A Matter of Time: Applying a Data–Users–Tasks Design Triangle to Visual Analytics of Time-Oriented Data

Silvia Miksch and Wolfgang Aigner

Institute of Software Technology and Interactive Systems, Vienna University of Technology, Austria

Abstract

Increasing amounts of data offer great opportunities to promote technological progress and business success. Visual Analytics (VA) aims at enabling the exploration and the understanding of large and complex data sets by intertwining interactive visualization, data analysis, human-computer interaction, as well as cognitive and perceptual science. We propose a design triangle, which considers three main aspects to ease the design: (1) the characteristics of the data, (2) the users, and (3) the users' tasks. Addressing the particular characteristics of time and time-oriented data focus the VA methods, but turns the design space into a more complex and challenging one. We demonstrate the applicability of the design triangle by three use cases tackling the time-oriented aspects explicitly. Our design triangle provides a high-level framework, which is simple and very effective for the design process as well as easily applicable for both, researchers and practitioners.

Keywords: Visual Analytics, interactive visualization, interaction design, temporal data mining, time-oriented data

1. Introduction

Due to the proliferating capabilities to generate and collect vast amounts of heterogeneous data and information we face the challenge that users and analysts get lost in irrelevant or inappropriately processed or presented information. The aim of Visual Analytics (VA) is to support the information discovery process from potentially large volumes of complex and heterogeneous data and information [1, 2]. Time itself is an exceptional data dimension with distinct characteristics (e.g., scale, scope, structure, viewpoint, and granularity [3]). In order to design and develop VA methods that effectively deal with the complexity of time, these special characteristics need to be considered within the intertwined interactive visualization and automated analysis process to explore trends, patterns, and relationships. Furthermore, to facilitate the design process, it is crucial to consider (1) the characteristics of the *data*, (2) the *users*, and (3) the users' *tasks* (cmp. Figure 1).

In the following, we first elaborate on this design triangle that is intended to ease the design process and covers data, users, and their tasks. Based on this, we illustrate its applicability by three examples from our own research.

2. Design Triangle: Data–Users–Tasks

The design space for VA solutions is manifold, because there is usually already a plethora of potential interactive visualization and automated analysis techniques available. Often it is not clear to the designer(s), which techniques eventually will lead to the intended results. We guide the designers in providing a design triangle framework, which addresses three main questions:

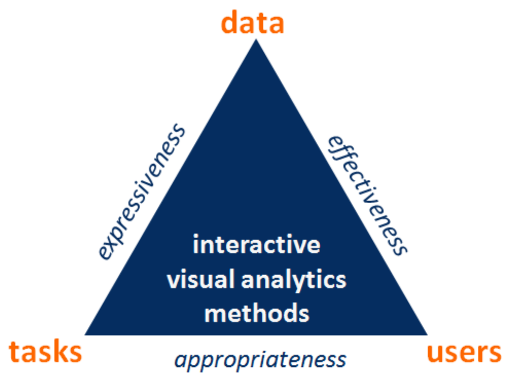

Figure 1: Design Triangle: Data–Users–Tasks. Major factors to be considered during the design and implementation of interactive VA methods.

- What kinds of data are the users working with? (*data*)
- Who are the users of the VA solution(s)? (*users*)
- What are the (general) tasks of the users? (*tasks*)

The answers to these questions largely determine which visual representation, analytical, and interaction methods are suitable. Along the sides of the design triangle shown in Figure 1, some of the major quality criteria for visualization are listed. These need to be satisfied in order to obtain useful results. *Expressiveness* refers to the requirement of showing exactly the information contained in the data; nothing more and nothing less must be visualized [4]. *Effectiveness* primarily considers

URL: `www.cvast.tuwien.ac.at` (Silvia Miksch and Wolfgang Aigner) the degree to which visualization addresses the cognitive capabilities of the human visual system, but also the task at hand, the application background, and other context-related information, to obtain intuitively recognizable and interpretable visual representations [4]. Finally, *appropriateness* involves a cost-value ratio in order to assess the benefit of the visualization process with respect to achieving a given task [5].

Data: Characterizing Data and Time. It is of fundamental importance to consider the characteristics of the data to design appropriate VA solutions. The available data characterization and modeling approaches are manifold and range from considering continuous to discrete data models (see [6]). According to Aigner et al. [3], we can distinguish between *scale* (quantitative vs. qualitative), *frame of reference* (abstract vs. spatial), *kind of data* (events vs. states), and *number of variables* (univariate vs. multivariate).

Time itself is an inherent data dimension that is central to the tasks of revealing trends as well as identifying patterns and relationships in the data. Time and time-oriented data have distinct characteristics that make it worthwhile to treat such data as a separate data type [3, 7]. Due to the importance of time-oriented data, its design aspects have been studied in numerous scientific publications (such as [3, 8, 9]). Aigner et al. [3] divided the aspects of time-oriented data into *general aspects* required to adequately model the time domain as well as hierarchical organization of time and definition of concrete time elements, also called *human-made abstractions*. We use the most important aspects in our design triangle: *arrangement* (linear vs. cyclic) and *time primitives* (instant vs. interval vs. span). *Linear time* corresponds to an ordered model of time, i.e., time proceeds from the past to the future. *Cyclic time domains* are composed of a finite set of recurring time elements (e.g., the seasons of the year). Basically, time primitives can be divided into anchored (absolute) and unanchored (relative) primitives. *Instant* and *interval* are primitives that belong to the first group, i.e., they are located on a fixed position along the time domain. In contrast to that, a *span* is a relative primitive, i.e., it has no absolute position in time and is a duration (of intervals).

Users. As VA is inherently user-driven, *meeting users' needs* is vital. In order to achieve this, a thorough understanding of users' needs, their tasks, and their work environment is needed. However, this is not easy to accomplish, since typical VA problems are not well defined, often ad-hoc, and far from routine tasks. Users can be described in many ways. On one level, user groups can be characterized along their specific context such as their application domain (e.g., health-care or environmental research), the physical environment (e.g., poor lighting), social factors (e.g., collaborative work or cultural specifics), as well as technical specifics (e.g., hardware, screen resolution). For example, in case of health-care, VA methods need to address the specific needs of health-care providers, physicians, nurses, patients, clinical researchers, or quality improvement analysts in different dimensions. In case of environmental research there are environmental analysts, managers, and government representatives. On another level, we need to consider individual factors such as the level of technical and domain expertise (e.g., experts, apprentices, or novices), specific metaphors and mental models that are used, or disabilities (e.g., color-blindness).

Tasks. Supporting effective, expressive, and appropriate VA methods requires an understanding of what particular tasks need to be carried out during the analytical reasoning process. One of the big challenges is to gain an understanding of the rich set of possible tasks during such processes. While there are a number of task taxonomies and typologies such as [10, 11, 12], there are only a few taxonomies available specifically for time-oriented visual data analysis [7, 13]. MacEachren [13] proposed a low-level task description specifically addressing the temporal domain. The tasks are defined by a set of important questions that users might seek to answer with the help of visual representations, like *existence of data element:* does a data element exist at a specific time? or *rate of change:* how fast is a data element changing over time? The task typology by Andrienko et al. [7] distinguishes between elementary and synoptic tasks on the first level. *Elementary* tasks address individual data elements. This may include individual values, but also individual groups of data. *Synoptic* tasks, on the other hand, involve a general view and consider sets of values or groups of data in their entirety. For example, elementary tasks are *direct lookup:* what is the value of glucose on March 21, 2013? or *direct comparison:* compare the value of glucose and the activity level on March 21, 2013. An example of synoptic tasks is *relation seeking:* find two contiguous months with opposite trends in the values of glucose.

In the next section we illustrate how this design triangle is applied in three different research projects, CareCruiser [14], ViENA [15], and TiMoVA [16].

3. Applying the Design Triangle: Three Examples

CareCruiser: Exploring the Effects of Medical Treatment Plans on a Patient's Condition. Medical treatment plans are recommendations for specific clinical situations containing precise instructions of clinical actions (e.g., the administration of a certain drug). The complex nature of these treatment plans calls for a plain and compact visualization of the underlying information. Furthermore, assessing the quality of such treatment plans is crucial to improve these plans. In particular, for the medical experts, there is a need to investigate the actual effects of the application of different treatment plans and clinical actions to optimize the choice of treatment actions.

Following the corners of the design triangle (Figure 1), we have identified the corresponding requirements:

- *Data*: patient data and treatment plans: data are quantitative, abstract, and multivariate, handling events and states; the time arrangement is linear and instants are used as time primitives.
- *Users*: medical experts and physicians.
- *Tasks*: exploring the effects of clinical actions on a patient's conditions.

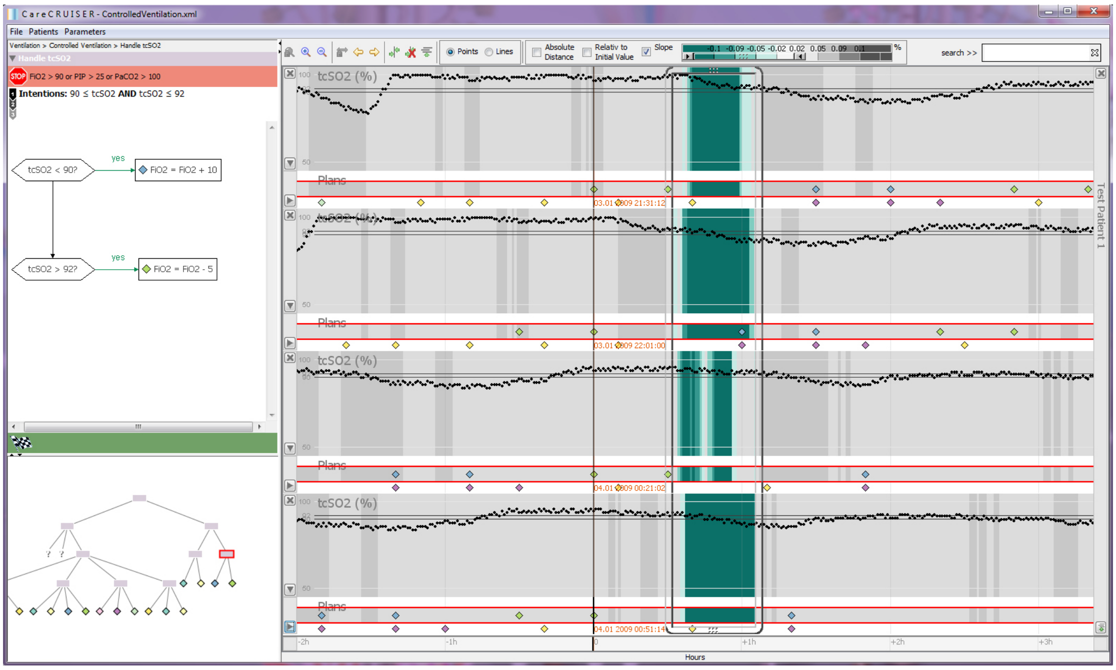

Figure 2: CareCruiser [14]. The temporal view (on the right-hand side) arranges patient parameters together with applied clinical actions (the little diamonds below the line charts) along a horizontal time axis. The logical and hierarchical views capturing the treatment plans are shown on the left-hand side.

CareCruiser [14] uses different visualization techniques to represent relevant characteristics and temporal constraints of computer-interpretable treatment plans in combination with time-oriented patient data. It provides three different views to communicate specific information: (1) the logical view to visualize the logic of steps within treatment plans, (2) a hierarchical view to show the nested structure of these plans, and (3) the temporal view to represent time constraints of treatment plans as well as the progress of patient parameters over time (Figure 2). CareCruiser provides several visual and interactive features to support a step-wise exploration: aligning clinical actions, color-coding curve events, filtering color-coded information, and a focus & context window for the detection of patterns of effects. Moreover, CareCruiser allows the comparison of multiple patients at the same time. Using color-coded distance information and slopes, CareCruiser supports the investigation of the effects of applied treatment plans on the condition of each patient.

CareCruiser was designed in close cooperation with physicians. We conducted a heuristic usability evaluation and interviews with medical experts using two different treatment plans about ventilating infants together with five data sets of real patient data [17]. We could show that CareCruiser meets the aimed quality criteria (sides of the design triangle as shown in Figure 1) for visualization and exploration. It provides means (1) to assess success or failure of previously applied treatment plans, (2) to explore the effects of each applied clinical action on the patient's conditions, and (3) to identify sub-optimal treatment choices to the physicians.

ViENA: Visual Enterprise Networks Analytics. Collaboration is the way work gets done in organisations. Social networks of different types, functions, and compositions have become an inevitable precondition of organizational performance. However, these networks are not static, but changing over time. Therefore, organisational analysts are seeking for appropriate methods to explore organizational changes within these dynamic networks.

Following the corners of the design triangle (Figure 1), we have identified the corresponding requirements:

- *Data*: temporal network slices with additional information, like organizational information (roles, skills, processes, etc.): data are both quantitative and qualitative, abstract, and multivariate, recording states; and the time arrangement is linear and the used time primitives are instants.

- *Users*: network experts, consultants, and business management users.

- *Tasks*: exploring organizational changes over time.

Within the applied research project ViENA (Visual Enterprise Network Analytics [15]), we designed and developed a VA prototype. It features three views: (1) a juxtaposition view, enabling the direct comparison of network structures over time, (2) a superimposition view, enabling the tracking of node trajectories, and (3) a two-and-a-half-dimensional view, aimed for synoptic tasks (Figure 3). Besides smoothly animated transitions between the views, specific interaction techniques ease the exploration of dynamic networks: a control of the layout stability to tune the mental map preservation amount, a dual-mode highlighting technique for tracking single nodes including their neighbours, and a vertigo zoom [18] to switch between

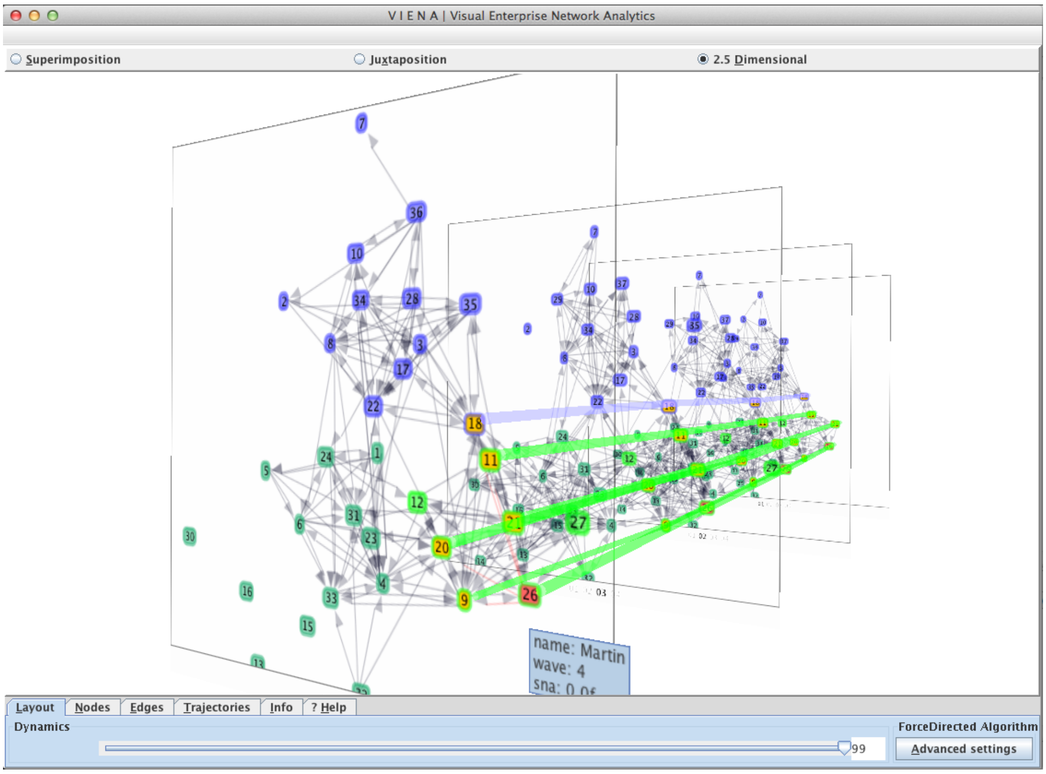

Figure 3: ViENA [15]. 2.5D visualization of dynamic networks. Traces represent change of network metric for person nodes over time.

relational and temporal perspectives. The computation of static network measures as well as of novel dynamic measures enriches the visualization and optimizes dynamic layouts.

We conducted two user-studies jointly with HCI experts to assess the utility of our prototype: a preliminary mock-up assessment and a final task-based qualitative evaluation [19]. The evaluation elicited few weaknesses and confirmed most of the design choices, showing how efficiently and effectively the appropriate combination of visual, interactive, and analytical techniques support complex tasks, according to different users' problem-solving strategies.

TiMoVA: VA for Model Selection in Time Series Analysis. Statistical time series analysis is a challenging task performed by experts in different domains, for example, a public health official predicting the number of people that need to be treated because of cardiovascular reasons in the next year. Usually, model selection is a cumbersome task.

Following the corners of the design triangle (Figure 1), we have identified the corresponding requirements:

- *Data*: time series in any domain (e.g., in economics or environmental research); data are quantitative, abstract, and univariate, handling states; the time arrangement is linear and the used time primitives are instants.
- *Users*: domain experts in any field applying time series analysis, experts in time series analysis.
- *Tasks*: time series model selection.

TiMoVA (Time series analysis, Model selection, applies VA methods) [16] provides an interactive exploration environment to guide users in the time series model selection processes. TiMoVA enables the domain expert to (1) select the model order interactively via the visual interface, (2) give the domain expert immediate visual feedback of the model results while selecting the model order, and (3) help domain experts with the visualization of the model transitions to decide whether or not the model improves (Figure 4).

TiMoVA was evaluated by usage scenarios with an example data set from epidemiology and interviews with two external domain experts in statistics [16]. It could be shown that the interactive visual exploration environment with short feedback cycles supports the model selection tasks in an expressive way.

4. Conclusion

We proposed a design triangle capturing the characteristics of the *data*, the *users*, and their *tasks*, which guides the design of VA methods tackling in particular time and time-oriented data. It provides a high-level framework for VA design, which is simple and very effective for the design process. As a consequence, it is easily applicable for researchers as well as practitioners.

TimeViz Browser. In collaboration with the University of Rostock we have built an interactive repository (`http://survey.timeviz.net`) that helps researchers and practitioners in finding available approaches that fit these requirements. This is

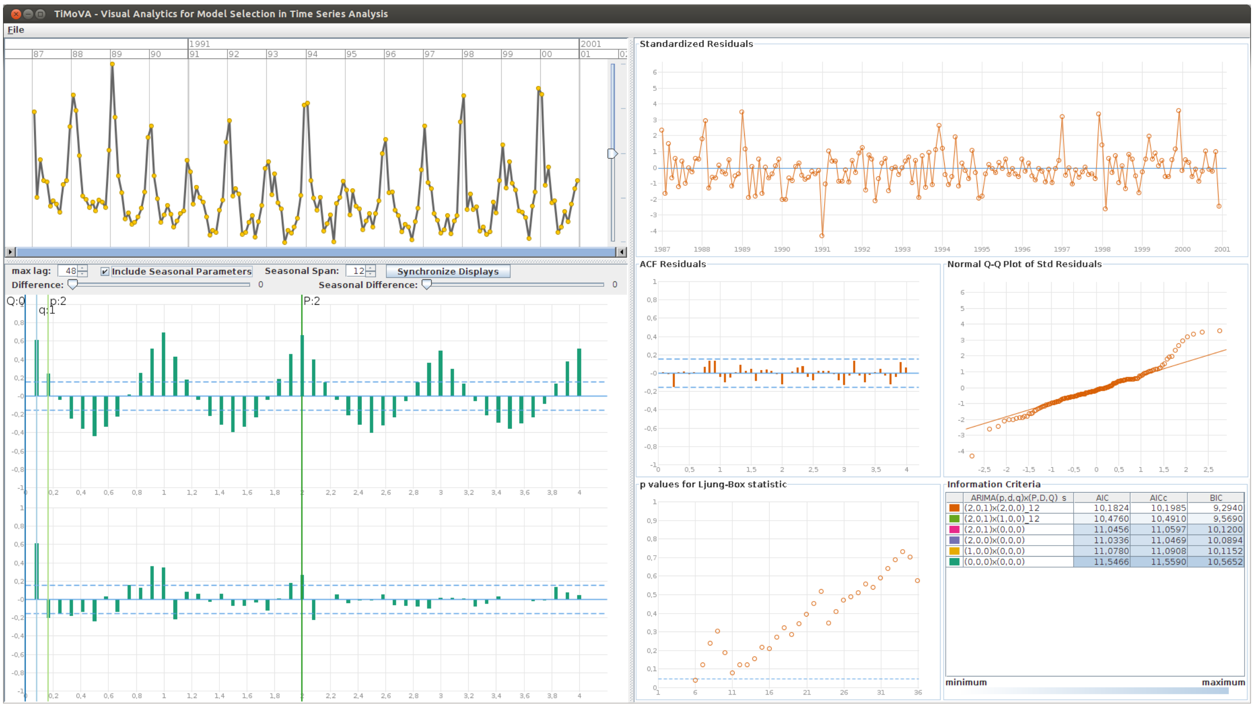

Figure 4: TiMoVA [16]. The interface shows the time series plot (input data; on the upper left), the model selection toolbox, and various other views to guide the model selection process.

achieved via faceted browsing where characteristics of time and data can be specified interactively and a number of applicable techniques is shown including a screenshot, short textual description, and references.

Acknowledgments. This work was supported by the Austrian Science Fund (FWF) through HypoVis (No. P22883), the Austrian Federal Ministry of Economy, Family and Youth via CVAST, a Laura Bassi Centre of Excellence (No. 822746), by the Austrian Federal Ministry of Transport, Innovation, and Technology (FIT-IT initiative) through ViENA (No. 820928), and the Vienna University of Technology through the Doctoral College for Environmental Informatics.

References

[1] Thomas JJ, Cook KA, editors. Illuminating the Path: The Research and Development Agenda for Visual Analytics. IEEE; 2005.
[2] Keim D, Kohlhammer J, Ellis G, Mansmann F, editors. Mastering The Information Age - Solving Problems with Visual Analytics. Goslar, Germany: Eurographics; 2010.
[3] Aigner W, Miksch S, Schumann H, Tominski C. Visualization of Time-Oriented Data. London: Springer; 2011.
[4] Mackinlay J. Automating the Design of Graphical Presentations of Relational Information. ACM Transactions on Graphics 1986;5(2):110–41.
[5] Van Wijk JJ. Views on Visualization. IEEE Transactions Visualization and Computer Graphics 2006;12(4):421–33.
[6] Tory M, Möller T. Rethinking Visualization: A High-Level Taxonomy. In: Proceedings of the IEEE Symposium on Information Visualization (InfoVis). Los Alamitos, CA, USA; 2004, p. 151–8.
[7] Andrienko N, Andrienko G. Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach. Berlin: Springer; 2006.
[8] Bettini C, Jajodia S, Wang XS. Time Granularities in Databases, Data Mining, and Temporal Reasoning. 1st ed.; Secaucus, NJ, USA: Springer; 2000.
[9] Frank AU. Different Types of “Times” in GIS. In: Egenhofer MJ, Golledge RG, editors. Spatial and Temporal Reasoning in Geographic Information Systems. New York, NY, USA: Oxford University Press; 1998, p. 40–62.
[10] Shneiderman B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: Proceedings of the IEEE Symposium on Visual Languages. Los Alamitos, CA, USA; 1996, p. 336–43.
[11] Brehmer M, Munzner T. A multi-level typology of abstract visualization tasks. IEEE Transactions Visualization and Computer Graphics 2013;19(12). Forthcoming.
[12] Schulz HJ, Nocke T, Heitzler M, Schumann H. A design space of visualization tasks. IEEE Transactions Visualization and Computer Graphics 2013;19(12). Forthcoming.
[13] McEachren AM. How Maps Work: Representation, Visualization, and Design. New York, NY, USA: Guilford Press; 1995.
[14] Gschwandtner T, Aigner W, Kaiser K, Miksch S, Seyfang A. CareCruiser: exploring and visualizing plans, events, and effects interactively. In: Proceedings of the IEEE Pacific Visualization Symposium (PacificVis). 2011, p. 43–50.
[15] Federico P, Aigner W, Miksch S, Windhager F, Zenk L. A visual analytics approach to dynamic social networks. In: Proceedings of the 11th International Confonverence Knowledge Management and Knowledge Technologies (i-KNOW). ACM; 2011, p. 1–8.
[16] Bögl M, Aigner W, Filzmoser P, Lammarsch T, Miksch S, Rind A. Visual analytics for model selection in time series analysis. IEEE Transactions Visualization and Computer Graphics 2013;19(12). Forthcoming.
[17] Gschwandtner T, Aigner W, Kaiser K, Miksch S, Seyfang A. Design and evaluation of an interactive visualization of therapy plans and patient data. In: Proceedings of the 25th BCS Conference on Human-Computer Interaction (HCI). Newcastle upon Tyne, UK; 2011, p. 421–8.
[18] Federico P, Aigner W, Miksch S, Windhager F, Smuc M. Vertigo zoom: combining relational and temporal perspectives on dynamic networks. In: Proceedings of the 11th International Working Conference on Advanced Visual Interfaces (AVI). ACM; 2012, p. 437–40.
[19] Smuc M, Federico P, Windhager F, Aigner W, Zenk L, Miksch S. How do you connect moving dots? insights from user studies on dynamic network visualizations. In: Huang W, editor. Handbook of Human Centric Visualization. Springer; 2014, p. 623–50.