitates this effort."

7.2. Tactical vs. Strategic Analysis

Tactical analysis is traditionally considered to be the pursuit of short term analytical goals using resources at hand, while strategy is the longer-term pursuit to understand a group or a situation at a higher level. Our experts from law enforcement confess that the large numbers and density of reports they receive on a daily basis force them to think tactically, rather than strategically. They agree that an exploratory visual analytics system could help to reduce the amount of noise they have to sift through in order to see the broader picture and hone in on the suspicious outliers. Two experts explicitly pointed out that it has become more and more important for tools to help analysts take a step back from tactical analysis to strategic analysis. They agreed that our tool will assist this analytical pattern by visually providing a global pattern as well as details on demand. According to the law enforcement analyst from the DEA, both outliers and global patterns are critical to analysis. "It is very useful to directly mark those targets with their strategic shapes. Your system could benefit the Federal government's interest in investigating both local terrorist groups and their attempts to develop ties with other international groups."

7.3. Reducing Communication Gaps

Conveying knowledge and investigative results visually can drastically reduce the amount of effort spent on communication and reduce ambiguity. Since our system is designed to display the five W's, it can present a complete and coherent picture of the current state of an analytical process. According to the analyst from NICB, the current practice of using text-based reports for sharing insights and hypotheses found during an investigation process tends to be error-prone and time-consuming. "[Your system] will greatly shorten the catch-up time between police shifts and guide them to focus more on things that they are interested in."

Along the same lines, the expert in criminology praised our system for its intuitiveness and ease of use. Since the GTD is a large and complex dataset containing tens of thousands of records with over a hundred dimensions, it has been difficult for non-experts in terrorism investigation to utilize this wealth of data. With a fully interactive and exploratory interface, our system effectively shortens the distance between terrorism experts who prepare and collect the data and the users of varying levels of expertise and backgrounds who seek to identify patterns and trends in terrorist activities around the world.

8. Discussion

Encouragingly, the experts from law enforcement who participated in our evaluation were eager to use our system for

their own purposes. During the evaluation, both of these analysts were already picturing their own data visualized in our system. They both consider the structure and nature of the GTD to be very similar to the financial and criminal reports that they investigate, and could see the tool having an immediate impact on multiple aspects of their day-to-day jobs, including tactical analysis of incidents, strategic thinking about global trends and patterns, and communication and reporting of their investigative findings with peers and superiors.

More importantly, they foresee an unexpected use of the system in predicting future trends and activities. Our system's ability to depict temporal trends clearly in relation to both geographical and other patterns suggests possible future directions of events. Although the predictive capability of the system is not one that we had considered, we are very excited about its promise and potential benefits in the field of law enforcement. Based on these positive responses, we plan to place our system in the hands of these and other experts for continued use and evaluation.

The only major criticism of our system concerns the visual representation of overlapping incidents and overplotted lines in the map view. For instance, when multiple events occur in the same city or region, it is not easy to see the number of events that overlap each other (e.g., there are more than 700 incidents in Beirut alone). This may lead the user to underestimate the degree to which a specific location has been attacked. Similarly, with overplotted lines in a condensed region, the shape of the activities and their temporal relationships can easily be lost. We are investigating new visual representations that may alleviate these issues.

9. Conclusion

We present a visual analytics system for exploring terrorist activities in the Global Terrorism Database. We employ a coordinated multi-view approach with high interaction to depict the five W's (*who*, *what*, *when*, *where*, and *why*) of investigative analysis. With our tool, analysts can both explore tactical incidents and identify strategic patterns and trends. It can also reduce ambiguity and effort in communicating the results of analysis.

References

- [Fre00] FREEMAN L. C.: Visualizing social networks. *Journal of Social Structure* 1 (2000).
- [GLM07] GUO D., LIAO K., MORGAN M.: Visualizing patterns in a global terrorism incident database. *Environment and Planning B: Planning and Design* 34, 1 (2007), 767–784.
- [HHWN02] HAVRE S., HETZLER E., WHITNEY P., NOWELL L.: Themeriver: Visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics* 8, 1 (2002), 9–20.
- [LD07] LAFREE G., DUGAN L.: Introducing the global terrorism database. *Terrorism and Political Violence* 19, 2 (2007), 181–204.
- [Per06] PERER A.; SHNEIDERMAN B.: Balancing systematic and flexible exploration of social networks. *Transactions on Visualization and Computer Graphics* 12, 5 (Sept.-Oct. 2006), 693–700.
- [Pla05] PLAISANT C.: Information visualization and the challenge of universal usability. In *Exploring Geovisualization* (2005), Elsevier.
- [She92] SHENON P.: For filipino new dealer, a hard road to recovery. *New York Times*, 1992.
- [SMER06] SHEN Z., MA K.-L., ELIASSI-RAD T.: Visual analysis of large heterogeneous social networks by semantic and structural abstraction. *IEEE Transactions on Visualization and Computer Graphics* 12, 6 (2006), 1427–1439.
- [Zhu07] ZHU WEIZHONG; CHEN C.: Visual analysis of terrorism events extracted from the public knowledge bases. *Transactions on Visualization and Computer Graphics* (2007).