AnnotatedTimeTree, Dodeca-Rings Map & SMART: A Geo-Temporal Analysis of Criminal Events

Chen Guo[1], Jing Xia[4], Jun Yu[1], Jieqiong Zhao[2], Jiawei Zhang[2], Qiaoying Wang[3], Zhenyu Cheryl Qian[3], Yingjie Victor Chen[1], Chen Wang[5], David Ebert[2]

[1]Computer Graphics Technology, [2]Electrical and Computer Engineering, [3]Interaction Design, Purdue University

[4]State Key Lab of CAD & CG, Zhejiang University

[5]New Media Technology & Art, Harbin Institute of Technology

Index Terms: [Human-centered computing]: visual analytics, geographic visualization, information visualization; [Information Systems]: spatial-temporal systems.

1 Introduction

The 2014 VAST Grand Challenge required us to find victims, suspects, and criminal motivations based on three separate datasets. We developed three VA tools (AnnotatedTimeTree, Dodeca-Ring Map and SMART) to facilitate the understanding of heterogeneous multivariate datasets. These tools were integrated to gain insights into the source data and find connections among complex information (Fig. 1). AnnotedTimeTree aims to identify the cause clues and timeline of kidnapping based on analysis of text and network data. Dodeca-Rings Map allows analysts to interact with geospatial, temporal, and card transaction data to find suspicious personal behaviors and social networks. SMART is a visual analytics tool that enables text stream analysis by dynamically visualizing microblog data on the map over time.

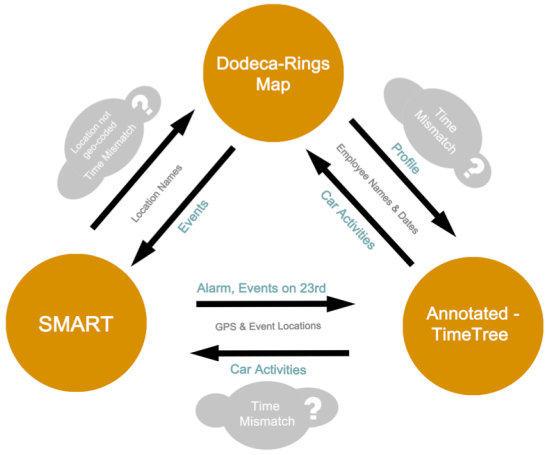

Figure 1: This Grand Challenges is solved by the collaboration of three visual analytics systems

To address this Grand Challenge, we collaborated widely with VA researchers from Purdue University, Harbin Institute of Technology, and Zhejiang University from China. The background of team members ranges from Computer Graphics Technology, Interaction Design, to Computer Engineering.

1.1 AnnotatedTimeTree – Mini Challenge 1

AnnotatedTimeTree is a time-critical document visualization tool designed for organizing large collections of reports and annotations. In order to capture the key events that happened in historical news and reports, we used a vertical tree-structure timeline to organize documents in reverse chronological order; the top ones are the most recent. The deep shade of green in each cell represents the numbers of documents. AnnotatedTimeTree also provides semantic interactions such as searching, highlighting, annotating, and expanding for maintaining reasoning process and reasoning results in the documents.

1.2 Dodeca-Rings Map – Mini Challenge 2

Dodeca-Rings Map is designed to analyze geo-temporal traffic problems. We used dodecagons that divide 24 hours into two-hour intervals to visualize one car's 24-hour activities. The color segments denote the stop periods in each location and color dots represent activities such as credit card transactions and twitter messages. The activity temporal chart opens all the Dodeca-Rings with the sample temporal length, and allows analysts to compare activities with different ordering choices. The social relationship matrix helps analysts to understand the general social connections.

1.3 SMART – Mini Challenge 3

SMART is a social media analytical tool to investigate temporal trend, geographical distribution, social networks, and semantic evolution. The top timeline presents statistical distributions of streaming data. The text clouds on the right side visualizes key words extracted from microblog data within filtered time ranges. On the background map of Abila City, blue dots show geo-tagged microblog data and the red dots represent call center data. The network view demonstrates the retweet or reply relationships. The analyst can use content lens to illustrate prominent keywords associated with selected microblog users.

2 Data Integration and Collaboration

Openness, connectivity, and flexibility are the key techniques we used for analyzing with the three tools. The original data sets of the three Mini Challenges have different formats. The Grand Challenge asked us to integrate network, text, temporal, geospatial, and stream data as well as explore the storyline of the disappearance incident. From the introduction, we can see that the whole story is a spatial-temporal event that involves strong underlying social networks. Thus we need to combine information gathered from all the three systems to make the final picture.

* email address: {guo171, yuj, jieqiongzhao, zhan1486, wang1925, qianz,victorchen@purdue.edu};xiajing@zjucadcg.cn;chwang@hit.edu.cn; ebertd@ecn.purdue.edu

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

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

To identify the clues and investigate the events, we needed to allocate the information of key events' when (time), where (location), and who (the social network). In most cases the datasets only provided partial information of an event, for example, in MC3, we were give the location and time of standoff, but who stayed in the suspicious black van was not clear. Combining the partial information by matching the known pieces provided us with more information. Dodeca-Rings Map provided found that their houses received visits from security people in some very early mornings (Fig.2 A). Several security people routinely visited an executive's home at approximately 3:30am and then left at 8am. From the social relationship matrix, we also found that this group of security personnel has a close relationship, but are not friends with those executives. The suspicious early morning visits led us to form the hypothesis that these security personnel were involved in the kidnapping event.

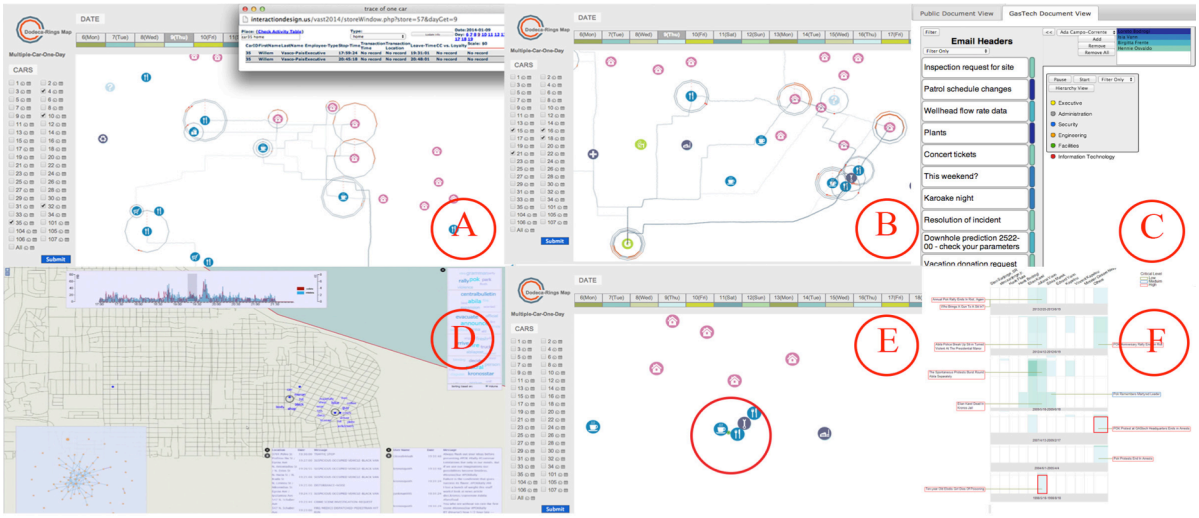

Figure 2: A: The executives' houses received visitations from security people in some very early mornings. B: Highlight several security people's names in Dodeca-Rings Map. C: Pass security people's names to AnnotatedTimeTree. D: Highlight the hit and run geolocation in SMART. E: Pass the suspicious location to Dodeca-Rings Map. F: POK has been protesting GasTech and Kronos Government for a long time.

spatial-temporal analysis and hints of several social networks within GAStech. Data from SMART and AnnotatedTimeTree contain rich spatial-temporal information about events. Linking up the events with networks is essential. AnnotatedTimeTree and Dodeca-Rings Map are connected by passing names between the two systems (Fig.2 A, B & C). Analysts can select employees' names in Dodeca-Rings Map and find relevant information in the document view of AnnotatedTimeTree, and vice versa. SMART and Dodeca-Rings Map are connected by passing geolocations (Fig.2 D & E). By clicking the suspicious geolocation in SMART, the same location will be marked on the Dodeca-Rings Map.

Dodeca-Rings Map has complete spatial-temporal information of the given two-week period. It can accept partial event information and provide a more complete picture of what happened before the disappearance incident. For example, given the person (car) and time of an event, the location of the event can be identified using Dodeca-Ring Map if the time is within the 2-week range. Similarly, given the person and location, the time of the event can be estimated at the period of the stay in the location.

3 Analysis Process

To investigate the events of January 20 related to the disappearance of the GAStech employees, we used the three visual analytics tools collaboratively. Analysts used AnnotedTimeTree and Dodeca-Rings Map to develop the initial hypotheses. SMART was able to provide evidence and support for current investigations and prove the various hypotheses.

At first, in the Dodeca-Rings Map system, we examined the executives' car activities in the multiple-car-one-day mode, and

By passing all the security people's names to the AnnotatedTimeTree's GAStech view, we found that two emails are significant. One was related to the protest of POK, and the other one attempted to obscure the motivation of the first email as a virus. These emails mentioned a person named Edward Vann, who shared the common family name with the key persons in POK. We also found that POK had been protesting GasTech and Kronos Government for a long time (Fig.2 F). Hence we assumed that the security department's employees, who were indeed POK members, kidnapped the executives.

Searching "kidnap" from SMART, we found that on January 23rd, a van hit a car and a bicyclist and then proceeded to head West on Egeou St. at approximately 7:30pm. By passing the hit and run geolocation to the Dedoca-Rings Map, we found that the place is near a car repair store called Frydos Autosupply n' More. Five security people frequently visited this store after work. The unusual activity further supported our assumption that several security people kidnapped four executives in GAStech on 20th.

4 Conclusion

The three visual analytics tools construct a powerful analysis system that helps analysts discover suspicious spatial-temporal patterns and behaviors in the data sets. Each system feeds information to others and helps to explore the unknowns. By passing different types of data among systems, we are able to retrieve and even create the missing temporal or geospatial information. For the future developments, we will build more direct linkages between the three systems and allow them to pass keywords, networks, archive documents, and filter information.