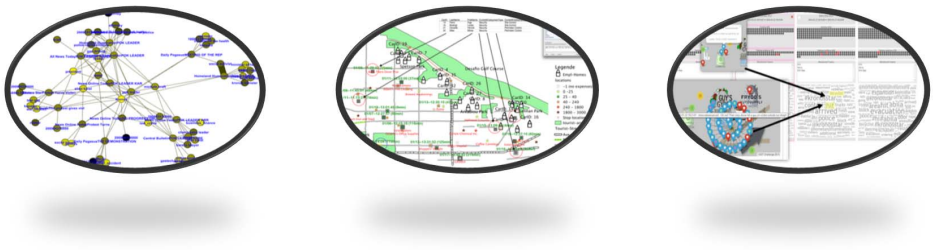
Erschienen in: 2014 IEEE Conference on Visual Analytics Science and Technology (VAST) : Proceedings ; Paris, France, 9-14 October 2014 / Min Chen ... (Hrsg.). - Piscataway, NJ : IEEE, 2014. - S. 301-302. - ISBN 978-1-4799-6227-3

Using Visual Analytics to Support Decision Making to Solve the Kronos Incident (VAST Challenge 2014)

Fabian Fischer, Florian Stoffel, Sebastian Mittelstädt, Tobias Schreck, Daniel A. Keim

Data Analysis and Visualization Group
University of Konstanz, Germany
{firstname.lastname}@uni-konstanz.de



ABSTRACT

Gaining insights from different heterogeneous data sources is one of the biggest challenges in decision making support. The large volumes of data can only be combined by sophisticated automatic methods. However, unexpected patterns can only be identified with the help of human intuition. In this paper, we present our visual analytics work-flows and tools to process heterogeneous data such as social networks, text streams, and geo-temporal data. We apply these tools on the VAST Challenge data and present our findings and assumptions that we identified in our analysis.

1 INTRODUCTION

The fictional scenario of VAST Challenge 2014 was the so-called *Kronos Incident* in which several employees of a company named *GAStech*, located at the island of *Kronos* went missing. Because of an ongoing conflict between an organization known as the *Protectors of Kronos (POK)*, they are suspected in the disappearance. In the grand challenge the focus is on combining all provided data sources and to summarize the events of the incident for an overview. Further, the challenge is to identify the existing networks and possible suspects as well as the locations the police should focus their investigations. Therefore, heterogeneous data such as social networks, text streams, and geo-temporal data is provided and the challenge is to analyze and combine the insights of these sources. In the following, we will present our visual analytics work-flows and tools before we present our final insights and assumptions that we would hand over to the decision makers.

The grand challenge is a classical visual analytics problem, where analysts are "being asked to make decisions on ill-defined problems. These problems may contain uncertain or incomplete data, and are often complex to piece together" [1]. To answer the grand challenge questions, we made use of our novel visual analytics tools and the insights, which are described in the individual entries for MC1, MC2, and MC3. For example, during the analysis of the data stream, we realized the fire and used the data and insights from MC2 to identify, who is living in that region - which was identified as Dancing Dolphin Apartment Complex.

1.1 Making Sense of Networks (MC1)

MC1 was approached by constructing an undirected graph from the given documents. The sense making process was driven by visualizations of specialized sub graphs, which were created by querying the graph based on an analysis question. Further evidence has been searched by using an ElasticSearch instance, which allows flexible full text search and inspection of the results. All facts found by examining both, the graph visualizations and text search results, were then used to reformulate the queries resulting in a graph structure, which is then visualized again. This is a classic drill down technique, which combines relation and distance information from the graph and textual information from the documents.

1.2 Making Sense of Geo-Spatial Data (MC2)

The presented Geographic Information System (GIS) is aimed to interactively analyze the complex geo-temporal data of MC2. Most of the preprocessing steps are implemented as KNIME workflows. To reduce the GPS-data we extract the stop-locations of cars by movement thresholding. Thus, complex issues such as the wrong booking times are handled relating the stop of the person's car at the according places. To estimate the approximate coordinates of locations like shops, GAStech headquarters, and "homes" we joined the credit-card and GPS-data of all persons. The data is visualized on a map that shows the stops of suspects and locations (see Figure 1). A time series indicates the amount of movement over time that guides the user to interesting time frames. Further, suspects and location types can be interactively filtered to discover unexpected behaviour patterns.

1.3 Real-Time Visual Analytics for Data Streams (MC3)

To solve MC3 we used *NStreamAware*, which is our real-time visual analytics system to analyze data streams. We make use of various modern technologies like Apache Spark Streaming and others to provide high scalability and incorporate new technologies. Furthermore we developed a novel web application, called *NVisAware*, to analyze and visualize the given microblog and call center messages in real-time, to help the analyst to focus on the most important time segments. We extracted so-called sliding slices, which are aggregated summaries calculated on a sliding window and represent them in a small-multiple like visualization containing various small visualizations (e.g., word clouds) as seen in Figure 2.

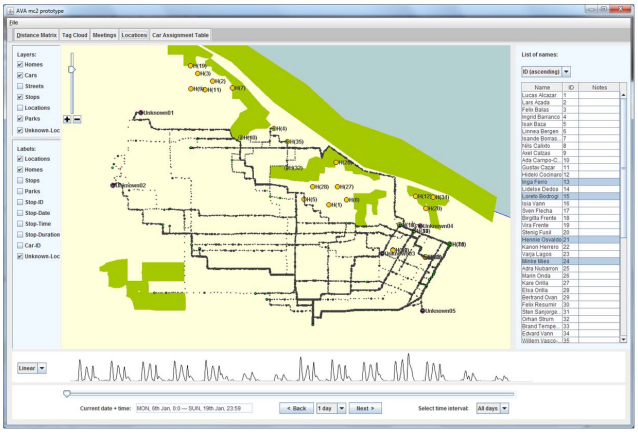


Figure 1: Using visual analytics to extract suspicious locations, which are possible hiding places of the terrorists.

Figure 2: NVisAware: Sliding slice visualization with interactive geographic map to compress the data stream to visually identify events and answer complex questions.

2 Results and Findings

In the following we summarize and describe major events of January, 2014 related to the disappearance of the GAStech employees in chronological order:

2014-01-07	**APA** Night observations of possible hostages (MC2).
2014-01-08	Start of regular **APA** meetings at 5 suspicious locations (MC2).
2014-01-09	**APA** Night observations of possible hostages (MC2).
2014-01-11	**APA** Night observations of possible hostages (MC2).
2014-01-14	**APA** Night observations of possible hostages (MC2).
2014-01-20	10:00 - Fire Alarm at GAStech (MC1).
2014-01-20	12:25 - **Reports on missing GAStech** executives (MC1).
2014-01-23	17:00 - **POK** rally to start in Abila City Park (MC3).
2014-01-23	18:41 - **Fire** at Dancing Dolphin Apartment Complex (MC3).
2014-01-23	19:20 - APA **black van** uses the fire confusion relocating hostages.
2014-01-23	19:40 - APA black van stuck at at Gelato Galore. Cop got shot.
2014-01-23	21:21 - APA members in van surrender, two hostages free (MC3).
2014-01-23	21:30 - Explosion at Dancing Dolphin Apartment Complex (MC3).

2.1 Hypothesis Testing to Identify Interesting Events

After extracting various insights and the evaluation of hypothesis with our visual analytics tool, we came to the following assumptions: Generally, we suspect the APA (Army of People of Asterian) to be responsible for the terror in Kronos. The APA does not see any other possibility to bring justice to the country and decided to kidnap or even kill GAStech executives to eventually stop the company. They started night observations and preparations to kidnap the different employees and also met at various places, where they created hiding places for future hostages. We were able to identify the persons belonging to the different groups.

The APA activates the fire alarm at GAStech utilizing the confusion to deport the executives to a nearby hiding place. Three days later they use the confusion during the POK rally and a fire to move the hostages to other places and/or to get them out of the country. However, one of their vans got caught by the police and the two POK members surrender after a firefight in which a police officer was killed. One of them might be Inga Ferro. Based on a message in MC3 we know that a person named Rachel might be missing. This is probably Rachel Panatal. She could be one of the hostages in the black van, who is free again. She is executive assistant and is reporting to chief corporate officers. So she is aware of all the meetings and might be supporting the terrorist or is forced to support them. She reports to the CIO, who is Ada Campo-Corrente, who might actually be an hostage.

2.2 Identification of Significant Networks

We have identified four groups of networks. The first one are the POK, which are persons organizing a peaceful rally. The second one are the APA, which are involved in the ongoing terrorist activity. POK seems to be too peaceful for them, so they obviously decided to really attack the GAStech company. In the group of hostages are the executives of GAStech. Then, we identified some persons who might or might not belong to POK who use the kidnapping of the APA to blackmail GAStech for money, even if they are not involved in the kidnapping at all. We think that these persons do not belong to the APA, because the APA is not interested in money. They want to destroy GAStech bringing it to an end to stop poisoning the environment.

2.3 Support the Final Decision Making

From the data we know that the highly suspicious suspects (from APA) met at five previously unknown locations as seen in Figure 1. These might be the hiding places for the hostages as well. Locations labeled in Figure 1 as *Unknown01* and *Unknown02* in the above figure seem to be the best choices to start with. The first one was the last place the terrorist went to. And the other one is located closest to the airport, which might be a reason they picked that place as well.

3 Conclusions

We developed various visual analytics tool and could address all mini challenges to answer the questions for the grand challenge. The combination of suitable visualizations and the use of modern and highly-scalable infrastructures resulted in tools, to provide actionable insights to the analysts. Future work would have to be done to integrate the different tools developed for MC1, MC2, and MC3 to provide an integrated visual analytics system. Currently, the different tools are loosely coupled and more should be done to optimize the insight management in a collaborative working environment with multiple analysts.

Acknowledgements

We would like to thank our course students Michael Hundt, Natascha Siirak, and Manuel Wildner for their great work on MC2.

References

[1] J. C. Roberts, D. A. Keim, T. Hanratty, R. R. Rowlingson, R. Walker, M. Hall, Z. Jacobson, V. Lavigne, C. Rooney, and M. Varga. From Ill-Defined Problems to Informed Decisions. *EuroVis Workshop on Visual Analytics (2014)*, 2014.