Henkin, R., Kachkaev, A. & Slingsby, A. (2014). Summarising the structure of an organisation and reconstructing a chain of events - VAST 2014 Mini-Challenge 1 Submission Honourable Mention for Novelty in Visualization. Paper presented at the IEEE Conference on Visual Analytics Science and Technology, 09-11-2014 - 14-11-2014, Paris, France.

Summarising the structure of an organisation and reconstructing a chain of events

Rafael Henkin, Alexander Kachkaev and Aidan Slingsby

Fig. 1. Summary of the structure of a terrorist organisation using a *chalkboard* metaphor. Changes in membership and leader are shown above the horizontal line.

Index Terms—Visual analytics, information visualization

1 Introduction

The 2014 VAST mini-challenge 1 asked participants to summarise the structure of a terrorist organisation and how it has changed over time, reconstruct the chain of events of a kidnapping and to provide two possible explanations.

2 Summarising the structure of a terrorist organisation

Our design for communicating key information about structure of a terrorist network to the "busy leaders of the investigation", shown in Fig. 1 won an Honorable Mention. It uses the metaphor of a chalkboard to convey important information in form that is familiar and reminds us that our knowledge often transient and incomplete.

Layout is an important aspect of our design. Members of the terrorist network were displayed *above* the horizontal line, with the left/right position indicating changes in membership and leadership over time. Other members with suspected links are indicated *below* the line. Red text indicates confirmed terrorist network membership, yellow indicates that there is evidence of links to the network, and green indicates that there is some evidence of this, but it is poor.

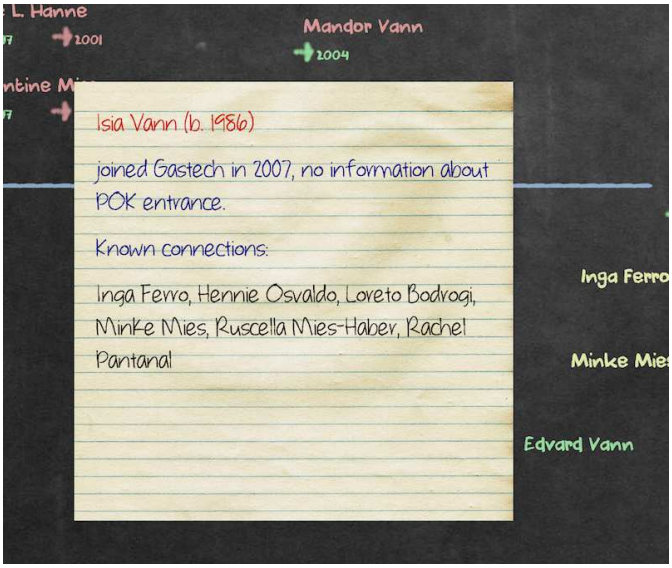

Fig. 2. Supporting evidence as details-on-demand (excerpt).

- *Rafael Henkin, Alexander Kachkaev and Aidan Slingsby are at giCentre, City University London. E-mails: rafael.henkin.1@city.ac.uk, alexander.kachkaev.1@city.ac.uk, aidan.slingsby.1@city.ac.uk*

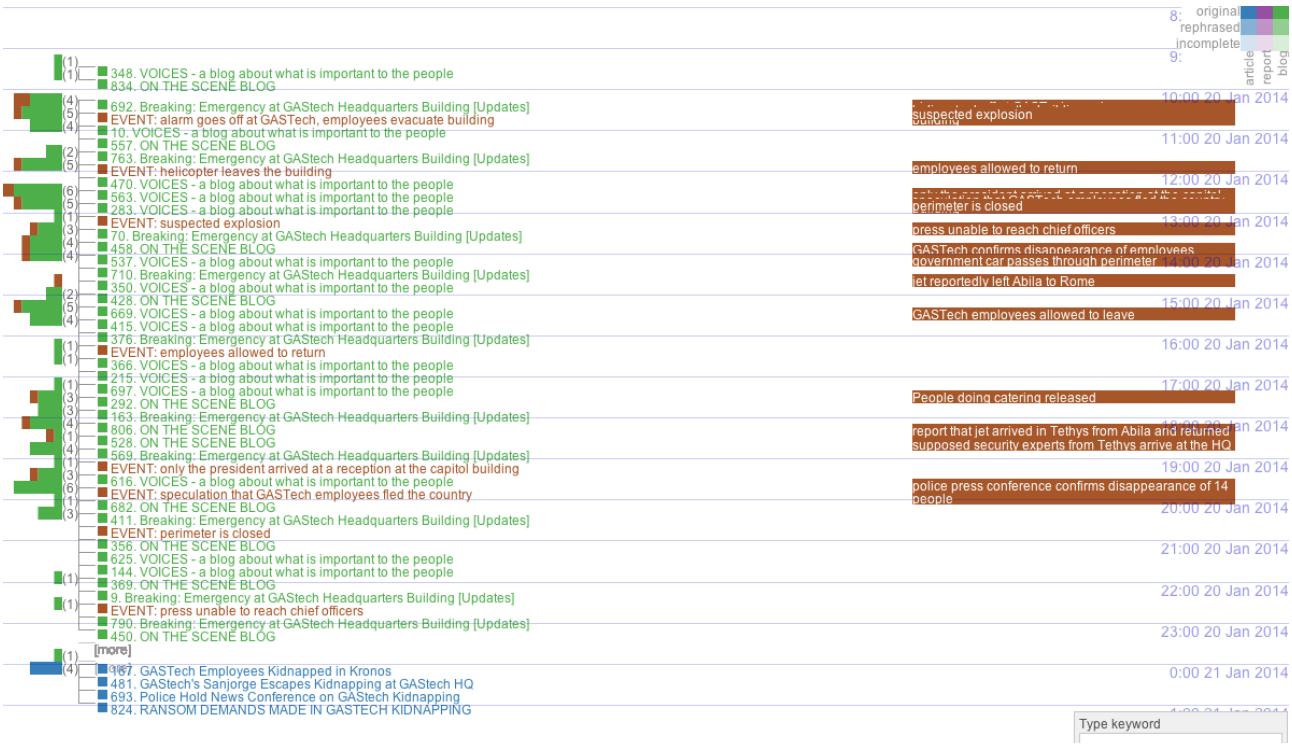

Fig. 3. The document (left) and event viewer (brown; right). Documents are coloured by type (article=blue; report=purple; blog-green) and their lightness indicates their quality (original, rephased or incomplete) and they can be filtered by these.

In a departure from the chalkboard metaphor, *interaction* is used to provide details-on demand (Fig. 2), including evidence, personal information, documents in which the person is mentioned and potential connections with other people. In a large multi-user touch-screen context, there is potential for incorporating gestures to make these interactions more natural.

In order to construct this structure, we were provided with a number of heterogeneous data files. Log files of email exchanges helped us establish interactions between individuals. We used matrix- and graph-based representations to help us interpret these. Reading through the resumes of individuals enabled us to add context to suspicious organisations (Fig. 2). Long historical documents about the history of the organisation were split into more digestible document sections and were browsed along with the news documents provided in our document timeline viewer described in the next section. Information from these document provided important context.

3 Reconstructing a Chain of Events

We were supplied with 845 timestamped documents that were news articles, other reports and blog posts. We used these to help collate relevant information for the structure of the organisation (previous section). These were also key for reconstructing the series of events during a kidnapping.

Initially, we explored the documents through the Jigsaw application (http://www.cc.gatech.edu/gvu/ii/jigsaw/). There were some problems with Jigsaw's entity identification and we wrote scripts that identified the entities of interest and annotated the articles in Jigsaw's XML-based file format. The *temporal sequence* of documents was important but we were not able to view this to our satisfaction in Jigsaw. So we produced our own viewer (Fig. 3) that reads the XML file and enables documents to be rapidly browsed, read and marked in their temporal context (see video at http://vimeo.com/101531951).

We noticed characteristics, problems and inconsistencies with some of the documents. Around the time of the kidnapping (20-21 January), distinct "report" and "blog" styles of documents exist. On the basis of various characteristics we identified, we categorised the documents and were able to filter these in our application. We also noticed that some of these documents appeared to be poorly-written or be incomplete versions of others. To help manage the number of documents to read, we wrote scripts to automatically identify and classify these. Language quality was assessed through the service provided by After the Deadline (http://www.afterthedeadline.com/). This quality score allowed us to identify the probable 'original' articles. We then identified the sources for rephrased or incomplete copies using sentence pattern extraction. Our tool enables documents to be filtered by type and quality. Type and quality are encoded using colour with hue indicating type and lightness indicating quality; see legend at the top right of Fig. 3. Filtering by document type and keyword (bottom right of Fig. 3 enabled us the control we needed to effectively browse the documents.

Finally, we added functionality to allow us to identify 'events', display these on the timeline (on the right, in brown) and linked to all the supporting documents. These are quickly and conveniently created using YAML and we used this to identify the events of the kidnapping. Using the linked articles, we were then able to identify two possible explanations for the kidnapping.

4 Conclusion

This mini-challenge required us to design interactive visualisation for *two types of user*. Visual analytics techniques in existing and custom tools helped us - as analysts - interpret the supplied data, identify the structure of a terrorist organisation, and the events of a kidnapping. A friendly- and non-technical-looking interactive graphic was used to provide important aspects of the structure of the terrorist organisation to "busy leaders of the investigation". This illustrates two important roles of visualisation: *(a)* conveying relevant and known information and *(b)* facilitating exploration and interpretation of data.