VAST 2014: Summary on Grant Challenge Work

Yang Liu*
ASTRI

ABSTRACT

VAST 2014 grant challenge required a big picture on a fictional company employee missing event. This paper introduces the software we developed to tackle the challenge. In addition, how we integrate the visual analysis results from related information identified is included. And, the integrated analysis results (motivations, suspects, and related locations) are illustrated to help further investigations.

Keywords: VAST 2014, grant challenge, visual analysis.

Index Terms: [Human-centered computing]: Visualization—Visualization application domains—Visual analytics

1 INTRODUCTION

VAST 2014 proposed a theoretic story. The Tethys-based gas company GAStech has made huge profits in the island country of Kronos. However, Kronos people suffered from pollution related to the GAStech commercial activities. On January 20th 2014, several employees of GAStech were missing after the company after the management group annual meeting. Protectors of Kronos (POK), one local environment protection organization seemed to be the suspect. Accordingly, three mini challenges are presented as well as the grant challenge. The first mini challenge demands us to describe events on January 20th and 21st based on numerous text such as news reports, resumes, and email headers. As a contrast, mini challenge two focuses on abnormal patterns derived from car GPS records, credit card transactions and loyalty card usage data in the two weeks before January 19. While in mini challenge three, we were asked to find important events in a streaming data (twitter like text) of 4.5 hours on January 23 afternoon. Finally, grant challenge requires us to combine all the information to help further investigation, such as the network enrolled in the situation, current important information gaps, and the locations worth digging more. [1]

Since we already worked on the three mini challenges, some software developed there were carefully selected to reply the grand challenge questions. In details, first of all, we studied the motivations, as well as the suspects in the disappearance with a cross validation on the answers to mini challenge one and mini challenge two. Secondly, we tried to describe the whole story by selecting important events found in the three mini challenges. Finally, based on the suspect behaviors and timeline findings, we proposed our observations and opinions for further investigation

The rest of the paper is organized as follows. In sections two, the software developed for three mini challenges are illustrated separately. In section three, the analysis procedure and results are presented. Finally, we summarized the findings.

2 SOFTWARE

In mini-challenge one, the data is basically text. We develop a java-based application to automatically build ontology models of Kronos, Tethys, POK and GASTech for their members, features and relations from related documents and extract keywords from news articles. Then we combine all the analysis results and provide users with an integrated platform to query the information of and the relations among organizations or people. This platform provides convenient interfaces for users to see all the relevant information to assist their investigation at a glance such as displaying all the relevant event keywords for the query subject (Figure 1), showing the original news text by clicking a keyword, and revealing the relations by Java D3 library (such as collapse tree and bubble chart).

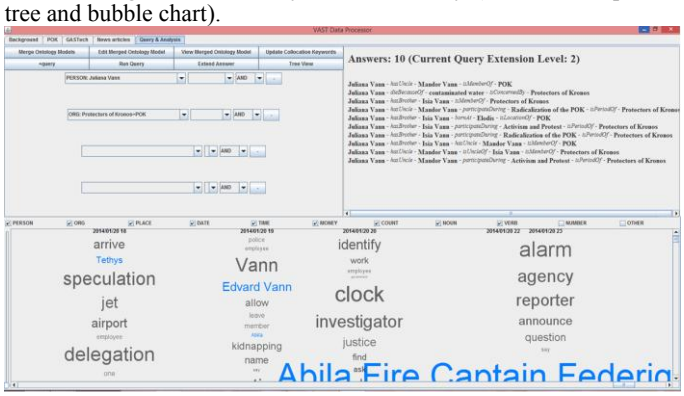

Figure 1: Mini Challenge One Software Interface Example.

In mini-challenge 2, we build a database to deal with the GPS data, credit card transactions and loyalty cards records by extracting the information from raw data. The data base is stored in SQLite. In order to link up the credit card and loyalty card transaction records with the GPS data, the following assumptions are made accordingly. First, we assume the employees drive their own cars for shopping, meals, and work. And, we assume after they park the car with a reasonable amount of time, (for example, 15 minutes), they will use their credit/loyality card. Finally, we assume employees stay at their apartments in the mid-night. With the locations are linked with the GPS data, we use HTML5 to draw a line chart to show the routine pattern with classified locations that we identified (Figure 2). Also, Python imaging tools are used to draw the real routes in a same scale. And, the consumption records of loyalty card and credit card are visualized in the view of consumer and location. Therefore, it is convenient to identify multiple abnormal patterns for further investigation by checking the figures generated.

* yangliu@astri.org

IEEE Symposium on Visual Analytics Science and Technology 2014
November 9-14, Paris, France
978-1-4799-6227-3/14/$31.00 ©2014 IEEE

Figure 2: Mini Challenge Two Software Traffic Pattern Example.

In mini-challenge 3, we need to analyze the streaming data to find events for the further investigation. We develop a Java-based application to support both the real-time processing on streaming data from network and off-line processing on the data from batch files. To unearth an event, we try to abstract the key components from the raw data: who, how, where and when. After processing the messages, we aggregate the results and display them with multiple charts. Our chart based platform can help users to clearly see the key subjects, actions, locations, the timeline, and their relationship. Therefore, it is convenient to identify an event and its progress to assist their analytics (Figure 3).

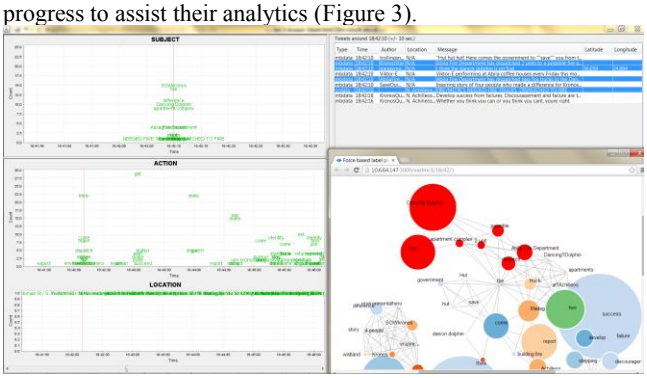

Figure 3: Mini Challenge Three Software Interface Example.

3 Discussion

3.1 Analysis Procedure

Isia Vann is our start point since he is considered suspect in both mini challenge one and mini challenge two. In mini challenge one, he is the only one that is an employee in GAStech as well as a key member in POK (Figure 1). In mini challenge two, he is one of those people with multiple middle night driving records (Figure 2).

Accordingly, Isia Vann's email traffic is checked in mini challenge one. The suspects list is then extended to the receipts of Email circle "RE: FW: ARISE - Inspiration for Defenders of Kronos": Isia Vann, Inga Ferro, Loreto Bodrogi, Hennie Osvaldo, Minke Mies. In addition, the abnormal multiple middle night traffics check in mini challenge two returns a list includes Isia Vann, Minke Mies, Loreto Bodrogi and Hennie Osvaldo.

Following the cross validated suspect list, we analysis the related consumption and locations. The location information revealed in mini challenge three is also included in this phase.

After combining all the related information we have in the three mini challenges, we tried to reveal the suspects, the victims, the events, and the motivations.

3.2 Analysis Results

First of all, we think the disappeared GAStech employees include executive group of GAStech and the kidnapper suspects. The executive group is composed of Ingrid Barranco, Ada Campo-Corrente, Orhan Strum, and Willem Vasco-Pais. The kidnapper suspects include Isia Vann, Inga Ferro, Loreto Bodrogi, Hennie Osvaldo, Minke Mies. The news in mini challenge one said the latest missing list was reduced to ten people. Therefore, we guess Isia's uncle, Mandor Vann was the last one in the missing list, as a suspect. But the guess was not indicated in our submission since no direct proof was found.

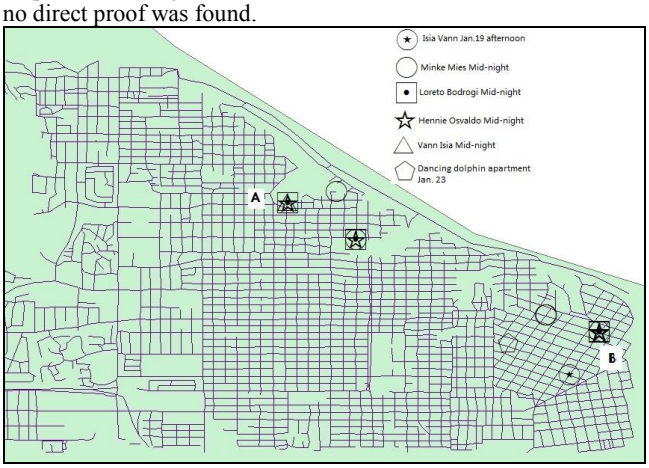

Figure 4: Location Analysis Example.

Secondly, the multiple places visited in the middle night by the suspects are plotted in Figure 4. In addition, the dolphin apartment fire observation point indicated from mini challenge three (Figure 3) is also shown in Figure 4. Considering the observation spot should be a certain distance from the fire, we guess with some confidence that dolphin apartment is one of places that the suspects frequently visited at right of Figure 4 (point B).

Thirdly, the related events, as well as some analysis, can be listed. The employee disappearance happened at Jan. 20, around 10:00 am. Before that, the executive group was having the annual meeting at GAStech building starting from 9:00 am. After 10:00 am, the suspects made a call to claim a false fire alarm, and took the victims away in the evacuation. Probably, the the executive group and some kidnappers left Abila by private planes at Jan. 20 afternoon. The suspects may plan the actions several time at the locations listed during Jan. 7 to Jan. 20 in Figure 4. Point A and Point B are the most visited places by the suspects in Figure 4. And, at Jan. 23, some of the suspects may try to set a real fire at dolphin apartment (point B in Figure 4) to remove the evidence. After that, the suspects flee caused a traffic accident, as well as a subsequent gunfire with local police. Therefore, we recommend point A and point B in Figure 4 for further investigation.

But, the biggest issues remained are the motivations. It seems POK has the strongest motivation, politically and personally, as we suggested in the submission. But, another possible scenario is someone else did that under POK's name. For example, if the executive group of GAStech wishes to jump off the company with the profit they already got, they can hire part of POK members to have the false kidnapping. However, again, we do not indicate the non-POK motivation in the submission due to the lack of direct proof.

4 Conclusion

To the help of our visual analysis tools, we understand the VAST 2014 story and found some clues for later investigation efficiently. But further study should be still needed to reveal the real motivations as well as the criminals.

References

[1] http://vacommunity.org/VAST+Challenge+2014