Interactive Visualization with User Perspective: A New Concept

Quang Vinh Nguyen
School of Computing and Mathematics
University of Western Sydney
Australia

vinh@scm.uws.edu.au

Simeon Simoff
School of Computing and Mathematics
University of Western Sydney
Australia

S.Simoff@uws.edu.au

Mao Lin Huang
Faculty of Engineering and IT
University of Technology, Sydney
Australia

maolin@it.uts.edu.au

ABSTRACT
With an astonishing amount of data generated for processing on a daily basic, it is essential to provide an effective methodology for understanding, reasoning and supporting decision making of large information spaces. This paper presents a new concept that provides an intelligent and interactive visualization in supporting large scale analysis. This aims to provide a much greater flexibility and control for the users to interactively customize the visualizations according to their preferences. A simple prototype is also presented to demonstrate the concept on hierarchical structures.

Categories and Subject Descriptors
H.1.2 [**User/Machine Systems**], H.5.2 [**User Interfaces**].

Keywords
Interactive Visualization, Visual Analytics, Intelligence, Hierarchical Data, Visualization Model.

1. INTRODUCTION
Our knowledge-driven technology-mediated environments generate a vast amount of data that need to be dealt with on a daily basis. Acquiring such an understanding is a difficult problem, especially those huge datasets from real-world sources. For the last few years, there has been a rapidly increasing interest in modelling and understanding such systems and processes through network models [1]. Data analysis brings a lot of benefits in a wide range of application domains, such as *economics/finance analysis, homeland defence, knowledge management, organisation development, social science, and health informatics*. For example, analysing a social network analysis for large cooperate organisations could identify the unofficial organisational roles, such as central connector, boundary spanner, information broker and peripheral specialists.

One of a promising approach to analyse large and complex data sets is *Visual Analytics* which is formally defined as the science of analytical reasoning facilitated by the interaction with visual representation of abstract data [2]. Technically, visual analytics combines automated analysis techniques (such as data mining, statistical analysis, compression and filtering) with interactive visualizations, analytical reasoning and other related disciplines to achieve effective understanding, reasoning and decision making. There are a number of good techniques that are quite capable of visualizing large graphs or networks of thousands of nodes and edges. Unfortunately, most of the available techniques often rely on computer-generated visualizations that limit the involvement and contribution of the domain experts in the analysis processes. Scalability and effectiveness are key challenges as they determine the ability to process large data sets because of computational overhead, appropriate layout and rendering techniques. As the fast growth of information, it is crucial to develop a better methodology and techniques that can cope with high scalability of data volumes and data dimensionality.

One of the best approaches to tackling the above problem is to utilise both human and computer strengths with a mixed-initiative visualization where both can collaborate in the exploration and creation of knowledge. The interactive visualization helps the human analyst to gain knowledge of the data through our powerful human visual perception and reasoning skills, ideally driving the system toward more focused and more adequate analytical techniques. Visualization should provide not only a simplified abstract view of the entire data set, but also display in detail the information at a particular focus point. Dually, the visualization should also allow analysts to interact and explore in order to make further discovery and to gain the insight of the information. From this, they can obtain a better understanding of the data and the structures so that they can contribute their domain expertises into the knowledge discovery process. In other words, the analysts can evaluate and refine automated analysis as well as interactive visualization processes.

Large-scale visual analytics could be enabled by intelligent and interactive visualizations whose layouts and interactions are responded to analyst preferences, behaviours, and other constraints. As a common property of fractal, the characteristic of a small set of data could be similar to a much larger one. Therefore, the use of supervised and un-supervised learning-based classification in data visualization would be a solution for analysing a huge data set by examining a small and manageable one. In addition, machine learning and logic based deductive reasoning can be used as an underlying layer to assist human analysts to make sense, discover knowledge and insight as well as validate the effectiveness. The development of new intelligent interfaces should replace the algorithm-specific control with simple customizable interfaces supporting by an invisible layer of complex intelligent algorithms [5]. Machine learning and artificial intelligence can be used to learn visualization tasks via user behaviour and preference. This could provide a new reasoning behind the learning process for further refine and optimize the visualization.

2. RELATED WORK

Effective integration of human and computer strengths is still one of the open topics in the research community. The cooperation of both the user and the computer in visual analytics was initially proposed by Ankerst in his PhD thesis in 2000 [3]. The author introduced a user-centred approach to decision tree construction that visualizes training data used for classification. An analyst could interact directly, search for algorithms, provided adequate data visualization and thus would have a deeper understanding about the information. In other words, the user provided the domain knowledge and evaluated the algorithms while the computer created the layout to satisfy user's constraint. This cooperation was ranging from completely manual over combined to completely automatic classification. Although this idea was very novelty and was high-potential applicable in the large-scale data analytics, the work has not been investigated further since then.

There is a growing interest in the development of the intelligent interface using both computer and human strengths for the analysis, especially at the *Annual IEEE Symposiums on Visual Analytics Science and Technology*[1]. Ma expressed his viewpoints of the importance of machine learning to boost the next generation of visualization [4]. Although some applications were used to illustrate the idea in the paper, it was technically somewhat not concrete and the author did not concern about interaction in his works. Gard et al presented a visual analytics infrastructure that used logic programming for analyst-guided discovery from high-dimensional visual interfaces [5]. From this platform, the analysts could supply patterns from the visual interface so that the system can learned automatically. Fuchs et al also presented a similar work for multivariate data exploration that provided a heuristic search algorithm for generating related features and hypotheses [6]. Works for combining automated analysis and visualizations have been also applied to high-dimensional data based on *Scatterplots* and *Parrallel Coordinates* Visualization [7]. Other interesting related works are the human cognition model [8] and the entity-based collaborative analytical tools [9]. Both techniques provided guidelines for visual analytics designs and collaborative intelligence analysis respectively. Further related works and concept in visual analytics can also be found at Simoff, Bohlen and Mazeika's book [10].

Although the above related works have shown their potentiality in some applications, most of these works are at their early stages and/or limit within specific domains. There are still no significant contribution and concept that provide effective intelligent and interactive visualization which is mentioned in this paper.

3. A MODEL OF INTELLIGENT INTERFACES FOR KNOWLEDGE DISCOVERY

The authors believe that intelligent interfaces will play an important role in the knowledge discovery processes. Data visualization will no longer provide only the static displays of the information with limited interactions. The future visualization and visual analytics techniques should work as robust platforms for collaborative knowledge sharing, discovery, data manipulation and feedback response among analysts. Figure 1 shows our proposed model of visual data exploration for knowledge discovery using intelligent interfaces. Technologically, we use rules and classifications as the underlying layer to enable the intelligent, portable and customisable interactive visualizations.

- **Machine learning using rules and classifications**: the rules and the hypotheses for producing new visualization can be learned automatically or can be specified manually by the user during the exploration based on his/her behaviour and preference. The captured rules and hypotheses can be recorded using a visualization language. With this, they can be portable so that the analysts can use the outcomes with the similar applications or with different data sets. This would open a new way for visualizing and analysing very large scale data sets because it is required to process on smaller training data sets and the results can be generalised to larger scale. The portability also utilise the power of multi-analysts who involving in the analysis.

- **Collection of Layout and Navigation algorithms**: computer-generated layouts play an important role in the interactive visualization because they can produce the layout of a large data set within a limited amount of time. It could take days to process or it is even impossible for a human to produce a similar readable display. In addition, preliminary result from usability evaluations has shown people tend to prefer computer-generated layouts to hand-drawn ones [11]. The collection of layout and navigation algorithms would effectively enable the system to provide a suitable output corresponding to nature of the data sets and user preference. The list of algorithms can be updated and refined through the interaction.

- **Interactive visualization**: using the collection of layout and navigation algorithms with the constraint of the rules and the hypotheses, the interactive visualization could provide the most suitable display for a data set as well as for an analyst's preference. In other words, it should create the layout and the interaction based on the nature of applications, data sets, analyst's preference and underlying sense-making rules. The innovation also lies in the creation of an effective mechanism to integrate different interactive visualization techniques so that it can provide a best-possible display corresponding to an analyst preferences and behaviour.

The use of portable visualization language would provide a medium for deploying pervasive visual analytics environments. This environment also enables the reusability of knowledge and findings to a new application by transferring the rules, hypotheses, and layout and navigation algorithms amongst the analysts. The portability could bring widely distributed benefits from the visual discovery to a variety of analysts. This also allows sharing their employed reasoning strategies while interacting with analytic environments.

[1] http://vis.computer.org

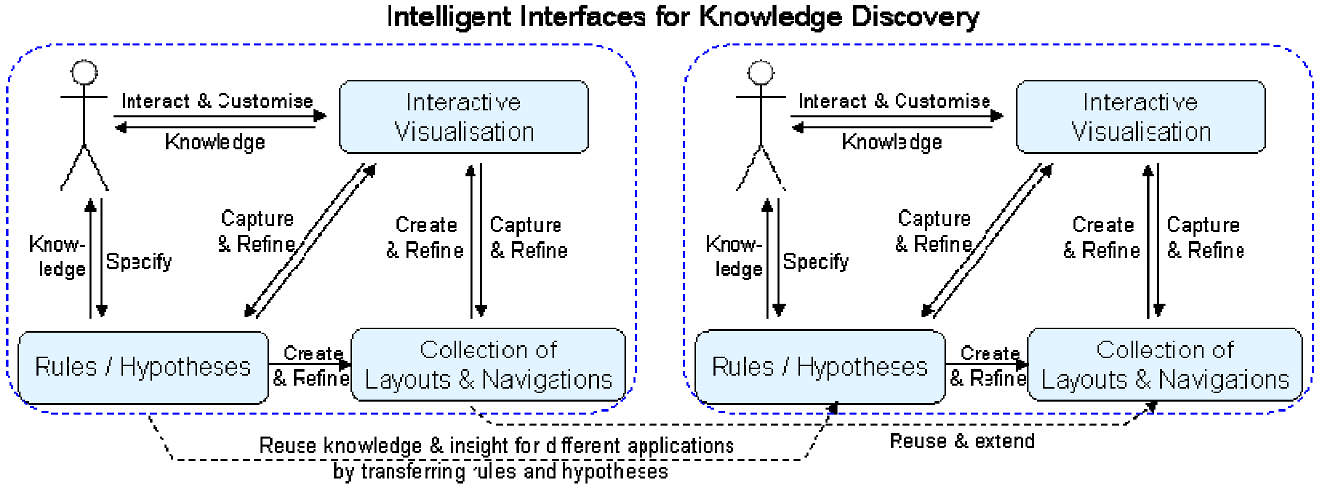

Figure 1. A Model of the Concept of Visual Data Exploration for Knowledge Discovery

4. VISUALIZATION

We have developed a small prototype to illustrate our new concept of the intelligent interfaces. Although the concept can theoretically be applied to any general data sets, at this stage, we only apply the prototype on hierarchical or tree structures for the demonstration purpose. The layout uses a modified fractal approach to generate the tree visualization at a fast speed. The user can interact with the visualization by relocating and/or rotating any particular nodes. The property of new visualization corresponding to the modifications can interactively be applied to sub-structures. This would produce new layout according to the user's preference when he/she interacts with the visualization. We next describe the algorithms in detail.

4.1 Layout

The layout of the relational structure is constructed as a tree where child nodes are branched from their parent node. Initially, the root node is placed at the centre of the display and is drawn as a vertical segment. The child nodes are next placed a long the root node at an angular faction. Technically, each node *N* is specified by a rectangular boundary *B* which is defined by the position of the bottom-centre *{p}*, width *{w}*, height *{h}* and the rotation angle *{α}* in comparison to vertical axis. The width of boundaries is proportional to the weight of a node (e.g. the number of nodes of a sub-structure) in comparison to

Suppose that we need to calculate the layout of a structure whose rooted *N* has been defined by its boundary *B{p, w, h, α}*, the partitioning algorithm is as followings:

procedure layout (Node N)
begin
  for all child-nodes of N
  begin
    C_i = current child of N
    B_i = boundary of C_i
    place C_i at either at left or right respectively of N where
    C_i.p is defined by N.p + increment_segment
    C_i.w = increment_segment
    C_i.h = N.h / reduce_factor
    $C_i.\alpha$ = N.α + rotation_angle
    for all child-nodes S_j of C_i
    begin
      layout(S_i)   // repeat to all other sub-structure
    end
  end
end

Figure 2 and figure 3 show two examples of the visualization of a small data set and a large data set respectively. Figure 2 shows clearly the hierarchical structure while there is quite significant overlapping in the visualization in figure 3.

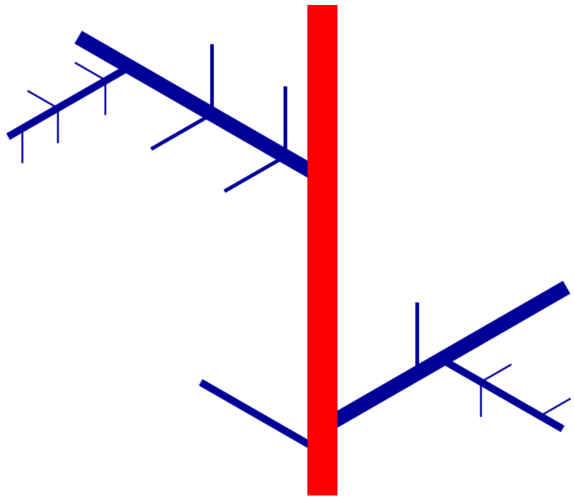

Figure 2. An example of the visualization of a small data set

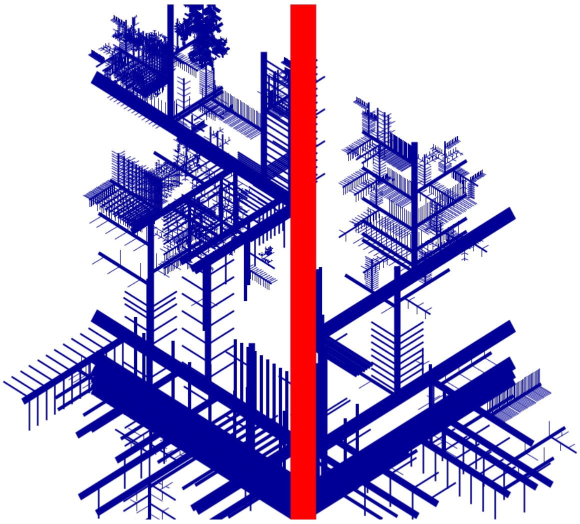

Figure 3. An example of the visualization of a large data set.

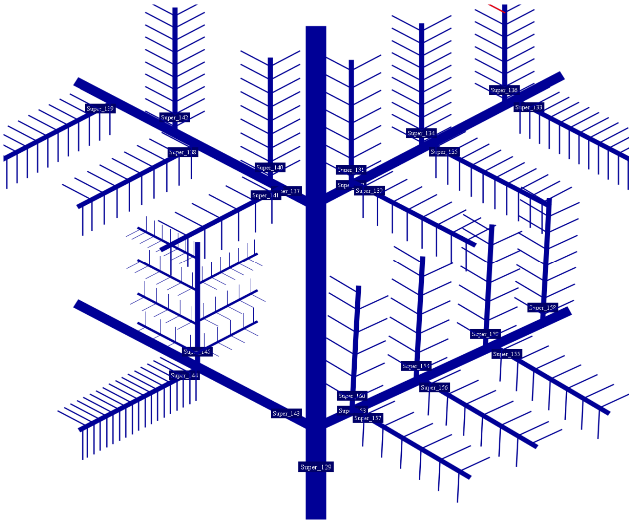

Figure 4. An example of a visualization before the interaction.

4.2 Interaction

The prototype includes two types of interactions in the current implementation, including movement and rotation. Each interaction is associated with a transformation process so that the capture property of the interactions can be further applied to any desirable nodes and sub-structures (or crossing structures). Our implementation, however, simplifies this process by applying to siblings and/or to descendants of the interaction nodes. Applying to siblings means the same property is used for all sibling nodes as well as their entire sub-structures while applying to descendants means the sample property is used for corresponding child node and its descendants.

- *Movement interaction:* this interaction corresponds to the movement of a node or a sub-structure from one location to another location by clicking and dragging the mouse. This interaction allows the user to arrange the position of nodes and sub-structures in their preference.
- *Rotation interaction*: this interaction corresponds to the rotating a node or a sub-structure around its position by clicking and rolling the middle mouse-wheel. This interaction allows the user to arrange the angle of nodes and/or sub-structures in their preference.

The interaction provides an effective mechanism so that the user can customise the visualization according to their preference. By using different type of interactions, they can reorganise the layout at any particular sections, such as move the interesting structure to a new location, change the layout preference of nodes. Further more, our system also allows the same property of a particular action to be applicable to any section of the geometry, for example, to every siblings or corresponding descendants in the structure. This brings a much greater ability to customise very large structure because the interaction does not require to be performed on the entire visualization, but could be on a small portion the structure. The same geometrical property is then generalised to other sub-structures in the visualization.

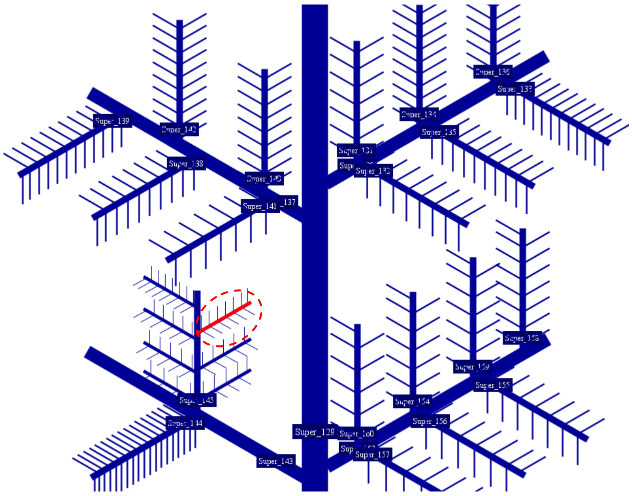

Figure 5. An example of the movement of sub-structures where all four child nodes move up/down along the root node.

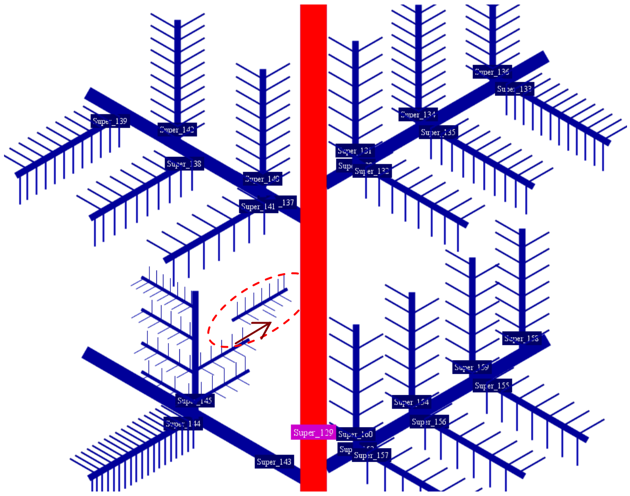

Figure 6. An example of the transformation when the movement of a node is applied to its corresponding descendant.

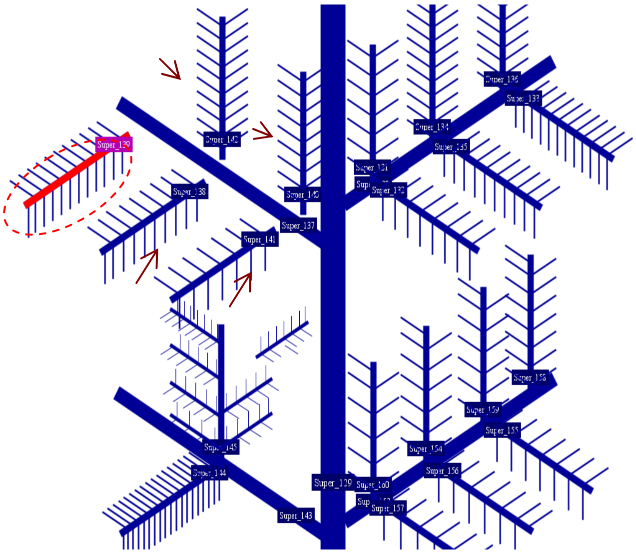

Figure 7. An example of the transformation when a) the movement is applied to all siblings

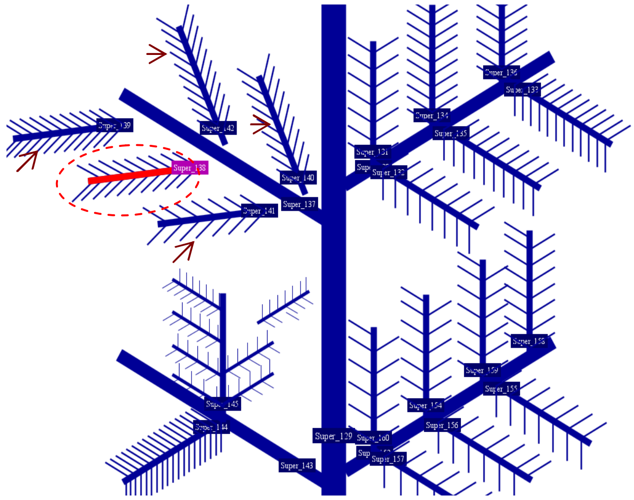

Figure 8. An example of the transformation when the rotation is applied to all siblings and their sub-structures.

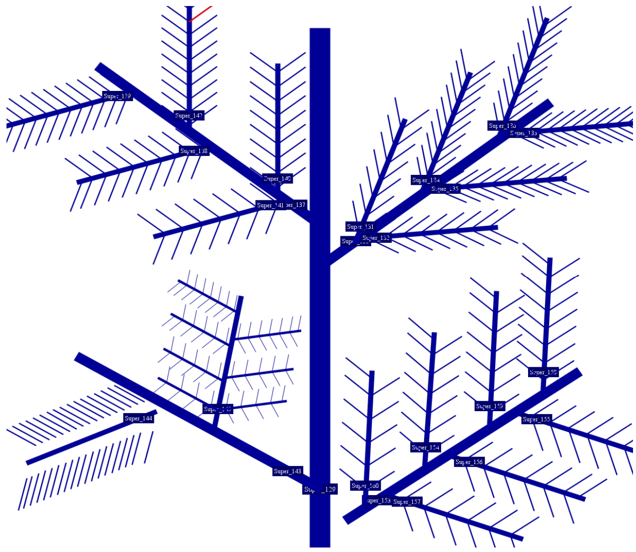

Figure 9. An example of a final display after interaction and transformations.

Figure 4 shows an example of a structure before the interaction. This figure shows a tree visualization with four sub-structures where each has a number of descendants. Figure 5 shows example of a movement of four sub-structures along the root node. This figure also indicates the user is focusing on a sub-structure which is circled. Figure 6 shows an example of the transformation when the movement is applied to the corresponding descendant. This figure indicates that the selected node (i.e. circled by a red boundary) is moved away from its original position and the corresponding child node also followed the manner. Figure 7 is an example of the transformation when the movement is applied to all siblings. This figure shows the selected node (i.e. circled by a red boundary) is relocated a distance from its original position and as well as all its siblings and their sub-structure. Figure 8 is an example of the transformation when the rotation is applied to all siblings and their sub-structures. This figure shows that the angle of the sub-structures is reduced according to the change of the selected node. Finally, figure 9 shows the display after a number of interaction and transformations. This figure provides a much clearer view of the display compared to the figure 4, and it could reveal some interesting property of the structure according to the user's visual preference.

5. CONCLUSION AND FUTURE WORK

We have presented a new conceptual model for intelligent interactive visualizations. The model could provide a greater flexibility for user to customise the visualizations according to the nature of data and user's preference. We believe that this approach will open a new opportunity for the development of new intelligent visualization in supporting large scale data analysis.

We will next investigate new technology that uses a simple visualization language as the underlying computing machinery to encode the captured logics and rules from the analyst's behaviours. From this underlying layer, the analyst will be able to construct the new visualization based on the available model, refine the model and visualization if necessary, and importantly query the model to automatically visualize and analyse the other data sets with much larger size

6. REFERENCES

[1] Newman, M. Watts, D., Barabási A.-L. 2006. *The Structure and Dynamics of Networks*, Princeton University Press.

[2] Thomas, J.J., Cook, K.A. 2005. *Illuminating the Path*. IEEE Computer Society Press, Los Alamitos.

[3] Ankerst, M. 2000. Visual Data Mining. *PhD Thesis*, der Ludwig-Maximilians-Universität München, Berlin.

[4] Ma, K. L. 2007. Machine Learning to Boost the Next Generation of Visualization Technology. *IEEE Computer Graphics and Applications*, 2007, pp. 6-9.

[5] Garg S., Nam J. E., Ramakrishnan I.V., Mueller K. 2008. Model-Driven Visual Analytics. In *Proc. IEEE Symposium on Visual Analytics Science and Technology*, Columbus, Ohio, pp. 19-26.

[6] Fuchs R., Waser J., Groller M.E. 2009. Visual Human + Machine Learning. *IEEE Trans. On Visualization and Computer Graphics*, 15(6), 2009, pp. 1327-1334.

[7] Tatu A., Albuquerque G., Eisemann M., Schneidewind J., Theisel H., Magnor M. Keim D. 2009. Combining Automated Analysis and Visualization Techniques for Effective Exploration of High-Dimensional Data. In *Proc. IEEE Symposium on Visual Analytics Science and Technology*, Atlantic City, New Jersey, pp. 59-66.

[8] Green T. M., Ribarsky W., Fisher B. 2008. Visual Analytics for Complex Concepts Using a Human Cognition Model. In *Proc. IEEE Symposium on Visual Analytics Science and Technology*, Columbus, Ohio, pp. 91-98.

[9] Bier E. A., Card S. K., Bodnar J. W. 2008. Entity-Based Collaboration Tools for Intelligence Analysis. In *Proc. IEEE Symposium on Visual Analytics Science and Technology*, Columbus, Ohio, pp. 99-106.

[10] Simoff, S., Bohlen, M. and Mazeika, A. (eds) 2008. *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics*, Springer, Heidelberg.

[11] Dwyer T., Lee B. Fisher D., Quinn K. I., Isenberg P., Robertson G., North C. 2009. A Comparison of User-Generated and Automated Graph Layouts. *IEEE Trans. On Visualization and Computer Graphics*, 15(6), pp. 961-968.