Global Terroristic Attacks

Visual Analytics

Sapienza Università di Roma

Ludovico Ottobre, 1712005
Gianmarco Evangelista, 1711818

March 02, 2021

Figure 1: Home page: Global Terrorism Attacks in 2017

Abstract – In these years we have often heard about terrorism and the consequences it has around the world. Terrorism is, in the broadest sense, the use of intentional violence for political or religious purposes. Wanting to analyze this problem, we have developed a visual analytics tool to allow the user to analyze terrorist attacks in the world from 1970 to 2017. The multiple interactive views of the project provide both low level and high level information and help the user in understanding terrorism events patterns.

1 Introduction

Terrorism is a serious problem that affects the whole world. In the recent years we have perceived how even European countries can be victims of attacks. Although strongly rooted in the Middle East, in general this phenomenon is present everywhere.

In this article we present Global Terrorism Attack, a visual analysis system for evaluating data with respect to terrorist phenomena around the world. In particular, the system is designed to allow the user to analyze the data of each country for a period of 47 years. In addition to the possibility of having an analysis on a general scale, it is also possible to make comparisons between two different countries and evaluate how they have been victims of terrorism by analyzing the methods of attack, the type of attack and the terrorist group.

Our main goal is to allow the user to obtain information and evaluate how this problem has developed over the years and in which forms it has expressed itself. We have focused our attention on *why, where, when, who, what*.

2 Dataset

In order to implement the idea described above, we needed lots of details about the type of attack. At the moment of the creation of our system, we were able to find a single open source dataset that includes all the information we need: `https://www.kaggle.com/START-UMD/gtd`.

The Global Terrorism Database (GTD) is an open-source database including information on terrorist attacks around the world from 1970 through 2017. The GTD includes systematic data on domestic as well as international terrorist incidents that have occurred during this time period and now includes more than 180,000 attacks.

Time period: 1970-2017, except 1993; *Unit of analysis*: Attack; *Variables*: more than 100 variables on location, tactics, perpetrators, targets, and outcomes.

2.1 Preprocessing

In order to manage in the right way the dataset, we needed of preprocessing to establish which variables we needed and to remove some events that didn't give a complete information respect the ones we have used. At the end of the preprocessing we have about 80,000 attacks in fact we have deleted all the attacks with a number of kills equal to zero - it would have been useless for our analysis goal - and all lines with missing data. In the final part of the preprocessing, PCA algorithm is performed in order to calculate two PCA components and store them as two new columns of the dataset. The algorithm is performed on different features of the dataset (not all) that have been already standardized.
The GTD requires missing value processing and conversion of characters to numeric values and numerical processing, In order to correctly compute the two components of the PCA, we shared and adopted the choice made during the preprocessing of the same data in [1] and in particular we converted char-

acter fields to numeric fields. Following the study carried out in the paper, the data mapping was calculated based on the frequency of occurrences of the respective data within the database. The PCA is used in our tool for dimensionality reduction and to classify terrorist events.
To avoid to overload the computation cost of the project, PCA is just computed once before the start and they will not change anymore then. Those principal components will be later used in the scatterplot view to give a 2D representation.

3 Related Works

Analyzing this phenomena, we have seen how there is an extensive literature devoted to the study and analysis of terrorism. How can this work located in with respect to existing literature about similar topics?
We have analyzed various documents that have allowed us to create a tool that can be a good solution to study the trend of this global problem in recent years. Considering [2] as the main reference, we have studied how the authors of this project have worked on this topic. In particular we have started from the same concepts with respect to the investigative analysis: the five W's (*why, where, when, who, what*).
Having as goal the study of the terrorist trend over the years, [2]'s team provides an analysis strictly linked to the long-term trends. Although they perform an analysis with respect to the terrorist groups involved, this provides information that allows the user to make a very broad as well as generic assessment. For this reason, we have considered necessary to implement a tool in order to allow the user a simple and accurate analysis by country and between countries year by year to evaluate the data in the best possible way and provide reliable results.
In addition to an approach based on long-term scales, thus allowing a general evaluation, and not on annual scales, [2] doesn't provide to the user a compare modality and therefore leaves to the user the task of carrying out this type of analysis which we have considered of fundamental importance especially for evaluating the trends in neighboring countries.
A different choice of the graphs allowed us to solve a common problem, in fact we have used bar chart in order to provides numerical data about attacks and kills to show in the right way all the events since in [2] and in our tool we have an overlapping in the map and the overplotted lines in the parallel coordinates when multiple events occur in the same country/city and they don't provide a clear view.
Although with different interactions and analyses, there are some similarities between the two applications that allow us to evaluate our project as a valid tool for scientific, investigative and social analyzes.
Another paper that we have analyzed is the [3] in which the main goal was to evaluate what were the ways, weapons and targets over the years, in order to understand how terrorist strategies have changed over time. The choices made however provide a macro analysis without the possibility of evaluating case by case and we strongly wanted it in our homework.
However, taking into consideration various papers that have analyzed in

particular the terrorist trend in certain countries, our tool results efficient and reliable.

4 Visualization

Different views are used in this project to build a complete tool for analyzing terrorist attacks.
Those views are built using D3.js and exposes a wide of information.

4.1 Header

While not being a chart or a graph the header too plays an important role in the visualization. The header comes with filtering year or group selection capabilities. In order to make the system easy and dynamic, we have used *vue.js* so that even the selected boxes were also updated with respect to data changes. Those proportions are computed at run-time and changes only when the user changes the year or the group. When the user want to filter data using the country selection, it can see only the groups with at least an attack in those country.

4.2 Geographic Map

The geolocation of the events is represented by a geographic map in which events are placed by their latitude and longitude.
Even if it is a very simple view, it helps the users giving them the position of the events which otherwise is hard to get it.
In this map the attacks are represented on each place with a circles which can potentially change color when the events are selected. On each circles, positioning the mouse on it, the user has the possibility to see more details on this particular attack.
The user has also the possibility to manage the map using the zoom.

4.3 Parallel Coordinates

The parallel coordinates chart is the graph that coordinates the other graphs. Through this graph the user can modify the values of the bar chart, scatter-plot and the other graphs. It allows to compare the feature of several individual observations on a set of variables. Each vertical bar represents a variable and has its own scale. Values are then plotted as series of lines connected across each axis so we have different attributes of the dataset in a single plot.
The axis represents the following attributes: *Region; N°Kill; Attack type; Weap type; Target Type.*
The user can interact with the graph and change the filters of visualization on the dataset. He can select a number of lines of the graph and through this (brush selections on each axis), he changes also the visualization of all the other elements of the page. The lines of the graph represents the attacks.

4.4 Bar Chart

The bar chart shows the top 15 countries ranked by the number of terrorist attack events that occurred inside their borders.
To better analyze them, the user can click on the button in order to see which are the most cities with more attacks or the number of kills with respect to country/cities.
During the compare mode, through the bar chart, we can see all the pre-

vious stats but for both the selected countries in different colors in order to have a graphic comparison.
If the user wants to perform a cross analysis, selecting country and group, in the bar chart we have that the bars change their color and their trend in order to provide data based on the requests made. The group selection also affects the map and we the same color encoding for both.

4.5 Graphs

In order to make a better analysis, the user can use the drop down menu to choose among different views. Let's see in detail each of them:

4.5.1 Global Line chart

In the initial window the user can see a line chart that identify the global trend of this phenomena during the period under analysis. The colors chosen follow the trend of the line, in fact warmer colors have been chosen to identify the peaks.
During the compare mode, this graph shows the attacks trend of the two selected countries during the 47 years.

4.5.2 Kills Line chart

Using a line chart viewer, the user can evaluate the kills trend over the 47 years. This graph follows the same color rule of the previous one.

4.5.3 Group Line chart

Following the previous approaches, the user can select a group in order to see when and with which frequency it develops itself.

4.5.4 Group Bar chart

If there aren't a country and a group selected, it shows the bars relative to the most hit target for year.
If the user selects a group, the showed bars are relative to the group selected for that year.
If there are a country and a group selected, the user can see the attack of that group in that year in this selected country.
If the user is in compare mode, it can see the sum of the attacks in these countries of the selected group, categorized for target.

4.5.5 Box Plot

For some distributions, you will find that you need more information than the measures of central tendency (median, mean, and mode). So we used the box-plot. A box plot is a graph that gives you a good indication of how the values in the data are spread out.
We used a box-plot for the kills distribution in the selected place. If the user is in compare mode, we have two box-plots, one for each country.

4.6 Scatter plot

The scatterplot chart is used to represent the elements of the dataset after a dimensionality reduction on two dimensions. In particular, this graph represents the first two components of the execution of PCA. Each point of the scatterplot represents a terrorist event, projected on the first two principal components.
By observing this chart the user can find similarities over the events looking at the distance between the points and

searching for clusters.
What emerges is that points are quite all clustered and only some of them are far from the main cluster representing the so called outliers.
In order to add more information on the graph, we added a chromatic encoding of the points, representing the relative region in which the event happened. For a deeper analysis, the user can select a zone of interest on the graph, this selection is highlighted on the parallel coordinates graph and the geographic map.

5 Interactions and Analytics

5.1 Interaction

The project includes multiple possibilities of interactions.
These interactions allow the user to perform a deeper analysis of the dataset.
We already introduced them when talking about the views, here is a recap:

Header: performs filter on the dataset on the year and on the group terroristic name in order to display all the attacks arised in that year or made by this group.

Geographic Map: allows pan and zoom. With this graph the user can initialize the compare mode; clicking on a country, the user can deeply analyze data of that country. If the user enable the compare mode, the data selected will filter all the other parts of the project.

Parallel Coordinates: it has many axes, each one representing a feature, that allow the user to brush them to select an interval of data whose he is interested in. This selection is highlighted on the geographic map and on the parallel coordinates itself.

Bar Chart: the bar offers the possibility to analyze different levels of attack frequencies for the relative country. Using the two buttons available, the user can see others informations about the cities most affected and the number of deaths.

Graph: the user can select the desired view through the drop down menu.

Scatter Plot: allows brush selection on the events, this selection is highlighted on the geographic map and on the parallel coordinates chart.

While some of these interactions, influence the other graphs, there are other interactions that allow the user to do only a better analyze without any type of correlate updating.

5.2 Analysis

Analytics part involves some of our graphs and also the computation of other informations. Let's see them in a list:

Frequency: this first analytics is computed on the bar chart graph, at runtime the 15 countries with most terrorist attack events are computed and this result is ordered in a data structure. This data is then displayed on the graph. As for the country, in the same way we have a frequency about

cities and deaths.

Quantiles of box plots: the quantiles of the box plots are computed at runtime, these values are the principal component of the box plot chart.

Scale: many scale adaptation are used to adapt data to graphs.

Other computation: for example we had to recompute, each time the selected year or group changes, all the views, the size and the range of the axis.

6 Conclusion

However, currently we only have limited understanding of the causes, development, and diffusion of terrorism activities. Moreover, terrorism data are often incomplete or inaccurate, and only represent the outcome (eg incidents) but not the process. These factors limit our ability to formulate a valid hypothesis and to test the hypothesis. Therefore, we believe that exploratory analysis approaches, including both visual and computational methods, have much to offer in the exploration of terrorism data by revealing unknown trends or regularities, prompting new thoughts, and helping the analyst to gain insights to formulate better hypotheses and models.

References

[1] H. Z. Yu Jun Xian Tong and L. Yutong, "Hazard grading model of terrorist attack based on machine learning," *International Journal of Advanced Network, Monitoring and Controls*, vol. 04, no. 02, pp. 81–85, 2019. [Online]. Available: https://www.ijanmc.org/201902/2019-02-11.pdf.

[2] X. Wang, E. Miller, K. Smarick, W. Ribarsky, and R. Chang, "Investigative visual analysis of global terrorism," *Computer Graphics Forum*, vol. 27, no. 3, pp. 919–926, 2008. DOI: https://doi.org/10.1111/j.1467-8659.2008.01225.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2008.01225.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2008.01225.x.

[3] D. Guo, K. Liao, and M. Morgan, "Visualizing patterns in a global terrorism incident database," *Environment and Planning B: Planning and Design*, vol. 34, no. 5, pp. 767–784, 2007. DOI: 10.1068/b3305. eprint: https://doi.org/10.1068/b3305. [Online]. Available: https://doi.org/10.1068/b3305.