The Kronos Incident (Mini-Challenge 1)

Pilar Avila, Amalia Guaymás and Valeria Burgos

Abstract – The solving of the VAST Challenge 2014 MC1 case was performed through application of various visualizations. In this work, we looked at the use of information visualization for identifying key elements in the Kronos Incident using an extensive analysis and interpretation of data, data mining techniques and programming.

Index terms – Visual analysis, timeline, streamgraph, dynamic graph

1 INTRODUCTION

In the country island of Kronos, the increasing noxious effects on health and farming have been related to the uncontrolled activity of GAStech, a natural gas operator, supported by corrupt government officials. On January 20th, 2014, a corporate meeting is held to celebrate the new-found fortune because of the initial public offering of the company. However, a series of rare events occur that lead to the disappearance of several employees. The Protectors of Kronos (POK), a social movement organization that has been fighting against water contamination and government corruption, is suspected in the disappearance.

As analysts, we were assigned with several tasks in order to analyze the evolution of the social organization over time and highlight the important events surrounding the disappearance. For these purposes, we analyzed the available data and provided novel visualizations to communicate key information. The overall procedure is depicted in Figure 1.

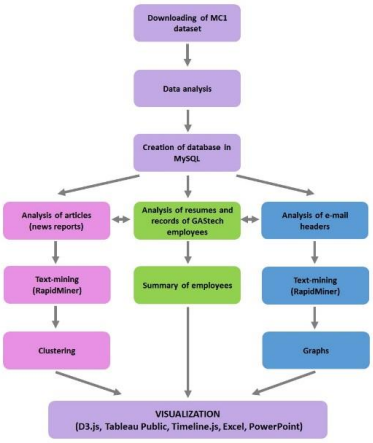

Fig. 1. Overview of the procedure

2 DATA DESCRIPTION

The available information consisted of several documents of both GAStech and POK: 1170 internal GAStech company e-mails headers from Jan 6 to Jan 17, 2014; resumes of 35 employees, 54 employee records, 2 historical reports and descriptions and 844 files of relevant current and historical news reports from multiple domestic and translated foreign sources.

- Pilar Avila is at University of Buenos Aires, Argentina. E-mail: piwica@gmail.com.
- Amalia Canavire is at University of Buenos Aires, Argentina. E-mail:acarolinagc@gmail.com.
- Valeria Burgos is at University of Buenos Aires, Argentina. E-mail: valaburg@gmail.com.

3 VISUALIZATIONS

We developed several static and dynamic visualizations to represent the structure and leadership of POK over time and try to elucidate the potential connections between POK and GAStech. We also implemented a timeline visualization to describe the events surrounding the incident. The visualizations were implemented using Tableau Public, JavaScript library D3.js, RapidMiner, Timeline.js and PowerPoint.

3.1 Graphs

Graphs are used to depict data in which information is comprised of objects and the relations among them. An interactive graph to represent the leaders of POK was created such as by clicking on any name, it would fade out a brief summary of information specific to that person. Changes in the structure of the organization over time were represented by a specific arrangement of the elements in the graph (Figure 2).

We also implemented a hierarchical edge bundling [1] to visualize e-mail communications among GAStech employees. Incoming e-mails in green and sent e-mails in red can be visualized by using the mouse over any of the nodes in the network. In this way, this dynamic visualization allowed us to represent the potential connections between GAStech and POK members (Figure 3).

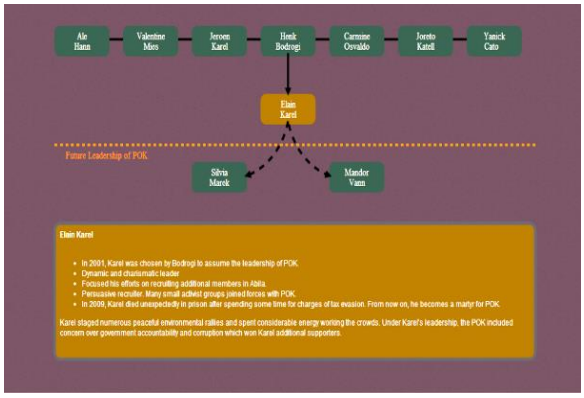

Fig. 2. Structure of the network of Protectors of Kronos.

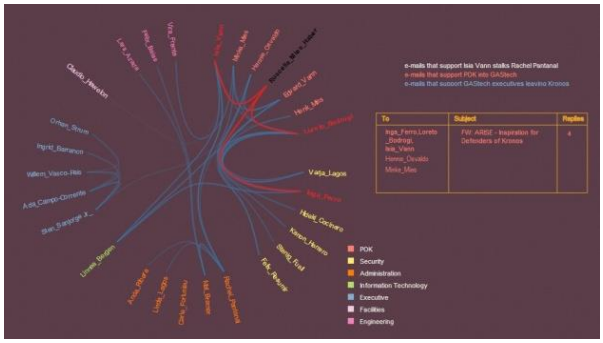

Fig. 3. Network of internal e-mails between GAStech employees.

3.2 Timeline

Timelines are used to visualize anything related to time. We used timeline visualizations that allowed zooming in by clicking on an element to reveal more in-depth information about an event of interest. As roughly depicted in Figure 4, specific information about changes in the structure of POK would be displayed by clicking on the colored elements.

On the other hand, in order to give a big-picture view while still providing full detail of the events surrounding the disappearance of GAStech employees, we used Timeline.js (Figure 5).

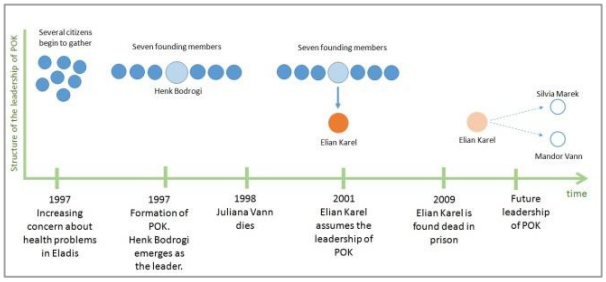

Fig. 4. Changes in the structure and organization of POK.

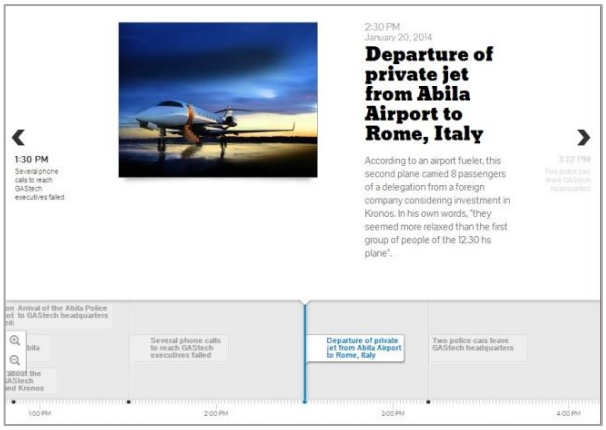

Fig. 5. Timeline of the events occurred on Jan 20-21, 2014.

3.3 Charts

We found that a good way to represent how news articles perceived both POK and GAStech over time was to search for the most frequent words associated to each entity and try to find a tendency.

The evolution of the organization POK was reflected into a streamgraph. This layout emphasizes the legibility of individual layers, arranging them in a distinctively organic form [2]. As depicted in Figure 6, at first POK was associated to "positive" words such as "leader", "activist" and "health" (indicated in green colors). However, from 2009 and on, the organization began to be associated to "negative" words such as "force", "violence" and "arrest" (indicated in warm colors).

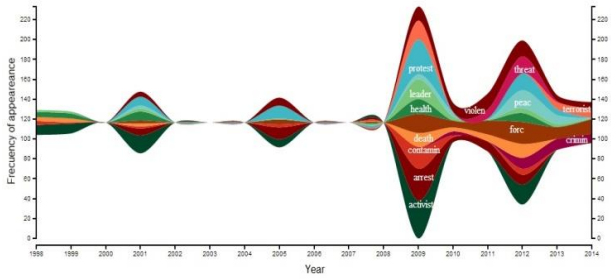

Fig. 6. Most frequent words associated to POK in news reports.

We used a frequency bubble chart to visualize the most frequent words associated to GAStech, as perceived by the media. This visualization allows us to compare the evolution of the perception comparing the bubble area. As shown in Figure 7, by the late 90s, the word "company" began to be associated to words such as "contaminant" and "water". And by 2010, the company showed associations to words such as "corrupt" and "government" (Figure 7).

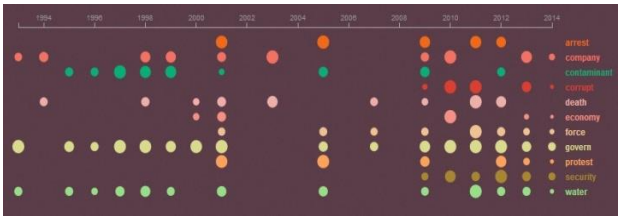

Fig. 7. Most frequent words associated to GAStech in news reports.

CONCLUSIONS

The case studied here resembles a real-life scenario and throughout the solving of the situation, we realized that data mining skills are required but also data analyst with good domain knowledge and visualization analysis. In order to support law enforcement in solving the disappearance, we provided several graphical answers to the questions made in the mini-challenge [3]. We believe that the visualizations presented achieved the objectives of communicating key information to solve the mini-challenge.

REFERENCES

[1] Danny Holten, Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. IEEE Trans. Vis. Comput. Graph., 12:741-748,2006.

[2] Lee Byron and Martin Wattenberg. Stacked Graphs – Geometry & Aesthetics, IEEE Trans. Vis. Comput. Graph.,14:1245-1252,2008.

[3] http://piwica.github.io/infovis/Presentation/index.html