Eurographics Conference on Visualization (EuroVis), Posters Track (2017)
A. Puig Puig and T. Isenberg (Editors)

ConTraffic Visual Analytics in Support to Customs Risk-Analysis

M. Poulymenopoulou and A. Tsois

European Commission, Joint Research Centre (JRC), Via E. Fermi 2749, I-21027 Ispra (VA), Italy

Abstract

Customs risk analysis is crucial for detecting fraud and contraband goods in the massive flows of internationally traded goods. Most of non-bulk goods are transported in shipping containers and, as customs can control only about 2% of them, efficient customs risk analysis is crucial. In support to EU customs, the Joint Research Centre of the European Commission has developed the ConTraffic visual analytics research prototype. This paper presents the main architectural elements of the application and some visualization and user-interaction techniques selected to enable the route-based risk analysis of large number of shipping containers.

Categories and Subject Descriptors (according to ACM CCS): Human-centered computing - Visual analytics

1. Introduction

Worldwide there are more than 20 million shipping container boxes transporting the majority of non-bulk internationally traded goods. Every year the logistics industry generates about 1 billion records that describe the status and movement of these shipping containers. These records basically define, for each container, the logistic operation, the location, the date and time of the operation and the involved vessel. The customs officers in the European Union have recently gained access to this massive amount of information and they have started using it to improve their risk analysis process. As customs have the capacity to control only about 2% of the goods transported in containers, it is only through high quality risk analysis that they can identify the fraud cases or illegal imports or exports of goods. To this direction, the ConTraffic visual analytics application is a research prototype that employs visual interfaces and interaction techniques aiming to provide to authorities the appropriate tools to explore this massive amount of container-traffic data in order to improve their risk-analysis.

2. ConTraffic visual analytics

The ConTraffic system deals with a database of around 3 billion records describing the movement of more than 17 million shipping containers [CDT12, MFC*16]. The real challenge is, in addition to any automated data analysis tools, to provide appropriate visualizations of the massive amounts of container data and interaction tools in order to identify those relatively rare cases of fraud or contraband goods that the automated processes failed to identify. Hence, there is a need to provide a general overview of large datasets of thousands of containers of interest in order the customs to create hypotheses and then through interaction to view details for small sets of targeted shipments in order to reject or verify hypotheses. The performance of the application is also an important requirement to enable real-time analysis of large data volumes. To accommodate these requirements, the ConTraffic Visual Analytics architecture has been designed as follows: a) database layer - stores, indexes and retrieves data, b) web services layer - transforms user requests into database queries and prepares the retrieved results, and c) client layer - enables users making complex data requests, visualizing and interacting with the results. The client layer is a single page web application with enhanced functionality that allows fast visualization and manipulation of the retrieved data through in-memory data clustering and filtering, reducing the need for high performance servers and high-speed network connections. This approach allows to serve multiple simultaneous users and with a certain degree of privacy that is highly appreciated in the customs domain. Hence, the main workflow in the visual analytics application is: a) the user first identifies a broad dataset to work on through the selection of a set of search criteria, (e.g. select all the shipments arriving in Rotterdam in March 2017); b) the data are retrieved from the server and visualized in the client web application, c) the user interacts with the visualizations or restricts the data using filtering on various attributes of the data. In the following we focus only on the visualization and interaction features implemented on the client layer.

2.1. Visual representations

In ConTraffic, a number of visual representations are used, each of which is best suited to highlight different dimensions of the spatio-temporal data [AAB*11, AAW07, FPV*13, KAF*08, Rob07, SMPS11]. The main visualization method is an interactive geographic map that represents, through markers, the locations where the containers handling operations took place. The containers flow is presented through geodetic lines or estimated maritime routes

DOI: 10.2312/eurp.20171177

linking the consecutive locations visited by the containers. The maritime routes are estimated using a predefined spatial maritime network that covers the globe. To provide clear and understandable map representations, even when the selected datasets are large, spatial clustering and adaptive level of detail strategies have been employed. For example, Figure 1 (a) shows the data on the traffic of 400 containers that travel from Brazil to Spain.

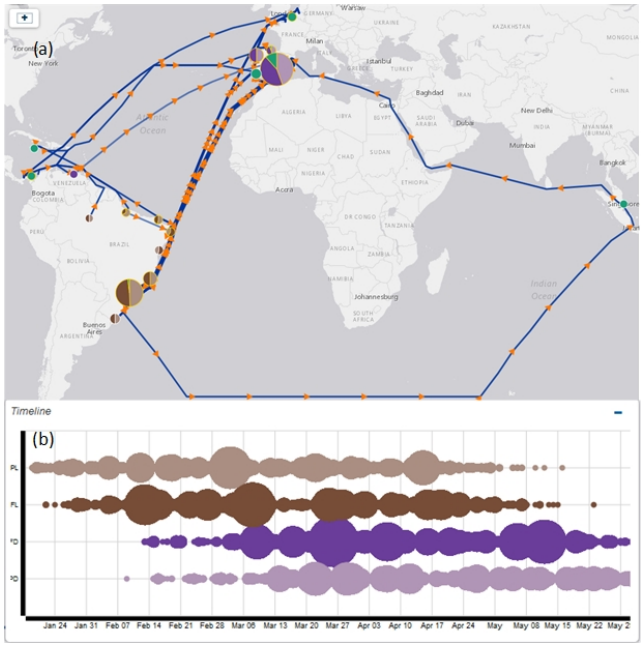

Figure 1: *(a) Map view and (b) bubble timeline view of 400 containers traveling from Brazil to Spain.*

The second visual representation is a bubble timeline chart, Figure 1 (b). This shows the time distribution of the different transportation phases (e.g. loading on vessel at departure port). When the number of selected containers is low, detailed timelines show the series of events for each container across the time, along with the duration of the various transportation phases and the overall duration. These can be useful to visually identify suspiciously long transportation phases or containers that travel together across the time. All views are cross-linked in the sense that any data selection or data filtering in one representation is propagated to the others.

2.2. Filtering

Filtering is an exploration tool for focusing on particular containers and verifying hypotheses created from the map visualization. Filters are constraints applied to attributes of the data like locations, trip dates, trip duration, vessel names and container IDs. Those are defined through drop-down boxes (Figure 2 (a)), where each selected filter is supplemented with a pie chart showing statistics of the filter values. Locations filters can also be applied through selections on the map in order to include or exclude (NOT) a set of locations (Figure 2 (b)). For example, the containers of Figure 1 (a) with stop on Singapore seem suspicious, hence, Singapore region is selected to be added into the filters in order to view in more detail those containers. Figure 2 (c) shows the filtered result and detailed timeline of the resulted container. Filtering is particular useful during the exploration process for examining in detail some cases or excluding some common cases. Complex filters can be defined using criteria on several attributes but the system prevents the user from selecting filters that would result in no data.

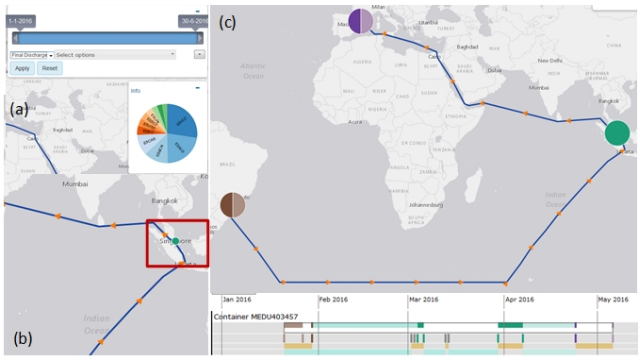

Figure 2: *(a) Filters, (b) Filter through map selection, (b) Filter results.*

2.3. External document integration

Customs during their risk management processes deal also with sensitive information which cannot be shared or uploaded to servers, like ConTraffic, outside their organization. In order to allow users to enrich the ConTraffic data with such information, a data loading mechanism has been implemented in the visual analytics client application. In such cases, the additional data are linked with ConTraffic data using the Container ID attribute and a text table is used to display the combined result. An important feature is the ability to use the additional attributes for filtering. For example, by loading a document with information on the type of goods transported in a set of containers, the custom officer can filter the visualized data using criteria on both the type of products and departure countries.

3. Implementation

For the ConTraffic client implementation, custom javascript libraries have been developed using open-source libraries (D3.js, Leaflet map). This implementation produces uncluttered map visualizations of 15.000 container trips records in <20ms and data filtering in <10ms, that are acceptable interaction rates for an immediate response perception [TC05]. External document integration has been successfully tested with documents of 15.000 records.

4. Conclusions and future work

The ConTraffic visual analytics application is already used experimentally by customs and has received favorable evaluations. Work is still in progress, several challenges and limitations are still to overcome and many improvements are planned.

References

[AAB*11] ANDRIENKO G., ANDRIENKO N., BAK P., KEIM D., KISILEVICH S., WROBEL S.: A conceptual framework and taxonomy of techniques for analyzing movement. *Journal of visual languages and computing 22* (2011), 213–232. doi:10.1016/j.jvlc.2011.02.003. 1

[AAW07] ANDRIENKO G., ANDRIENKO N., WROBEL S.: Visual analytics tools for analysis of movement data. *ACM SIGKDD Explorations Newsletter, Special issue on visual analytics 9*, 2 (2007), 38–46. doi:10.1145/1345448.1345455. 1

[CDT12] CAMOSSI E., DIMITROVA T., TSOIS A.: Detecting anomalous maritime container itineraries for anti-fraud and supply chain security. *Proc. 2012 European Intelligence and Security Informatics Conference. IEEE Computer Society* (2012), 76–83. doi:10.1109/EISIC.2012.39. 1

[FPV*13] FERREIRA N., POCO J., VO H., FREIRE J., SILVA C.: Visual exploration of big spatio-temporal urban data: A study of new york city taxi trips. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2149–2158. doi:10.1109/TVCG.2013.226. 1

[KAF*08] KEIM D., ANDRIENKO G., FEKETE J., GORG C., KOHLHAMMER J., MELANCON G.: Visual analytics: Definition, process and challenges. *Information Visualization LNCS 4950* (2008), 154–175. doi:10.1007/978-3-540-70956-5_7. 1

[MFC*16] MAKRIDIS M., FIDALGO R., COTELO A., TSOIS A., CHECCHI E.: A quality assessment framework for large datasets of container-trips information. *International Conference in Computer Information Systems and Industrial Management 9842* (2016), 729–740. doi:10.1007/978-3-319-45378-1_63. 1

[Rob07] ROBERTS J. C.: State of the art: Coordinated and multiple views in exploratory visualization. *International Conference on Coordinated and Multiple Views in Exploratory Visualization, IEEE Computer Society Press* (2007), 61–71. doi:10.1109/CMV.2007.20. 1

[SMPS11] SILVA R., MOURA-PIRES J., SANTOS M.: Spatial clustering to uncluttering map visualization in solap. *International Conference on Computational Science and its Applications, Lecture Notes in Computer Science 6782* (2011). doi:10.1007/978-3-642-21928-3_18. 1

[TC05] THOMAS J., COOK K. (Eds.): *Illuminating the path: The research and development agenda for visual analytics*. National Visualization and Analytics Center, 2005. 2