Facilitating insights with a user adaptable dashboard, illustrated by airport connectivity data

Ieva Dobraja,[a] Menno-Jan Kraak,[a] Yuri Engelhardt[a]

[a] *Faculty of Geo-Information Science and Earth Observation (ITC) of the University of Twente, Department of Geo-Information Processing, Enschede, The Netherlands, i.dobraja@utwente.nl; m.j.kraak@utwente.nl; yuri.engelhardt@utwente.nl*

Abstract: Since the movement data exist, there have been approaches to collect and analyze them to get insights. This kind of data is often heterogeneous, multiscale and multi-temporal. Those interested in spatio-temporal patterns of movement data do not gain insights from textual descriptions. Therefore, visualization is required. As spatio-temporal movement data can be complex because size and characteristics, it is even challenging to create an overview of it. Plotting all the data on the screen will not be the solution as it likely will result into cluttered images where no data exploration is possible. To ensure that users will receive the information they are interested in, it is important to provide a graphical data representation environment where exploration to gain insights are possible not only in the overall level but at sub-levels as well. A dashboard would be a solution the representation of heterogeneous spatio- temporal data. It provides an overview and helps to unravel the complexity of data by splitting data in multiple data representation views. The adaptability of dashboard will help to reveal the information which cannot be seen in the overview.

Keywords: Spatio-temporal movement data, global airline data, dashboard, question-driven exploration, adaptability

1. Introduction

It is well known that graphical representations of especially spatial data communicate better than textual descriptions. Therefore, gaining insights in data is not possible without visualization. However, the complexity of spatio- temporal data still makes it challenging.

There are strategies as well as methods and techniques developed to offer users an overview of data, which deal with the data complexity and clutter problem. Examples are Shneiderman's (1996): "Visual Information-Seeking mantra" following the Overview first, zoom and filter, then details-on-demand approach. Later this strategy has been adapted by Keim et al. (2006) as Visual analytics mantra: "Analysis First-Show the Important-Zoom, Filter and Ana- lyse Further-Details on Demand". Both approaches suggest to provide users first the general overview and then encourage for further explorations.

Following the strategies introduced by Shneiderman (1996) and Keim et al. (2006), and applying already developed methods to reduce data complexity and visual clutter, a summary can be realized in so-called dashboards. Few (2006) has defined the dashboard as "a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance". Dashboard provides an aggregated summary of data displayed in multiple representations. It visually identifies trends, patterns and anomalies and encourages the user to drill down to reach the information which is hidden due to aggregation. Similar to typical geovisual analytics environments dashboards consists of a combination of several linked views displaying a given types of visualization, for example, a map and / or diagrams. However, the

dashboard aims at "the most important" while a geovisual analytics exploratory functionality is key. Users may interact with the views, for example select, filter and zoom-in to explore and analyze subsets of data. As the dashboard's views are linked to each other, performing action in one of the view will make the changes will appear in other views as well. Dashboards could be a solution because of its function to show the overview with the most significant trends and patterns and the linked perspectives through different data representations in one environment. Traditionally dashboard have a fixed layout. In this research we have implemented a user query based adaptable appearance of the dashboard. The user will have the possibility to explore the data based on the three components of spatio-temporal data and choose the data representation method.

2. Related Work

The following section will give an overview about spatio-temporal movement data, methods to reduce visual clutter and data complexity, dashboard concept and graphic representations to visualize spatio-temporal data.

2.1 Spatio-temporal movement data

Spatio-temporal movement data can be characterized by three components – spatial (where), temporal (when) and attribute (what). Movement data consists of time (set of moments), a set of entities that move and space (a set of location that can be occupied by the entities). When analyzing movements, users are interested not only in the spatial and temporal changes of the movements but also in the characteristics of the moving object (Andrienko et al. 2008).

These components are related to each other and can be formed in a Spatiotemporal Triad scheme.

Based on the three components, users can formulate three types of basic questions related to each of the components. These questions will lead to information about the location of where the event took place, the time when it happened and information about the object that was present at a certain location and time (Peuquet 1994):

- when + where →what: describes the objects or set of objects that are present at a given location or set of location at a given time or set of times;
- when + what →where: describe the location or set of locations occupied by a given object or set of objects at a given time or set of times;
- where + what →when: describe the times or set of times that a given object or set of objects occupied a given location or set of locations.

The basic questions can vary in complexity, Bertin (1967) defined these as reading levels. There are three levels of reading – Elementary, Intermediate and Overall, depending if the question refers to an individual element of the data, to a subset of elements or to the whole data set. Bertin's approach of reading levels and Peuquet's approach of question types will be used to implement the question-driven exploration data exploration environment.

In comparison to other spatio-temporal data, entities of movement data change their position in space with respect to a reference system which in the most cases is geographical space.

2.2 Data Complexity and Visual Clutter

In the exploration of spatio-temporal movement data users are often interested in finding spatio-temporal patterns for further analysis. If the image is cluttered, no patterns can be distinguished and no exploration or analysis is possible. In addition, users may misunderstand the represented information (Boyandin 2013) (van den Elzen & van Wijk 2014). Spatio-temporal data complexity in visualization is related to its size (number of elements) and number of attributes (Kocherlakota & Healey 2005). In addition data can be complex due to the spatial and/or temporal distribution of the phenomena represented (Buschmann et al. 2016). This makes it challenging to create a clear overview, and may result in a cluttered image. Therefore, it is important to provide users with a clear summary of the data or offer options that allow exploration.

Rosenholtz et al. (2007) in the context of the design of a user interface for information visualization have defined clutter as: "The state in which excess items, or their representation or organization, lead to a degradation of performance at some task." In such a case, excess items or their representation may cause several problems such as crowding and occlusion, which decrease information perception. Based on the clutter definition by Rosenholtz et al. (2007), Berg et al. (2009) have stated that information density and layout are two factors that contribute to visual clutter. They suggested that visual clutter can be reduced by reducing the information density or by changing the visualization layout.

Researches by Rosenholtz et al. (2007) and Berg et al. (2009) were done on clutter on displays in general, not specially for spatio-temporal movement data.

Visualizing spatio-temporal movement data, clutter can appear both spatially (Fig. 1) or temporally (Fig. 2a). Fig. 1 shows spatial clutter. In this case, we observe lots of movements above Europe that makes the image locally cluttered and no further exploration and analysis is possible without applying clutter reduction methods. Fig. 2a shows temporal clutter at a certain time. If the temporal component is involved in the visualization, we can see that at a particular time intervals there are more movements than at other times. This can make the image cluttered. For example, in Fig. 2a we can see that some of the flows above the Central Europe and the North America overlap each other. In comparison, Fig. 2b shows no temporal clutter as there are only few flights over the chosen period of time.

Fig. 1. Spatial clutter example

a,

b,

Fig. 2. Temporal clutter example

Currently developed methods and techniques to reduce data complexity and visual clutter of movement data can be divided in three categories:

- Algorithms, for example, bundling, clustering and aggregation. Using bundling algorithm, flow lines which are close to each other, are rerouted to form a bundle (Fig. 3.) (Boyandin 2013) (Hurter et al. 2014);
- Representations, for example, animation of movements over time (Fig. 4.) or various 2D and 3D representations (Nollenburg 2007) (Andrienko et al. 2010);
- Environment, for example, interaction. In the overview only part of the details are displayed,

next users can specify which part of the data they would like to further explore more and receive more details (Fig.5) (Keim et al. 2006) (Hurter et al. 2009) (Boyandin 2013).
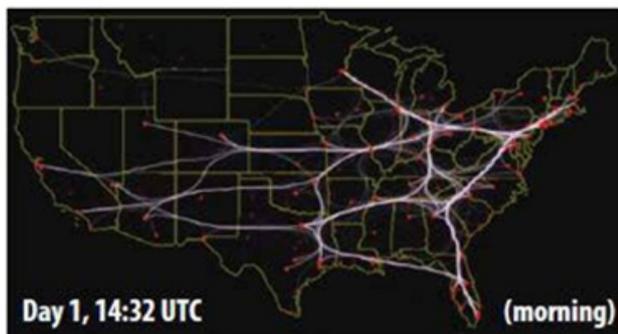

Fig. 3. Flow line clutter reduction by applying bundling (Source: Hurter et al. 2014)

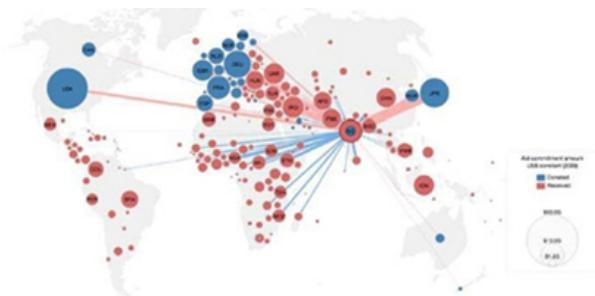Fig. 4. Animation of incoming and outgoing flights over time at Keflavik airport

Fig. 5. Visual clutter reduction by using interaction (Boyandin 2013)

2.3 Dashboard

A dashboard is a visual display that contains multiple graphic presentation. Its purpose is to communicate complex information by showing the summary of the data to gain an overview and encouraging the user for further exploration. (Few 2006) (Few 2013). It is essential to provide users with various data representation views because a single view conveys only part of the story, and the user should have the option to choose among several data representation as he or she may understand one data representation better that another (Keim et al. 2010) (Roberts 2000).

There are various graphic representations available. For example, a diagram and map. Both can appropriate to show the changes of the attribute over time. A diagram gives a shape to numbers and shows patters (Few 2006) and the maps tends to be suitable if a geographical component is required. The most popular map type to visualize movement is the flow map where the connections between an origin and a destination are displayed. Another example of diagram is parallel coordinate plot (PCP) which can be used to compare the profiles of several items. It is suitable for multiple variable visualizations (Lundblad & Jern 2013). All these data representations can be combined in a dashboard together with filters and highlighters for visualizing the data.

Fig.5 shows an example of a dashboard displaying information about flights to Schiphol airport. It shows a general dashboard that consists of a map, parallel coordinates plot and a bar graph. In addition, there is a time filter and a country highlighter available. The views, the filter and the highlighter are linked to each other.
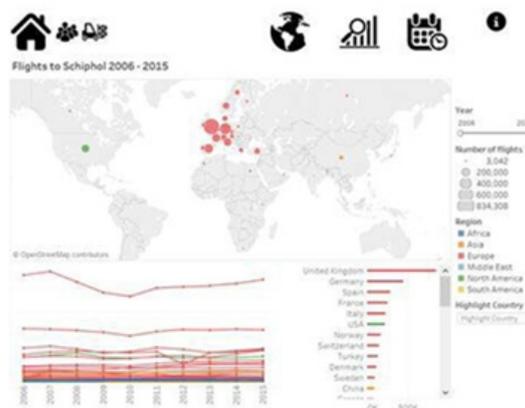

Fig. 5. Example of a dashboard (Source: Rahman 2017)

There are various ways of categorizing dashboards, but Few (2006) and Pappas & Whitman (2011) suggest to divide dashboards into three categories based on their role. These categories are listed below.

- Strategic – provides an overview of the parameters decision makers need to monitor to help achieve a successful development. This type of dashboard is built up from the information collected previously and there is no need for real-time data. In addition, no interactivity function is involved in this type of dash- board.

- Analytical – contains complex information about data; for example, various parameters and comparisons. An interactivity function for this type of dashboard is essential to support further analysis.

- Operational – requires simple display media to attract user's attention and to obtain a response in the case of an emergency. Interactivity may be involved to provide users with additional information regarding an emergency.

Our dashboard fit the analytical dashboard category. This type of the dashboard is the most suitable for displaying the summary of movement data as it allows to display complex information and supports interactivity that is essential to ensure the adaptability possibility.

In their research into dashboards Few (2006) and Pappas & Whitman (2011) mentioned requirements that have to be taken into account to create a successful dashboard.

First of all, displayed information has to fit on one screen without the necessity to scroll down to see all information. Such a screen helps to monitor all the information and gain an overall picture of the data. Therefore, if it is necessary to display a large amount of data, the main view can display the most important information in an aggregated summary view and other more detailed information can be displayed in linked views. Secondly, as the space on any screen is limited, unnecessary information or elements must be avoided. Instructions related to a dashboard usage can be displayed on request as they are not necessary all the time. Thirdly, there should be a greater focus on recent compared to historical information. This statement can be linked with suggestions introduced by Peuquet (1994) and Keim et al. (2006) regarding the spatial and temporal clutter. Furthermore, information should be organized in meaningful groups while avoiding fragmentation. For example, data representations regarding to one component has to be placed close to each other and must use the same color coding to avoid confusion.

Few (2006) in his research has introduced the concept of screen fragmentation based on a position of importance. As can be seen in Fig. 6, the upper-left corner is the most prominent position and the bottom-right corner is the least prominent. The information that requires immediate attention – such as important trends or patterns, has to stand out. It can be done by placing it in the upper-right corner or by displayed using different color intensity or size than other displayed information. In contrast, non-data elements that are necessary for a successful dashboard functionality can be located in a less prominent position or muted visually; for example, by applying a lighter color than for other elements.

Fig. 6. Positions on the dashboard according to their importance (Few 2006)

3. Case Study

Our approach is illustrated by a subset of a global airport and air traffic database that is used by researchers interested in the connectivity of airports. Our case study concerns Iceland's Keflavik (KEF) airport. KEF is located south- west from the capital Reykjavík and plays an important role as the hub airport linking two continents.

The data is supplemented by the airport's annual reports. The data includes information on the airports that can be reached from KEF with a direct flight or via a hub airport, the capacity of the individual flights, as well as

the type of airplanes and the airlines, the arrival and departure times, and the duration of the flights for each week.

4. Question-driven Approach for Spatio-Temporal Pattern Exploration

Researchers interested in airport connectivity have formulated questions at different levels of complexity. These levels can be referred to Bertin's (1967) levels of reading. Users questions will always involve the component from the Spatio-temporal triad (where, when or what) or a combination of them. These questions require and help to explore the data sub-levels which cannot be seen in the overview.

The default layout of the adaptable dashboard has views showing spatial, temporal, and attribute-related trends. This is assumed as Overall level of reading. The spatial view shows the airline networks and temporal view might show the passenger arrival distribution over the year showing a peak in summer. In the Overall level user may have the following questions:

- What is the overall passenger trend? (attribute and time)
- Which airports can be reached from KEF airport? (space)

Zooming-in into further subsets of dataset helps to discover patterns at Intermediate or Elementary level.

A researcher might like to zoom in and want to know what the daily trends are in direct connections to the origin and destination of the flights. This level is assumed as the Intermediate level of reading where the user might have following questions:

- What are the destinations and origins of flights at certain hours? (space and time)
- What are the origins and destinations of flights at morning/afternoon/evening hours? (space and time)
- At what time, there is no flights? (time and attribute)
- What are the incoming and outgoing flight patterns in morning and evening? (attribute, time and space)
- At what time of the day there are more passengers arriving/departing? (attribute and time)

In the Elementary level of reading user might ask:

- Which airline connects airport X with KEF? (space and attribute)
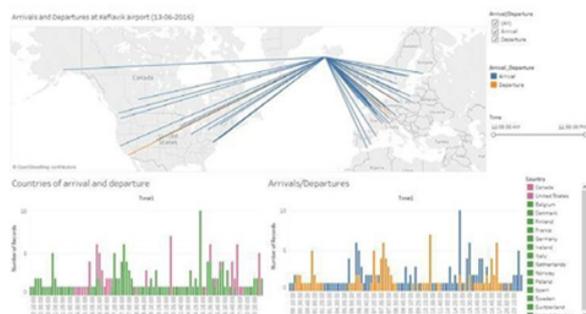- What are the incoming and outgoing flights at X hour? (space and time)

Fig. 7. Dashboard view example

Zooming into levels of reading, will change the layout of the dashboard, replace views, and emphasize the data representation regarding the component. For example, if the user is interested in space related questions, the map will be emphasized. For time related questions time line or graph displaying time will be emphasized.

In a follow-up question the researcher may be interested in the airlines and kind of passengers arriving. This again would swap some of the views and result in different visualizations for further interpretation.

5. Conclusions

An adaptable dashboard will allow users to explore data at three levels of complexity from various perspectives, depending on the chosen component. In addition to the functions of a traditional dashboard, adaptable dashboard will switch views depending on the user's questions and allows to gain insights on various levels of complexity of spatio- temporal components.

6. References

Andrienko, G. et al., 2010. Space, time and visual analytics. International Journal of Geographical Information Science, 24(10), pp.1577–1600. Andrienko, G. et al., 2013. Visual Analytics of Movement, Springer Berlin Heidelberg. Available at: http://www.springer.com/us/book/9783642375828.

Andrienko, N. et al., 2008. Basic concepts of movement data. Mobility, Data Mining and Privacy: Geographic Knowledge Discovery, pp.15–38. Berg, R. van den, Cornelissen, F.W. & Roerdink, J.B.T.M., 2009. Perceptual Interactions in Human Vision and Implications for Information Visualization. Journal of Vision, 9(4), pp.1–11. Available at: https://www.rug.nl/research/portal/.../thesis.pd.

Bertin, J., 1967. Semiologie Graphique: les diagrammes, les réseaux, les cartes, La Haye: Mouton.

Boyandin, I., 2013. VISUALIZATION OF TEMPORAL ORIGIN - DESTINATION DATA. University of Fribourg.

Buschmann, S., Trapp, M. & Döllner, J., 2016. Animated visualization of spatial–temporal trajectory data for air-traffic analysis. The Visual Computer, 32(3), pp.371–381. Available at: http://link.springer.com/10.1007/s00371-015-1185-9 [Accessed February 20, 2017].

van den Elzen, S. & van Wijk, J.J., 2014. Multivariate network exploration and presentation: From detail to overview via selections and aggregations. IEEE Transactions on Visualization and Computer Graphics, 20(12), pp.2310–2319.

Few, S., 2013. Information Dashboard Design: Displaying data for at-a-glance monitoring, Burlingame: Analytics Press.

Few, S., 2006. Information dashboard design. The effective visual communication of data, O'Reilly Media, Inc. Available at: http://books.google.com/books?id=qWER8Im-WYIC&pgis=1.

Hurter, C. et al., 2014. Bundled visualization of dynamic graph and trail data. IEEE Transactions on Visualization and Computer Graphics, 20(8), pp.1141–1157.

Hurter, C., Tissoires, B. & Conversy, S., 2009. FromDaDy: Spreading Aircraft Trajectories Across Views to Support Iterative Queries. IEEE Transactions on Visualization and Computer Graphics, 15(6), pp.1017–1024.

Keim, D. et al., 2010. Mastering the Information Age Solving Problems with Visual Analytics, Goslar: Eurographics Association. Available at: http://diglib.eg.org.

Keim, D. a. et al., 2006. Challenges in Visual Data Analysis. Tenth International Conference on Information Visualisation (IV'06), pp.9–16. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?ar number=1648235.

Kocherlakota, S.M. & Healey, C.G., 2005. Summarization Techniques for Visualization of Large Multidimensional Datasets,

Lundblad, P. & Jern, M., 2013. Geovisual analytics and storytelling using HTML5. In Proceedings of the International Conference on Information Visualisation. pp. 263–271.

Nollenburg, M., 2007. Geographic Visualization. Human-Centered Visualization Environments, pp.257–294.

Peuquet, D.J., 1994. It's about Time : A Conceptual Framework for the Representation of Temporal Dynamics in Geographic Information Systems. Annals of the Association of American Geographers, 84(3), pp.441–461. Available at: https://www.jstor.org/stable/pdf/2563777.pdf.

Rahman, A., 2017. Designing a Dashboard as Geo-Visual Exploration Tool for Origin-Destination Data. The University of Twente.

Roberts, J.C., 2000. Visualization Display Models – ways to classify visual representations. International Journal of Computer Integrated Design and Construction, 2(4), pp.241–250. Available at: http://kar.kent.ac.uk/21931/ [Accessed February 27, 2017].

Rosenholtz, R., Li, Y. & Nakano, L., 2007. Measuring visual clutter. Journal of vision, 7(2), p.17.1-22.

Shneiderman, B., 1996. The eyes have it: a task by data type taxonomy for information visualizations. Proceedings 1996 IEEE Symposium on Visual Languages, pp.336–343.