# StudentViz: A Tool for Visualizing Students' Collaborations in a Social Learning Environment

Alex Becheru, Andreea Calota, and Elvira Popescu

Computers and Information Technology Department, University of Craiova, Romania
becheru@gmail.com, popescu_elvira@software.ucv.ro

**Abstract.** Visualizations play an important role in learning analytics, supporting reflection and decision making. Network representations are commonly used for depicting social interactions between learners. While there are many network visualization platforms available, most of them are aimed at researchers, requiring social network analysis expertise. Our goal is to provide a simple tool for visualizing students' collaboration patterns in a social learning environment, which should be easy to use by the teacher. The paper presents a description of this tool (called StudentViz), some design and implementation details, and an illustration of its functionalities. It further shows that the tool adequately addresses the visualization needs of the instructors, fostering insight gaining.

**Keywords:** information visualization · visual analytics · learning analytics · graph plotting · social networks analysis · social learning environments

## 1 Introduction

Information visualization relies on the remarkable visual perception abilities of humans for pattern discovery [18]. It employs interactive visual representations in order to amplify cognition [12] and generate "insight" [5].

Visual approaches have been used in learning analytics, to help teachers and students explore learner traces from virtual learning environments. Various types of data can be included in a learning analytics dashboard, such as: *artifacts produced by learners, social interaction, resource use, time spent, test and self-assessment results*. The goal is to provide insight into learning data, supporting awareness and decision making, and increasing students' engagement and motivation [12].

In particular, social network analysis (SNA) and network visualizations have been used to investigate students' interactions taking place in educational environments [6, 16]. In this context, our goal is to visualize collaboration patterns between students in a social media-based learning space, which was less explored in the literature. More specifically, we are interested in studying learners' collaboration in our eMUSE social learning environment [15]. The platform integrates several social media tools (blog, wiki, microblogging tool) and provides



M. Chang et al. (eds.), *Challenges and Solutions in Smart Learning*, Lecture Notes in Educational Technology,
https://doi.org/10.1007/978-981-10-8743-1_12

value added services both for the students and the instructor (learner tracking, basic administrative services, data visualizations, peer assessment and grading support); more details about eMUSE can be found in [15].

We have already proposed a conceptual framework for knowledge extraction and visualization based on SNA in [2], which we have validated in [3]. *Gephi* network analysis tool [11] was used for all computations and graph visualizations, which adequately fitted researcher's needs; however, the tool was deemed too complex for instructors, who are not specialists in SNA or visualization. Therefore, we decided to build a simple network visualization tool, easy to use by the teachers and specifically designed to work in conjunction with our eMUSE social learning environment. The tool should provide useful and relevant visualizations from the instructor's point of view, therefore we first identified a list of visualization needs (VN) outlined by the teachers working with eMUSE:

- VN1. Visualize the general status of collaboration
- VN2. Visualize the status of collaboration for each community
- VN3. Visualize the status of collaboration for each learner.

Furthermore, the tool should support the processes of gaining insight through information visualizations, as identified in [18]:

- P1. *Provide Overview* - grasp the big picture of a dataset
- P2. *Adjust* - explore a dataset by changing the abstraction level or selection range (e.g., by filtering, grouping, aggregating)
- P3. *Detect Pattern* - find relationships, trends, or anomalies in the dataset
- P4. *Match Mental Model* - correlate the data with the user's mental model of it, in order to facilitate understanding.

Starting from these requirements, we designed and implemented our StudentViz tool, as described in the following sections. An overview of existing network visualization platforms is included in section 2. StudentViz functionalities and implementation details are presented in section 3. Its visualization capabilities are illustrated and validated in section 4. Finally, section 5 outlines some conclusions and future research directions.

## 2 Related Work

Networks or graphs are a common visualization method in educational settings [5]. They can be used to display information regarding students' interactions, which is particularly important in case of collaborative learning and social learning environments.

Many network visualization platforms (NVP) are available [7]; however, they are not specifically built for educational settings, so we wanted to investigate whether they can be used for our particular learning scenario, in conjunction with eMUSE platform. As our target users are instructors with limited technical expertise, we imposed some initial restrictions. The considered tools should be free, easy to install and use, and operating system independent. Moreover, they should provide high flexibility, so that instructors could adapt the visualization methods to their needs. Hence, our evaluation included the following platforms:

- *Cuttlefish*[1] is a very easy to use platform, but with limited capabilities and no flexibility; we also experienced some visualization glitches upon using the zooming functionality.
- *Cytoscape* [17] is an open source platform developed for molecular networks visualization, that has expanded its use across various network related research fields. Its standard features are relatively easy to use. However, the platform lacks flexibility of the visualization methods.
- *Visione*[2] has similar capabilities with *Cytoscape*, but it provides even less flexibility and the user interface is cluttered and non-intuitive.
- *Tulip* [1] is a visualization platform for relational data. It provides highly flexible visualization and a wide range of analysis capabilities for various research fields.
- *Gephi* [11] aims to be a general platform of analysis and visualization for all kinds of networks. Its clear design and resemblance with *Photoshop* make it very easy to use. Furthermore, its visualization capabilities are flexible and extensible through plugins.

Overall, we found *Gephi* and *Tulip* to be equally capable in terms of visualization functionalities; however, *Gephi* provided a better user experience, hence we chose to use it in [3].

Nevertheless, all platforms were considered too complex by the instructors, including many irrelevant functionalities for their purpose and requiring SNA expertise. Also, some of the desired visualizations (e.g., team and community perspectives) required significant effort in order to be generated with the existing NVP, starting from our available eMUSE dataset. Therefore, we decided to build a network visualization tool dedicated for the teachers, with a simple and intuitive interface, as described next.

## 3 StudentViz Tool Description

StudentViz is a data visualization tool purposely built to work in conjunction with our eMUSE social learning environment. Its aim is to provide suggestive visualizations of students' collaboration patterns, as they are recorded by the platform. More specifically, eMUSE integrates several social media tools that students use for communication and collaboration support. All students' social media traces are monitored and stored by eMUSE, and StudentViz uses these data

[1] http://cuttlefish.sourceforge.net
[2] http://www.visone.info/html/about.html

to draw the graphs depicting social media interactions between the students. A schematic representation of the data flow is shown in Fig. 1.

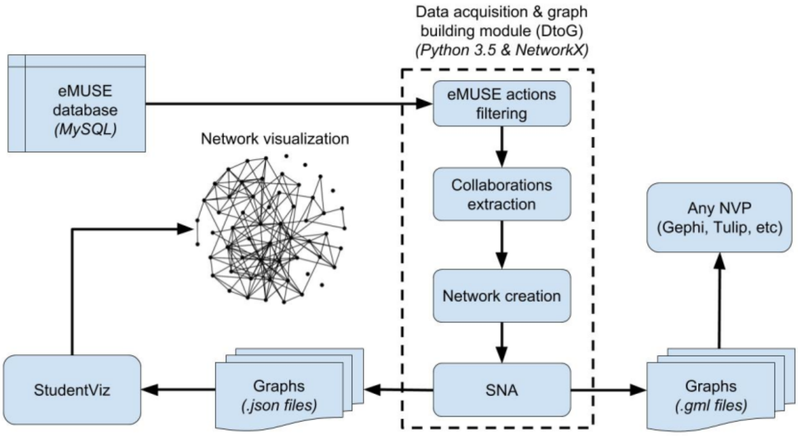


**Fig. 1.** StudentViz - network visualization data flow

The first step was to map students' collaborations as social networks. A *data acquisition & graph building* module (denoted *DtoG*) was designed to bridge the gap between the data source (eMUSE) and the visualization tool (StudentViz). DtoG processes the raw data collected by eMUSE, filtering the collaboration actions, and then creates various social graphs on which several SNA methods are applied. More specifically, directed graphs are built starting from students' interactions on the blog and microblogging tool; nodes represent learners and links represent messages exchanged through the social media tools integrated in eMUSE. The types of interactions (collaborations) taken into account on Blogger and Twitter respectively are detailed in [3]. These graphs can be exported in various formats (e.g., *.gml* or *.json* files), which can be subsequently input into any NVP, including StudentViz. As far as implementation is concerned, DtoG was built using Python 3.5 programming language and *NetworkX* graph analysis library [10].

In an attempt to reduce instructors' effort in using StudentViz, we decided to conceive it as a web application, thus eliminating the need of installation, configuration and manual updates. PHP, HTML5, CSS and *Cytoscape* JavaScript library were used for implementing the tool.

We also discussed with the instructors in order to agree on a set of graph plotting conventions that would be most suitable for their needs. Indeed, as mentioned in the Introduction, visualization methods should be easily correlated with humans' mental map (insight gaining process P4 [18]), thus reducing the comprehension effort. Therefore, we used directed graphs, in which nodes represent learners and links represent messages sent between the learners (on blog or Twitter). In order to expand the dimensionality of the information rendered in the graph, we introduced a color schema and magnitude schema for each graph element. Nodes shall be colored according to their affiliation to a certain community, i.e., nodes representing learners of the same team / community shall have the same color. In addition, links shall be colored according to their source node, in order to represent link direction. For example, if student A (red-colored node) sent a message to student B (green-colored node), then the link between nodes A and B shall be colored in red. The magnitude of each node (i.e., diameter) shall be directly proportional to a chosen SNA ranking: the larger the node, the higher the ranking. Thus, instructors can easily compare students according to a selected SNA ranking method, e.g., *PageRank* [14]. Furthermore, the thickness of each link shall be directly proportional to the strength / intensity of the collaboration between the two students; this can be computed through various methods, the simplest being the number of exchanged messages. Finally, nodes shall be labeled with a unique learner ID, in order to map the node to a particular learner.

We also decided to use force directed methods (FDM) for graph plotting [9], which generally produce aesthetically pleasing results. These methods are based on attractive and repulsive forces inspired from physics. Such forces attract nodes with high connectivity and repulse those with low connectivity, making the observation of communities of collaboration very intuitive. Moreover, the distance among nodes is inversely correlated with the strength of their influence on each other. Another advantage of FDM is their adaptability to various network traits, so they can be optimized from case to case.

Finally, StudentViz interface was designed in a simple and intuitive way. Two views are available, similar to *Gephi*: the *Main view* and the *Data view*. The *Main view* is further divided into 3 areas: *Options* area (left side), *Plotting canvas* area (center) and *Additional information* area (right side). Figure 2 provides a screenshot of StudentViz *Main view*. The *Data view* consists of a sortable grid of learners' attributes, including various SNA metrics. Similar views can also be found in *Gephi, Cytoscape* or *Tulip* under various names (e.g., *Data laboratory* in *Gephi*).

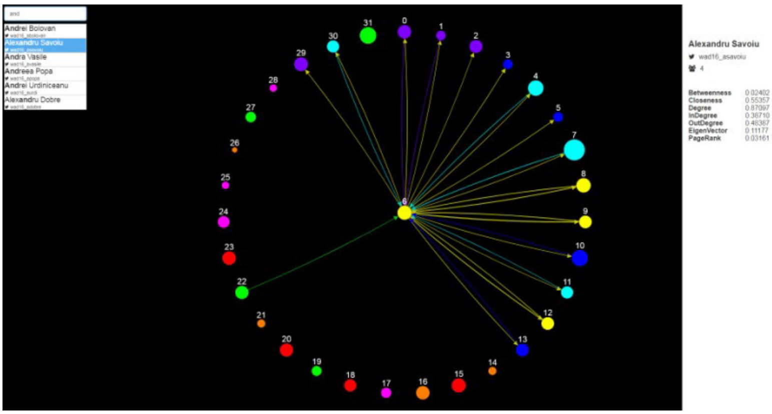


**Fig. 2.** StudentViz *Main view* - *Focus-circular* layout is employed; nodes' diameters are proportional to their *PageRank*, while their colors depict affiliation to a specific team

In what follows we present the *Main view* in more detail. The *Options* area allows instructors to interact with the visualizations and adjust them through various settings. Thus, as collaboration cannot be quantified by just one SNA metric, the teacher has the possibility to choose from several metrics: *betweenness, closeness, degree, in-degree, out-degree, Eigenvector and PageRank*. Through *betweenness* an instructor can determine the students that bridge community silos, those that facilitate the exchange of knowledge between communities. Students with high *closeness* values are positioned on various communication paths, playing an important role in knowledge diffusion. *Degree, in-degree* and *out-degree* centrality metrics can be used to determine the most / least active learners. Both *Eigenvector* and *PageRank* are measures of nodes importance that take into consideration qualitative and quantitative aspects; an important student is defined as one that has multiple collaborations with other important students. Additional information about these centrality metrics can be found in [4].

Another functionality provided in the *Options* area allows the instructor to select the graph plotting algorithm; available choices are: *WebCola*[3]*, Cose-Bilkent* [8], *circular* and *focus-circular*, which will be discussed in the next section. Furthermore, the instructor can also choose the focus of the visualization: individual learners, teams or communities. This functionality is achieved by applying a *reduction transformation* on graphs that include all students; learners of the same team are represented as one node, while filtering out intra-team collaborations. Furthermore, the nodes' color can depict team or community affiliation; teams are predetermined from the beginning of the semester, while communities are non-formal and self-regulated. Community detection is computed using a Laplacian method [13].

An additional option available to the instructor is to load various graphs created by the DtoG module (e.g., graph containing all social media interactions among students, graph containing only collaborations on the blog / Twitter). Finally, for easy identification of each student / team, an autocomplete search box is provided, in addition to the full list of students.

The center area of the *Main view* consists of a black canvas on which the graph is plotted. The canvas color was chosen in order to provide a high contrast for the graph nodes and edges. The instructor can reposition nodes through drag-and-drop functionality; he can also select one node for detailed inspection, which sets the graph plotting algorithm to *focus-circular* and opens the *Additional information* area.

Finally, the right side area of the *Main view* provides information regarding the specific node selected: student name, team, SNA metrics values. This area is only displayed upon selection of a node, otherwise it is hidden, to allocate a larger space for the plotting canvas.

[3] http://marvl.infotech.monash.edu/webcola

## 4 Illustrating Visualization Functionalities in StudentViz

In what follows we show how StudentViz answers instructors' visualization needs, as they were specified in the Introduction (VN1 - VN3). It also provides support for the general processes through which people gain insight when using an information visualization system (P1 - P4) [18].

The context of use is a course on Web Applications Design, taught to 4th year undergraduate students from the University of Craiova, Romania, during 2016-2017 winter semester. 32 students used eMUSE platform (and the associated social media tools) for communication and collaboration support, in a project-based learning (PBL) scenario. Students worked in teams of 4 peers in order to develop a relatively complex web application. Based on the social media traces collected by eMUSE, a total of 2224 collaboration links were extracted (263 having distinct source-target pairs). Therefore, a social graph with 32 vertices and 263 links was built. More details regarding the PBL scenario and the process of extracting the collaboration links from blog and Twitter can be found in [3]. That paper also provides various graphs rendered by *Gephi* tool, which required specific SNA expertise to produce; here we present the graphs rendered by our StudentViz tool, which can be easily used by the instructor with no specialist knowledge.

Thus, Fig. 3 and 4 provide a birds-eye view on learners' collaboration, by using *Cose-Bilkent* and *WebCola* FDM respectively. These algorithms have low computational workload in case of small graphs and are suited for interactive applications, as they avoid the overlapping of nodes. As seen in Fig. 3 and 4, the general pattern of collaboration can be easily spotted, thus supporting VN1 and P1. High and low density areas of collaboration can be easily identified in both figures. Learners in the high density area (with nodes tightly plotted together) are significantly involved in collaboration with members of diverse teams; students in low density areas are those that teachers should focus on in order to enhance collaboration.

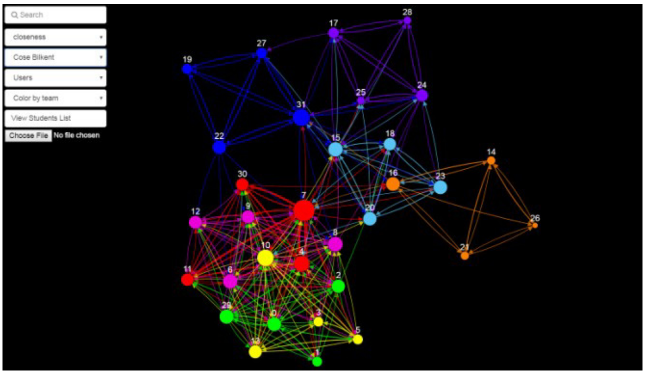


**Fig. 3.** Visualization provided by StudentViz using *Cose-Bilkent* plotting method. Nodes' diameters are proportional to *PageRank* metric, while their colors depict affiliation to a specific team.

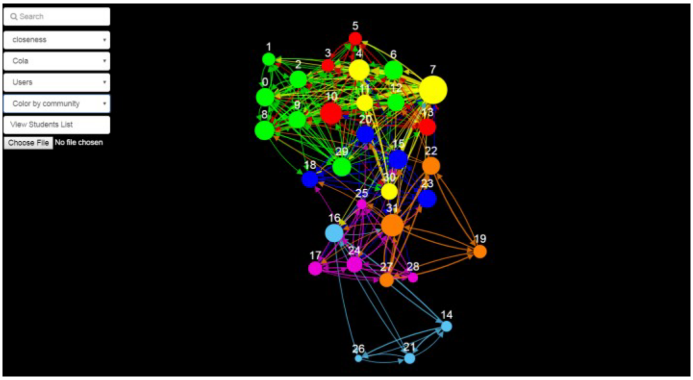


**Fig. 4.** Visualization provided by StudentViz using *WebCola* layout. Nodes' diameters are proportional to *PageRank* metric, while their colors depict affiliation to a specific community.

Although both algorithms produce similar visualizations, there are some variations that justify their complementary use. *WebCola* favors the identification of large communities of collaboration, as nodes are plotted in close proximity. However, this creates clutter, making smaller communities (teams) hard to spot. In turn, *Cose-Bilkent* favors the observation of smaller communities over large ones. Hence, *WebCola* and *Cose-Bilkent* also support VN2. In addition, these methods allow the discovery of the general structure and trends of collaboration, thus addressing P3.

Furthermore, instructors can better assess the collaboration between the teams by visualizing teams as nodes, like in Fig. 5 (which was obtained by applying the graph *reduction transformation*). Both Fig. 2 and Fig. 5 are illustrations of the *Focus-circular* plotting method, devised to clearly observe the status of a particular node (learner or team, respectively). The node of interest is positioned in the center of a circle on which the other nodes are plotted. Moreover, only the collaborations that involve the node of interest are rendered, to reduce unnecessary clutter and allow the instructor to focus on the particular node. Hence, these focus-circular visualizations address both VN2 (when nodes represent teams) and VN3 (when nodes represent students). Furthermore, P2 is also supported, as instructors can change the perspective on the dataset by selecting the node of interest.

Finally, the *Circular* plotting method, inspired by *Gephi's* circular layout, was devised to better observe the status of each team, as illustrated in Fig. 6. First, an average value of the selected SNA metric is computed for each team. The team with the highest ranking is plotted first, followed by the other teams according to their average metric rank, in a clockwise descending order. Individual nodes are also rendered in a clockwise descending order in the designated plot area for their team, according to their SNA metric value. Hence, the first plotted node depicts the learner with the highest selected metric value from the highest ranking team. This visualization method allows the teacher to observe the status of each team in comparison with other teams, but also the status of each learner in comparison with his fellow team members. Thus, both VN2 and VN3 are addressed by this visualization method; moreover, P3 is supported here, as instructors can discover collaboration patterns among teams and students.

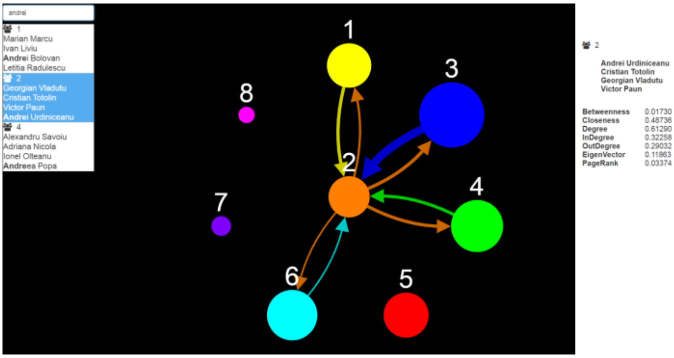


**Fig. 5.** Visualization provided by StudentViz using *Focus-circular* layout. Nodes represent teams and their diameters are proportional to *PageRank* metric.

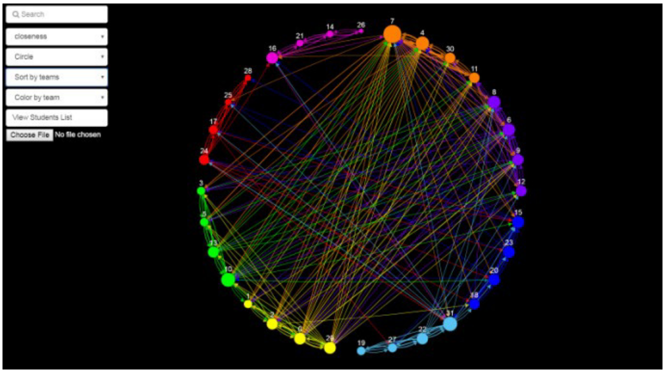


**Fig. 6.** Visualization provided by StudentViz using *Circular* layout. Nodes' diameters are proportional to *PageRank* metric, while their colors depict affiliation to a specific team.

## 5 Conclusions

Overall, with StudentViz an instructor is able to observe the collaboration status on different levels of network granularity, to emphasize different traits of the collaboration spectrum through various SNA metrics, and to turn their focus towards specific learners / teams. Hence, we consider that all basic visualization needs are successfully addressed, and an adequate support for gaining valuable insight is provided.

As future work, we plan to extend the tool with more visualizations, such as three-dimensional plotting methods and time-based representations. StudentViz could also be used to explore data from other social learning environments; the DtoG module is flexible and could easily be extended to accommodate different data sources.

## References

[1] Auber, D.: Tulip - A Huge Graph Visualization Framework. In: Graph Drawing Software, Mathematics and Visualization, Springer, pp. 105-126 (2004)
[2] Becheru, A., Popescu, E.: Design of a Conceptual Knowledge Extraction Framework for a Social Learning Environment Based on Social Network Analysis Methods. In: Proc. ICCC 2017, pp. 177-182 (2017)
[3] Becheru, A., Popescu, E.: Using Social Network Analysis to Investigate Students' Collaboration Patterns in eMUSE Platform. In: Proc. ICSTCC 2017, pp. 266-271 (2017)
[4] Boldi, P., Vigna, S., Axioms for Centrality. Internet Mathematics 10(3-4), 222-262 (2014)
[5] Bull, S., Duncan, A., Ginon, B., Kickmeier-Rust, M.: Educational Data Visualisation Approaches and Open Learner Modelling. Project report, available at: http://css-kmi.tugraz.at/mkrwww/leas-box/downloads/D4.1.pdf (2015)
[6] Crespo, P.T., Antunes, C.: Predicting Teamwork Results from Social Network Analysis. Expert Systems 32(2), 312-325 (2015)
[7] Desale, D.: Top 30 Social Network Analysis and Visualization Tools. Available at: www.kdnuggets.com/2015/06/top-30-social-network-analysis-visualization-tools.html (2015)
[8] Dogrusoz, U., Giral, E., Cetintas, A., Civril, A., Demir, E.: A Layout Algorithm For Undirected Compound Graphs. Information Sciences 179(7), 980-994 (2009)
[9] Fruchterman, T. M., Reingold, E. M.: Graph Drawing by Force-Directed Placement. Software: Practice and Experience 21(11), 1129-1164 (1991)
[10] Hagberg, A., Schult, D., Swart, P.: Exploring Network Structure, Dynamics, and Function Using NetworkX. In: Proc. SciPy2008 (7th Python in Science Conference), pp. 11-15 (2008)
[11] Jacomy, M., Venturini, T., Heymann, S., Bastian, M.: ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. PLoS ONE 9(6), e98679 (2014)
[12] Klerkx, J., Verbert, K., Duval, E.: Learning Analytics Dashboards. In: Handbook of Learning Analytics, SOLAR, DOI: 10.18608/hla17, pp. 143-150 (2017)
[13] Lambiotte, R., Delvenne, J.C., Barahona, M.: Laplacian Dynamics and Multiscale Modular Structure in Networks. arXiv:0812.1770 (2008)
[14] Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web. Technical Report, Stanford InfoLab (1999)
[15] Popescu, E.: Providing Collaborative Learning Support with Social Media in an Integrated Environment. World Wide Web - Internet and Web Information Systems 17(2), 199-212 (2014)
[16] Romero, C., Ventura, S.: Educational Data Mining: A Review of the State-of-the-Art. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 40(6), 601-618 (2010)
[17] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T.: Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research 13, 2498-2504 (2003)
[18] Yi, J.S., Kang, Y.A., Stasko, J.T., Jacko, J.A.: Understanding and Characterizing Insights: How Do People Gain Insights Using Information Visualization?. In: Proc. BELIV'08, article no. 4, ACM (2008)