A Visual Analytics Technique for Identifying Heat Spots in Transportation Networks

Marian Sorin Nistor
Department of Computer Science, Universität der Bundeswehr München,
85577 Neubiberg, Germany

Stefan Wolfgang Pickl
Department of Computer Science, Universität der Bundeswehr München,
85577 Neubiberg, Germany

Martin Zsifkovits
Department of Computer Science, Universität der Bundeswehr München,
85577 Neubiberg, Germany

ABSTRACT

The decision takers of the public transportation system, as part of urban critical infrastructures, need to increase the system resilience. For doing so, we identified analysis tools for biological networks as an adequate basis for visual analytics in that domain. In the paper at hand we therefore translate such methods for transportation systems and show the benefits by applying them on the Munich subway network. Here, visual analytics is used to identify vulnerable stations from different perspectives. The applied technique is presented step by step. Furthermore, the key challenges in applying this technique on transportation systems are identified. Finally, we propose the implementation of the presented features in a management cockpit to integrate the visual analytics mantra for an adequate decision support on transportation systems.

Keywords: Public Transportation System, Transportation Network Analysis, Munich Subway Network, Visual Analytics, Management Cockpit.

1. INTRODUCTION

The analysis of complex networks is an ongoing research field for various disciplines [1–3]. The concept of a complex network is used as a simplified frame of a complex system (e.g. a public transportation system). The nodes and the links of a network are represented by the entities and their interrelations in a system [4]. However, finding suitable visualization techniques for the structural information of complex networks is an open question [5].

Several studies showed so far the applicability of visual analytics in general [6–8] and to transportation systems in particular [9]. The paper at hand aims for developing a visual analytics technique to detect multiple vulnerable areas in a transportation network. This approach was already applied in bioinformatics to biological network analysis [10,11], but should not stay limited to this field of application.

Transportation networks are often the target of disturbances and attacks, which makes the knowledge about vulnerability and resilience even more important. We therefore aim for applying the biological network analysis to the field of transportation networks in order to analyze and visualize critical spots.

In section 2, relevant features for transportation networks are identified. Methodological background, such as the visual analytics process and applied network analysis measures for transportation networks, is presented in section 3. In section 4, the Munich subway network is used as an example to apply the proposed visual analytics technique. Thereby, several visual representations according to different measures are presented. Section 5 presents the key challenges of this technique and a final conclusion.

2. FROM BIOLOGICAL NETWORKS TO TRANSPORTATION NETWORKS

To translate the analysis of biological networks to transportation networks, we first have to define their main characteristics and differences. Compared to biological networks, transportation networks tend to be smaller. While biological networks can have several million nodes, the world-wide air transportation network has only 1.000 nodes and 35.000 links [12].

The example network in this paper, the Munich subway network has 100 nodes and 198 links [13]. This leads to the fact that the visualization of a transportation network is rather focused on information visualization than on network drawing. The network drawing is more oriented on the visual representation, while the information visualization is more oriented on operating the network hierarchies for various view perspectives and interactions between its nodes.

Compared to biological networks, the visualization of transportation networks tends to be more subjective. This holds especially for the purpose of gaining information and knowledge for decision making. Thereby, nodes and links in transportation networks can be associated with costs and/or causalities. Therefore, the loss or damage of just one can be very significant in such networks [14].

While most biological networks are time independent, analyzing a transportation network, e.g. in terms of passengers' flow, train traffic, financial revenues/losses, particular vulnerabilities, etc., is preferred to be studied on a time base. For example, the time window when a node reaches its maximum value in terms of passengers' flow. Also, when studying for example the network reliability, route alternatives, or the shortest path of a route, more information needs to be considered for transportation networks. Besides the length of a link, e.g. also the availability of trains on that link need to be considered [14]. When visualizing a transportation network as a directed network, the weights can store various information, e.g. the weights representing the number of trains traveling from one station to another in a subway network. These values are related to the degree of the nodes. Therefore, when considering parts of network with only one subway line, the number of trains varies in a small range, while for the dense areas where the stations are crossed by more than one subway line, the number of trains increase significantly with the number of subway lines.

The connectivity of a transportation network is rather loose compared to most biological networks. For most of the world-wide subway networks it would be enough to identify and cut one node, or one link, in order to disconnect the whole network [15].

3. BACKGROUND

Visualization, as a science, is mainly dealing with visualization techniques for an efficient interaction. Branch of this science is the information visualization. This one refers to the visualization of abstract data with no explicit spatial references available [16]. In the last decade, an interdisciplinary version of visualization arose: visual analytics. The reason is the strong need of understanding, and also visualizing, huge amounts of data. Visual analytics is an adapted version of information visualization which combines advanced data analysis algorithms. Therefore, it can be defined as *"an integral approach to decision-making, combining visualization, human factors and data analysis"* [17].

The visual analytics process is described as an adaptive process [9], where the user can be rather involved in the visual data exploratory loop, or to the automated data analysis loop. This process is applied to assess transportation networks in Section 4.

The network analysis combined with visual analytics is a key element for a proper understanding of a network. Classical network topology parameters can offer important structural information of the analyzed networks. These are recognized as relevant for network vulnerability measures [18,19]. Topology parameters, such as the number of nodes and links, diameter, network connectivity, girth, nodes and links connectivity, and cohesion, are compiled components for heuristic reliability indexes [20]. These indexes offer a quicker and insightful overview of the entire network vulnerability. This type of analysis was already conducted on transportation networks [15].

Structural measures, such as network entropies, can be considered as reliable measures to determine the structural properties of a network [21]. These measures capture the information structure of the complete neighborhood and the centrality properties of each node in the network [22]. Entropy measures have been successfully applied on transportation networks by using this information-theoretic method [23].

In this paper we refer to the recently introduced flow-weighted efficiency measure [14]. This measure calculates the efficiency of a transportation network by assessing two metrics weights: the length of links and the train traffic on each link. The most efficient nodes are here considered as being most vulnerable, as losing their regular flow results in a serious disturbance regarding the serviceability of the network [14]. We therefore propose applying the visual analytics technique to the flow-weighted efficiency measure. This enables the detection of network vulnerabilities from different visual perspectives: modularity, distances, train flow, and efficiency.

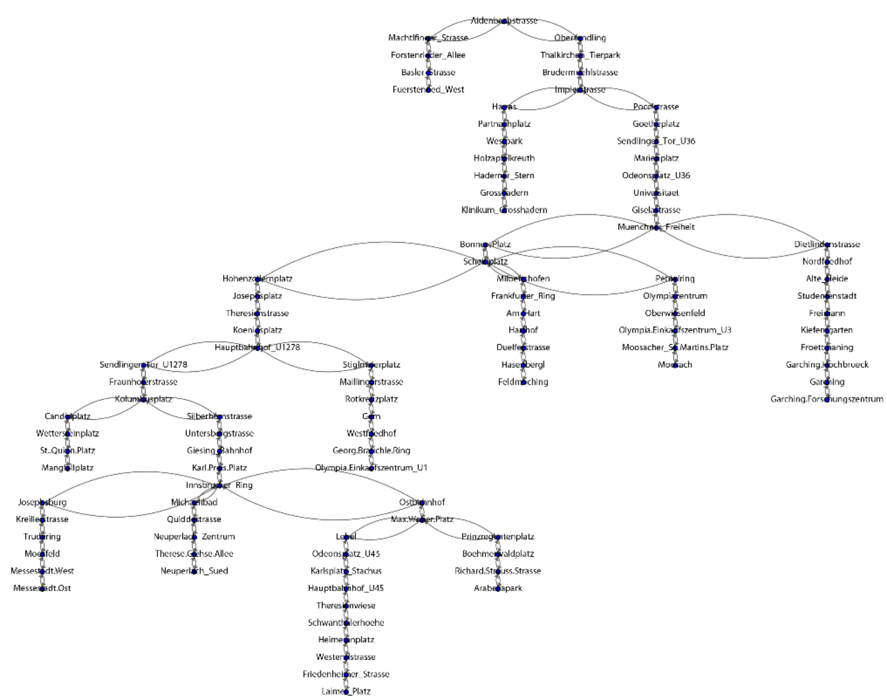

Figure 1. A simple tree-like visual representation of the Munich subway network.

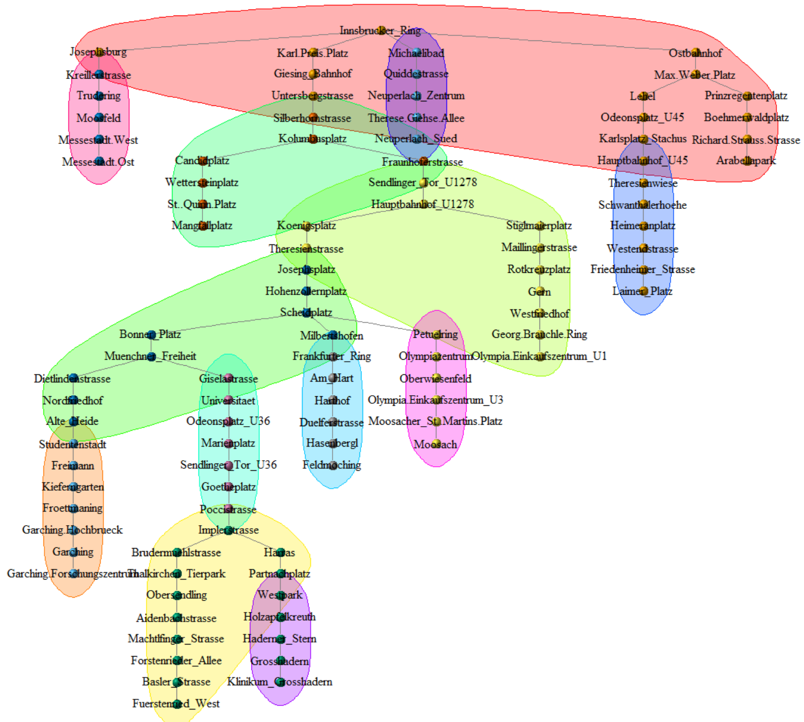

Figure 2. A modular tree-like visual representation of the Munich subway network.

4. APPLICATION: MUNICH SUBWAY NETWORK

The visual analytics process is applied to the Munich subway network. This is encoded as an adjacency matrix for a directed network. The subway network consists of 100 stations as nodes, and 198 connections between stations as links [13].

In this work, we consider the train traffic between every two linked stations in both directions on a daily basis. The collected numbers are public and available at http://www.mvv-muenchen.de/. The selected schedule is based on the weekday schedule for business days between Monday and Thursday. However, results might differ for the other schedules available on the network, Friday, Saturday, Sunday and Holidays [24].

For visualization, we follow the visual analytics mantra *"Analyze first, Show the Important, Zoom, filter and analyze further, Details on demand"* [25].

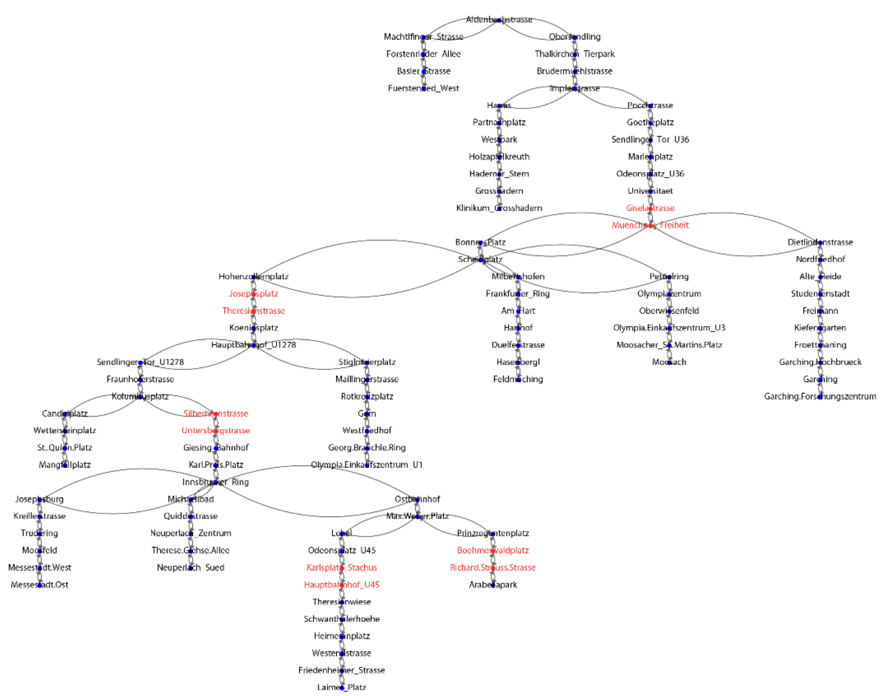

Figure 3. Top five shortest (real) distances of the Munich subway network highlighted in a tree-like visual representation.

In this paper, the focus is on a simple representation of the network. The nodes are the key of our visual analysis. The selected visualization layout is Reingold Tilford's [26] which is a tree-like layout. For this layout, some cycles of the network might be omitted. *Figure 1* is a simple representation of the network as directed graph with the selected visualization layout. The plots in this paper are generated with the *RStudio* [27] software and the *igraph* [28] package.

More insights on the placement of nodes according to the other nodes can be seen from a modular perspective, which is illustrated in *Figure 2*. A module contains a connected subgroup of nodes of the network selected on different criteria. In this case, the nodes are grouped based on the available connections between them. Two different types of groups can be spotted. The small ones highlight the linear paths of subway lines, while the big ones highlight the presence of hub nodes. The latter are the key nodes of the network. The big groups show their impact in the network in terms of connectivity. Hub nodes like *Innsbrucker Ring, Kolumbusplatz, Hauptbahnhof* (for the lines U1, U2, U7, U8), *Münchener Freiheit* and *Implerstraße* can be spotted in the figure.

Another interesting analysis is the perspective of distances in the network. *Figure 3* shows the top five shortest distances highlighted based on real life data [29]. The connections belong to the following groups of nodes: *Josephsplatz - Theresienstraße* 0.513 km, *Hauptbahnhof* (for the lines U4, U5) - *Karlsplatz Stachus* 0.521 km, *Böhmerwaldplatz- Richard Strauß Straße* 0.552 km, *Silberhornstraße – Untersbergstraße* 0.553km, and Giselastraße - *Münchener Freiheit* 0.579 km.

In *Figure 4* the top ten most demanded nodes in terms of train flow in the Munich subway network are highlighted for a weekday schedule. For this analysis the total number of trains stopped in a node per day are considered. The nodes with the highest daily train flow are:

- Innsbrucker Ring with 759 trains/day,
- Giselastraße with 719 trains/day,
- Implerstraße with 719 trains/day,
- Marienplatz with 719 trains/day,
- Odeonsplatz (lines U3, U6) with 719 trains/day,
- Sendlinger Tor (lines U3, U6) with 719 trains/day,
- Universität with 719 trains/day,
- Goetheplatz with 717 trains/day,
- Poccistraße with 717 trains/day,
- Sendlinger Tor (lines U1, U2, U7, U8) with 712 trains/day.

In *Figure 4* a very demanded route can be observed in the network with eight consecutive nodes in the top ten selection. The other two nodes highlighted represent central node hubs of the network.

In the last visual perspective of *Figure 5*, the top results of the flow-weighted efficiency measure [14] are highlighted. This measure is a combined analysis from the shortest distance of any pair of nodes in a network and the minimum number of trains available on that route per day. Thus, the measure uses the exact data processed for *Figure 3* and *Figure 4*.

The top ten highlighted nodes and their values from *Figure 5* are:

- Giselastraße - *6.32*,
- Münchner Freiheit - *6.32*,
- Hauptbahnhof (lines U4, U5) - *5.86*,
- Karlsplatz Stachus - *5.86*,
- Implerstraße - *5.83*,
- Poccistraße - *5.83*,
- Marienplatz - *5.71*,
- Odeonsplatz (lines U3, U6) - *5.71*,
- Goetheplatz - *5.37*,
- Universität - *4.91*.

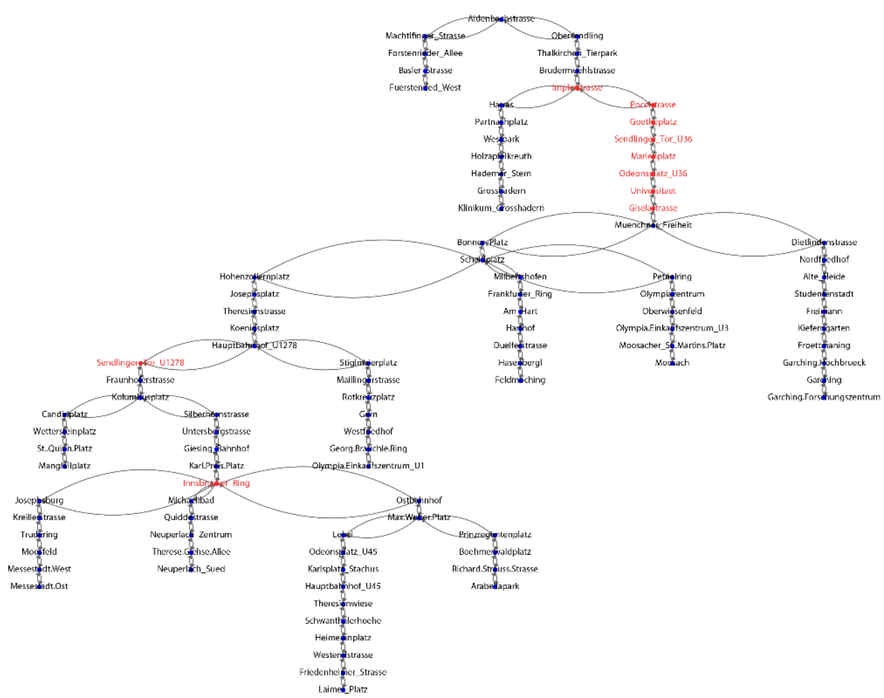

Figure 4. Top ten most demanded nodes in terms of train flow in the Munich subway network from Monday to Thursday highlighted in a tree-like visualization.

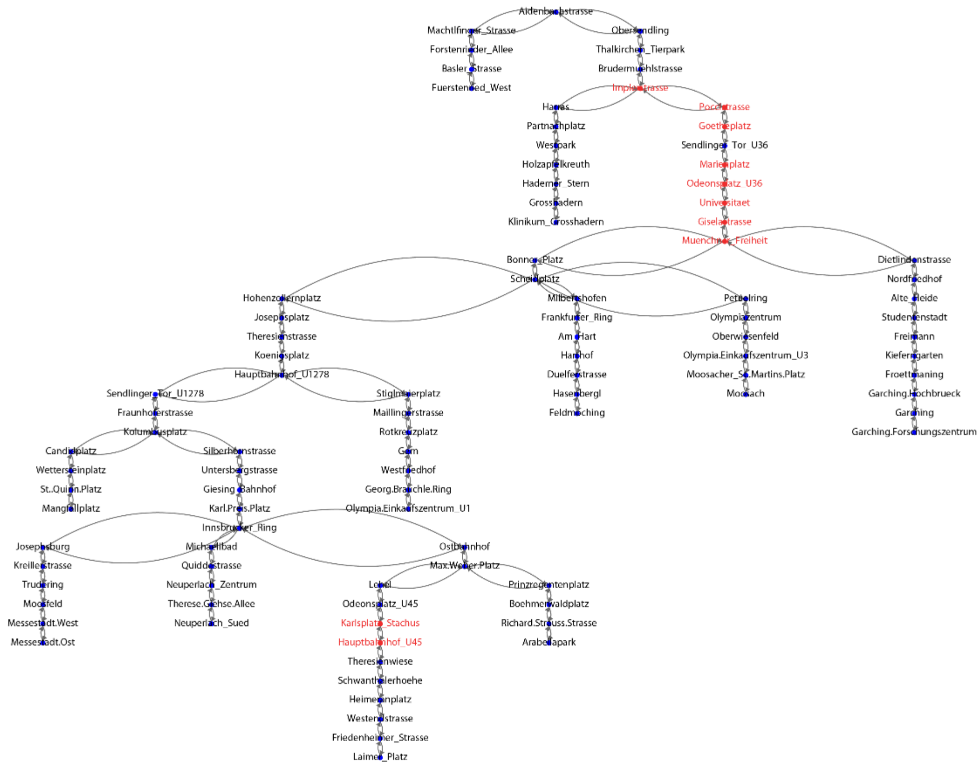

Figure 5. Top ten most efficient nodes in terms of train flow and shortest (real) distances in the Munich subway network from Monday to Thursday highlighted in a tree-like visualization.

For profound managerial decision making each analysis might make sense and give additional insights into the network structure. This shows that applying measures from biology definitively makes sense for other fields of application, such as transportation networks. However, the special architecture and design of a transportation network demands for further improvements of the measures. To this end, the flow-weighted efficiency measure was found to be very helpful and allowing for deeper insights [13].

A more convenient solution than working with heat maps is the extraction of highlighted nodes. This would lead to a further step via creating another network: the network of networks. In this way, the last step of the visual analytics mantra "*Details at demand*" can be applied. More precisely, the network analysis measures presented in the previous section could now be applied on the exact parts of the network on which decisions must be focused. However, the procedure stays unchanged and follows the approach presented here.

5. CONCLUSION

The paper at hand might be seen as one more proof that the application of visual analytics is favorable for several disciplines and might support managerial decision making in various fields. As critical infrastructures in general, and the rail-bound public transport in special, are essential for the functioning of a society and are therefore often target of disturbances and attacks. We analyzed the systems vulnerability and identified the most important spots that need special treatment in terms of safety and security, as well as recovery after interruptions. Thereby, we showed that the combination of two different measures from biology can be used to gain deeper insights into the serviceability of the system.

However, there are several challenges in applying the visual analytics process on transportation networks. Concerning the analysis, it is an open gap to find the most suitable tools for the structural interpretation of the networks. The same holds for visualization techniques for this type of networks. The solutions are rather subjective.

However, assessing multiple values of weights for links and the physical position of each node in relation to the others will improve the analysis of transportation networks. In this sense, the analysis will be more realistic when measuring classical topological measures, e.g. diameter, shortest path, or average path length. A physical position of the nodes can control the overlapping problem when plotting.

This type of analyzes can also be performed on a time base, being an extension from static networks to dynamic networks. Therefore, the most vulnerable spots of the transportation networks can be assessed for different time schedules.

In conclusion, visual analytics can be successfully applied to describe and visualize network structures and their vulnerabilities. However, decision makers do not have modelers available all the time. Therefore, we propose the automated analysis via the implementation of several measures for special types of networks, such as transportation networks, in a management cockpit. This integration of visual analytics into a novel decision support tool would allow for fast and detailed analyzes in such special fields. The integration into a management cockpit will be presented in a follow-up publication.

ACKNOWLEDGMENT

Feedback on the paper from Prof. Matthias Dehmer, Dr. Doina Bein, Prof. Wolfgang Bein and the two unknown reviewer is gratefully acknowledged.

This research was supported by the German Federal Ministry of Education and Research (BMBF), project RE(H)STRAIN (Grant No. 13N13786).

Research of author Marian Sorin Nistor, was funded by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA Grant Agreement Number 317382.

REFERENCES

[1] M. Dehmer, F. Emmert-Streib, **Analysis of Complex Networks**, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2009.

[2] S. Derrible, **Metros as Biological Systems: Complexity in Small Real-life Networks**, M. Dehmer, A. Mowshowitz, F. Emmert-Streib (Eds.), Advances in Network Complexity, Wiley-VCH, Weinheim, Germany, 2013, pp. 259–285.

[3] C. Correa, T. Crnovrsanin, Kwan-Liu Ma, **Visual Reasoning about Social Networks Using Centrality Sensitivity**, IEEE Trans. Visual. Comput. Graphics 18 (1) (2012) 106–120.

[4] A. Geiger, A. Kroll, **Modeling and analyzing technical systems as complex networks: Detecting inverse response**, Computational Intelligence in Control and Automation (CICA), IEEE Symposium, 2013, pp. 89–96.

[5] D.A. Keim, J. Kohlhammer, G. Ellis, F. Mansmann, **Mastering the information age-solving problems with visual analytics**, Eurographics Association, 2010.

[6] S. Jarukasemratana, T. Murata, **Recent Large Graph Visualization Tools: A Review**, Information and Media Technologies 8 (4) (2013) 944–960.

[7] J.R. Harger, P.J. Crossno, **Comparison of open-source visual analytics toolkits**, IS&T/SPIE Electronic Imaging, 2012, p. 82940.

[8] L. Zhang, A. Stoffel, M. Behrisch, S. Mittelstadt, T. Schreck, R. Pompl, S. Weber, H. Last, D. Keim, **Visual analytics for the big data era - A comparative review of state-of-the-art commercial systems**, pp. 173–182.

[9] M.S. Nistor, S.W. Pickl, M. Zsifkovits, **Visual Analytics of Complex Networks: A Review from the Computational Perspective**, The 2015 International Conference on Modeling, Simulation and Visualization Methods, Las Vegas, NV, USA, CSREA Press, San Diego, CA, USA, 2015, pp. 10–15.

[10] C. Zhou, L. Zemanová, G. Zamora, C.C. Hilgetag, J. Kurths, **Hierarchical organization unveiled by functional connectivity in complex brain networks**, Physical review letters 97 (23) (2006) 238103.

[11] A.-L. Barabasi, Z.N. Oltvai, **Network biology: understanding the cell's functional organization**, Nature reviews genetics 5 (2) (2004) 101–113.

[12] Northwestern University, **Researchers develop method that shows diverse complex networks have similar skeletons**, available at http://phys.org/news/2012-06-method-diverse-complex-networks-similar.html (accessed on July 29, 2015).

[13] Münchner Verkehrsgesellschaft mbH, **MVG in figures**, available at https://www.mvg.de/dam/en/mvg/ueber/unternehmensprofil/mvg-in-figures-s (accessed on August 13, 2015).

[14] M.S. Nistor, S.W. Pickl, M. Raap, M. Zsifkovits, **Quantitative Network Analysis of Metro Transportation Systems: Introducing the Flow-Weighted Efficiency Measure**, Proceedings of the International Conference on Economics and Management of Networks (EMNet 2015), Cape Town, South Africa

[15] M.S. Nistor, M. Dehmer, S.W. Pickl, **Network Exploratory Analysis on Subway Transportation Systems against Complex Threats Including a Human Factors Perspective**, Procedia Manufacturing 3 (2015) 6593–6598.

[16] R. Spence, **Information visualization**, Springer, 2001.

[17] D.A. Keim, F. Mansmann, D. Oelke, H. Ziegler, **Visual Analytics: Combining Automated Discovery with Interactive Visualizations**, J.-F. Jean-Fran, M.R. Berthold, T. Horváth (Eds.), Discovery Science: 11th International Conference, DS 2008, Budapest, Hungary, October 13-16, 2008, pp. 2–14.

[18] R. Wilkov, **Analysis and design of reliable computer networks**, Communications, IEEE Transactions on 20 (3) (1972) 660–678.

[19] H. Frank, I.T. Frisch, **Communication, transmission, and transportation networks** (1971).

[20] I.M. Soi, K.K. Aggarwal, **Reliability indices for topological design of computer communication networks**, Reliability, 30 (5) (1981) 438–443.

[21] M. Dehmer, **Information processing in complex networks: Graph entropy and information functionals**, Applied Mathematics and Computation 201 (1) (2008) 82–94.

[22] M. Dehmer, K. Varmuza, S. Borgert, F. Emmert-Streib, **On entropy-based molecular descriptors: Statistical analysis of real and synthetic chemical structures**, Journal of chemical information and modeling 49 (7) (2009) 1655–1663.

[23] M. Dehmer, M.S. Nistor, W. Schmitz, K.A. Neubecker, **Aspects of Quantitative Analysis of Transportation Networks**, B. Jürgen, Andreas Meissner, G. Jürgen (Eds.), Proceedings of the 10th Future Security, 2015, pp. 239–244.

[24] MVV GmbH, **Alle Informationen zu den Bahnhöfen im MVV**, 2015, available at http://www.mvv-muenchen.de/de/netz-bahnhoefe/bahnhofsinformation/index.html (accessed on August 13, 2015).

[25] D. Keim, F. Mansmann, J. Schneidewind, H. Ziegler, others, **Challenges in visual data analysis**, Information Visualization, 2006. IV 2006. Tenth International Conference on, 2006, pp. 9–16.

[26] E.M. Reingold, J.S. Tilford, **Tidier drawings of trees**, Software Engineering, IEEE Transactions on (2) (1981) 223–228.

[27] RStudio Team, **RStudio: Integrated Development Environment for R**, Boston, MA, 2015, available at http://www.rstudio.com/.

[28] Gabor Csardi, Tamas Nepusz, **The igraph software package for complex network research**, InterJournal Complex Systems (2006) 1–9.

[29] F. Schütz, B. Merath, **muenchnerbahn.de - Stationsabstände**, available at http://www.muenchnerubahn.de/netz/bahnhoefe/stationsabstaende/ (accessed on August 13, 2015).