

RESEARCH

Open Access



Analyzing students' collaboration patterns in a social learning environment using StudentViz platform

Alex Becheru, Andreea Calota and Elvira Popescu 

* Correspondence: popescu_elvira@software.ucv.ro; elvira.popescu@gmail.com

This is a substantially extended and revised version of the paper: Alex Becheru, Andreea Calota, Elvira Popescu, StudentViz: A Tool for Visualizing Students' Collaborations in a Social Learning Environment, Challenges and Solutions in Smart Learning, Lecture Notes in Educational Technology, Springer, pp. 77–86, 2018. Computers and Information Technology Department, University of Craiova, Bvd. Decebal, nr. 107, 200440 Craiova, Romania

Abstract

Understanding students' collaboration patterns is an important goal for teachers, who can thus obtain an insight into the collaborative learning process. Social network analysis and network visualizations are commonly used for exploring social interactions between learners. However, most existing network visualization platforms are deemed too complex by the teachers, who do not possess social network analysis expertise. Therefore, we propose an easy to use platform for visualizing students' collaboration patterns, called StudentViz. An overview of the tool, including a description of its implementation and functionalities, is provided in the paper. An illustration of how the tool can be used in practice, for investigating students' collaboration patterns in a social learning environment, is also included.

Keywords: Collaboration patterns, Collaborative learning, Social learning environments, Learning analytics, Network visualization, Social networks analysis

Introduction

Information visualization relies on the remarkable visual perception abilities of humans for pattern discovery (Yi et al., 2008). It employs interactive visual representations in order to amplify cognition (Klerkx et al., 2017) and generate “insight” (Bull et al., 2015).

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 (Klerkx et al., 2017).

In particular, social network analysis (SNA) and network visualizations have been used to investigate students' interactions taking place in educational environments (Crespo & Antunes, 2015; Romero & Ventura, 2010). Various SNA methods and metrics have been applied to measure and enhance collaboration and detect potential weak trainees (Maglajlic & Gutl, 2012), to rank learners and predict teamwork results (Crespo & Antunes, 2015), to create / interpret a social graph (Haythornthwaite & de Laat, 2010). In this context, our goal is to visualize collaboration patterns between students in a social media-based learning space, which has been less explored in the literature.

Indeed, during the past decade, social media tools have been increasingly adopted in educational settings, helping to build online learning networks, foster communication and collaboration between learners and encourage positive interactions (Dron & Anderson, 2014; Lumby et al., 2014). These technologies support contribution-based pedagogies, according to which students can learn efficiently by collaboratively creating learning resources and sharing them with peers (Hain & Back, 2008; Popescu, 2014). They are also in line with social learning theory (Bandura, 1977), according to which understanding is a product of participation in a community; social media tools can thus create an appropriate environment for conversations and interactions around specific problems, which trigger learning mechanisms (Ala-Mutka, 2009). Therefore, understanding students' collaboration patterns in such a social media-based learning space would bring valuable insight to the instructor (Becheru & Popescu, 2017a), helping to monitor the collaborative learning process and provide personalized interventions.

More specifically, in this paper we are interested in studying learners' collaboration in our eMUSE social learning environment (Popescu, 2014). The platform integrates several social media tools (blog, wiki, microblogging tool) that students use for communication and collaboration support. Learners' activity on these tools is tracked by the platform and used for various data visualizations (provided both for students and teachers), as well as for grading support. In addition, eMUSE offers basic administrative services and a peer assessment module for the learners. More details about the platform can be found in (Popescu, 2014; Popescu, 2016).

We have already proposed an SNA-based conceptual framework for knowledge extraction and visualization from a social learning environment in (Becheru & Popescu, 2017a), which we have validated in (Becheru & Popescu, 2017b). *Gephi* network analysis tool (Jacomy et al., 2014) was used for all computations and graph visualizations, which adequately fitted the 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 (Yi et al., 2008):

- 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 related research and existing network visualization platforms is included in section “[Related Work](#)”. StudentViz functionalities and implementation details are presented in section “[StudentViz Tool Description](#)”. Its visualization capabilities are exemplified and validated in section “[Illustrating Visualization Functionalities in StudentViz](#)”. Some insights on students’ collaboration that teachers can obtain by means of the tool are discussed in section “[Using StudentViz to Explore Students’ Collaboration Patterns](#)”. Finally, section “[Conclusions and Future Work](#)” outlines some conclusions and future research directions.

Related work

Networks or graphs are a common visualization method in educational settings (Bull et al., 2015). 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 (Desale, 2015); 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* (Shannon et al., 2003) 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* (Auber, 2004) is a visualization platform for relational data. It provides highly flexible visualization and a wide range of analysis capabilities for various research fields.
- *Gephi* (Jacomy et al., 2014) aims to be a general platform of analysis and visualization for all kinds of networks. Its clear design and resemblance to *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 (Becheru & Popescu, 2017b).

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 in the following section.

As far as research studies are concerned, there are several papers relying on SNA methods to investigate students' collaboration in educational environments. A review of early works is provided in (Romero & Ventura, 2010), which focuses on patterns of academic collaboration, the structure of online learning communities, as well as the use of collaborative filtering techniques for generating personalized recommendations for students. Haythornthwaite and de Laat (2010) also provide an overview on the use of SNA techniques in learning networks, including also a discussion on how to create/interpret a social graph.

De Laat et al. (2007) explored interaction patterns of students in a learning management system (LMS) by using SNA in conjunction with content analysis and student interviews. Macfadyen and Dawson (2010) also employed network analysis of asynchronous discussion forums to investigate students' engagement in a course; this was performed in the broader context of analyzing LMS tracking data to predict student grades. One of their findings was that students tend to cluster with peers of similar academic ability; moreover, they obtained insights into the development of the student learning community by identifying disconnected students, patterns of student-to-student communication, and instructor positioning within the network.

Maglajlic and Gutl (2012) used various SNA techniques (including *cliques*, *centrality* and *density*) to monitor and enhance collaboration in an educational environment, to detect potentially weak trainees and to assign trainees in the appropriate tutored group. Crespo and Antunes (2015) relied on diverse variants of *PageRank* algorithm in order to rank learners and predict teamwork results.

These studies further emphasize the importance and usefulness of applying SNA in computer-supported collaborative learning environments. However, all of them were performed by SNA experts and could not be applied by the average teacher, with limited SNA knowledge. Therefore the need to make these visualizations readily available to the instructor, by means of easy to use, dedicated tools. Few initiatives have been proposed in this respect: SNAPP (Social Networks Adapting Pedagogical Practice) (Dawson et al., 2010) is one such tool, which extracts student data from a LMS discussion forum and provides easily interpretable visualizations and social network metrics. Meerkat-ED (Rabbany et al., 2014) is another SNA-based tool designed for analyzing students' collaboration in discussion forums; it provides snapshots of students' interactions, their ranking (leaders / peripheral students) and collaborative groups, together with the term co-occurrence network. While these two platforms are aimed at teachers, CanvasNet (Chen et al., 2018) is another easy to use tool designed for students. It extracts data from Canvas LMS discussion forums and combines SNA for social interactions with lexical analysis for conceptual engagement (by providing snapshots of trending concepts in students' posts). The main goal of CanvasNet is to promote student reflection and foster student discussion and engagement in online classes.

It is worth mentioning that all the tools presented above analyze discussion forum data, which is extracted from learning management systems. By contrast, the tool that we propose relies on social media data (tweets and blog posts / comments), retrieved from eMUSE social learning environment. In addition, our tool is more flexible, providing also team-based collaboration visualizations, as well as comparisons between blog

and microblog-based student interactions. It thus covers a well defined niche in the context of social media-based learning spaces, as described next.

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

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 (sociograms) 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. As far as the blog is concerned, two types of messages are recorded: posts and comments. For each post we established links between the student who wrote it and all his/her fellow team mates, as the instructional scenario implies the existence of one blog per team. For each comment, we created a link between its author and the student who wrote the initial post. As far as Twitter is concerned, we established the links based on the built-in referencing mechanism to determine the addressee(s). For a more nuanced representation of interactions, we used several link weighting methods, taking into account the number of messages and some of their characteristics (e.g., re-tweeting mechanism or URL-sharing). More details regarding the types of interactions (collaborations) taken into account and the weighting methods are included in (Becheru & Popescu, 2017b).

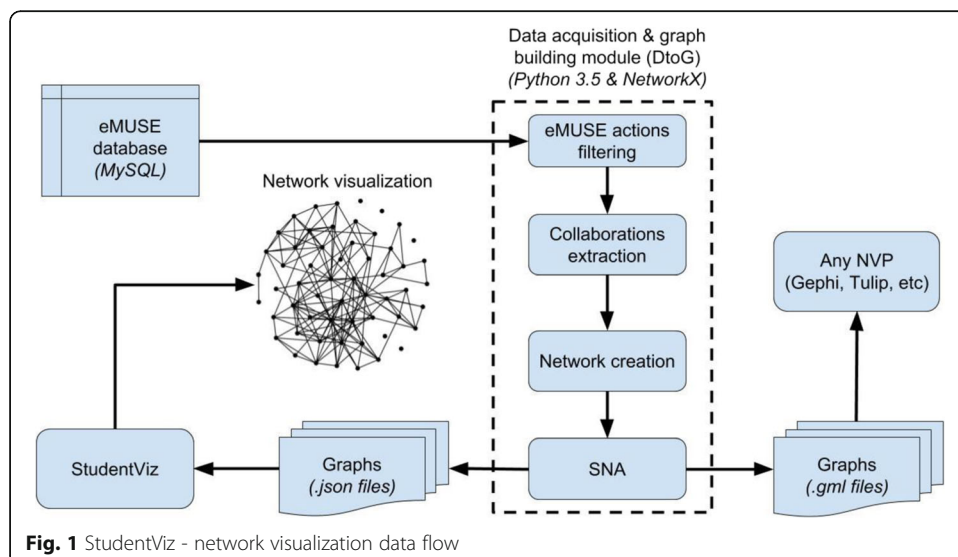


Fig. 1 StudentViz - network visualization data flow

In order to assess learners' performance and compute their status in the community, we rely both on classical graph theory metrics such as: degree, in-degree, out-degree, and SNA centrality metrics: closeness, betweenness, eigenvector centrality and PageRank. The degree-related metrics were chosen as they can assess a student's activity in terms of the number of collaborations he/she establishes (either initiated by him/her or by others). However, this number does not capture other aspects of collaboration, such as the position of a person in a group. Learners with high betweenness values connect community silos and allow information to traverse from one group to another. Also, closeness centrality depicts the involvement of a student in the entire learning environment. This metric can be interpreted as the distance from the core of learners that best support collaboration in their environment; hence, the smaller the value the better. PageRank and eigenvector centrality are measures that consider both the number of collaborations established and the position of each student in the social graph; in other words, not only the number of collaborations is important but also whom one collaborates with. These last two metrics can be used to determine an overall status for each student, while the others refer to specific aspects of the collaboration spectrum, together forming a comprehensive picture. Our choice of metrics was made taken into account their popularity across SNA studies as well as their comprehensibility from teachers' point of view, leading to a relevant set of measures for determining students' collaboration status (Becheru & Popescu, 2017a; Becheru & Popescu, 2017b). Details on the mathematical foundations of SNA metrics used can be found in (Boldi & Vigna, 2014).

Once the graphs/sociograms are computed by DtoG module, they 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 (Hagberg et al., 2008).

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 conferred 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 (Yi et al., 2008)), 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* (Page et al., 1999). Furthermore, in order to map the node to a particular learner, nodes shall be labeled with a unique learner ID.

In addition, 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. We have also considered more complex methods, which capture the qualitative aspects of collaboration. Thus, for Twitter interactions, we considered that the re-tweeting (RT) mechanism is a marker of more intense collaboration. This process requires a higher cognitive effort to review the content and signal that a piece of information is of particular interest. Link sharing is also a highly effective manner of collaboration, as the practice of recommending a tutorial/book/course is a widespread method of helping peers; hence, messages containing recommended links are given a higher weight when computing collaboration strength between two students.

For a higher graph granularity analysis, we employed a graph transformation method called *reduction*; this aggregates a set of nodes into one single node, representative for the entire set. In our scenario, reduction transformation was used to visualize collaborations between teams. In addition, StudentViz offers instructors the possibility to analyze collaborations for specific time periods, not just for the entire semester. This was done by restricting the collaborations considered for the construction of graphs to those that were established during a specific time period (e.g., 2 weeks). Finally, collaborations present in the graph can be filtered by their source: Twitter, Blogger or both.

We also decided to use force directed methods (FDM) for graph plotting (Fruchterman & Reingold, 1991), 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 (see Fig. 7). Similar views can also be found in *Gephi*, *Cytoscape* or *Tulip* under various names (e.g., *Data laboratory* in *Gephi*).

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

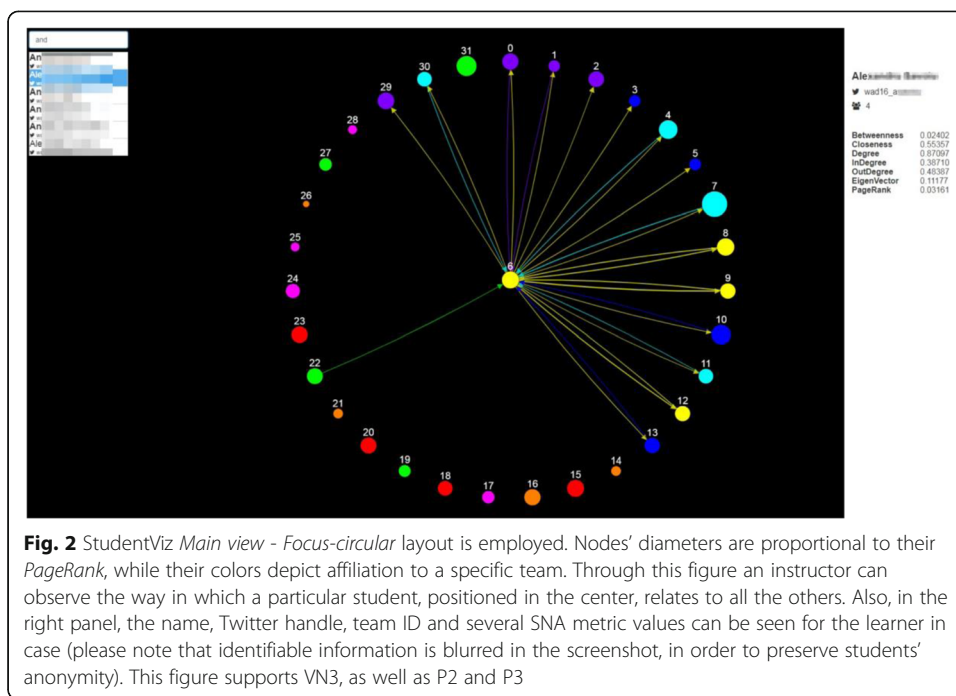


Fig. 2 StudentViz Main view - Focus-circular layout is employed. Nodes' diameters are proportional to their PageRank, while their colors depict affiliation to a specific team. Through this figure an instructor can observe the way in which a particular student, positioned in the center, relates to all the others. Also, in the right panel, the name, Twitter handle, team ID and several SNA metric values can be seen for the learner in case (please note that identifiable information is blurred in the screenshot, in order to preserve students' anonymity). This figure supports VN3, as well as P2 and P3

important students. Additional information about these centrality metrics can be found in (Boldi & Vigna, 2014).

Another functionality provided in the *Options* area allows the instructor to select the graph plotting algorithm; available choices are: *WebCola*,³ *Cose-Bilkent* (Dogrusoz et al., 2009), *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 (Lambiotte et al., 2008).

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; she/he can also select one node for detailed inspection, which sets the graph plotting algorithm to *focus-circular* and opens the *Additional information* area.

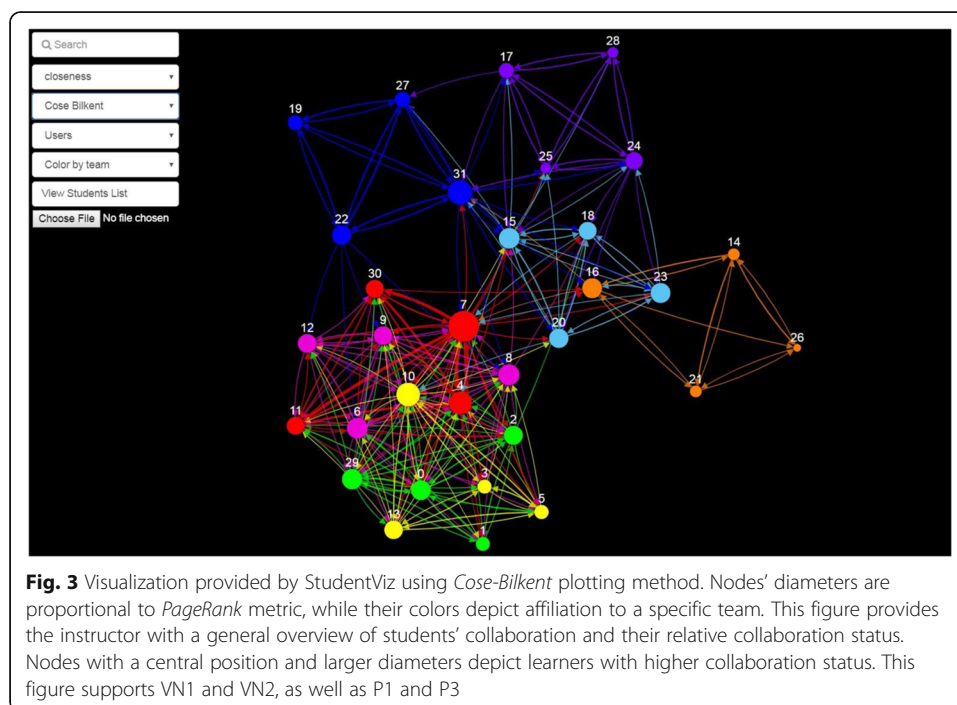
Finally, the right side area of the *Main view* provides information regarding the specific node selected: student name, team, SNA metrics values. This area is only displayed upon selection of a node, otherwise it is hidden, to allocate a larger space for the plotting canvas.

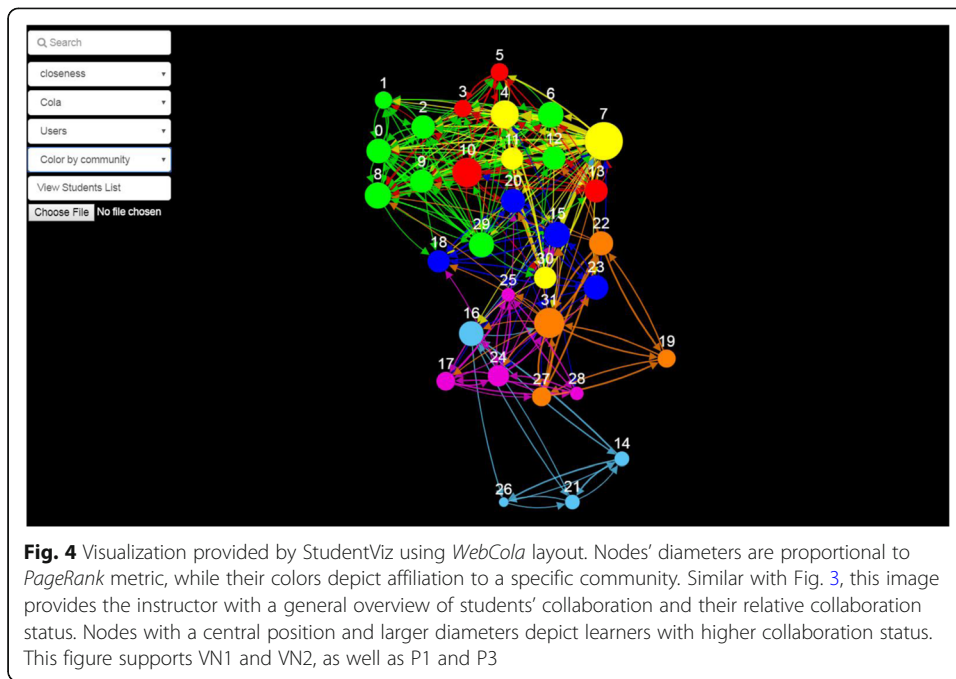
Illustrating visualization functionalities in StudentViz

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) (Yi et al., 2008).

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 (Becheru & Popescu, 2017b). 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.

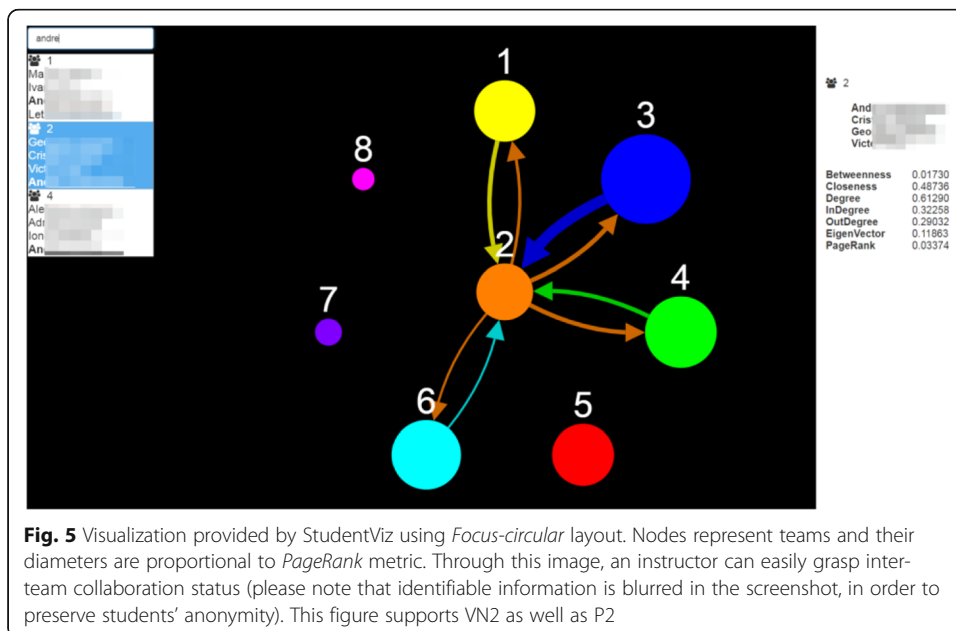
Thus, Figs. 3 and 4 provide a birds-eye view on learners' collaboration, by using the option of *Cose-Bilkent* and *WebCola* FDM respectively. These algorithms have low computational workload in case of small graphs and are suited for interactive applications, as they avoid the overlapping of nodes. As seen in Figs. 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.





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, as in Fig. 5 (which was obtained by applying the graph *reduction*



transformation). Both Figs. 2 and 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/her 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.

Using StudentViz to explore students' collaboration patterns

In what follows, we discuss some of the insights on students' collaboration patterns that teachers can obtain by using StudentViz. By presenting these insights and their manner of discovery we aim to provide better evidence on the utility of our proposed tool. Our analysis refers to the dataset presented in the previous section; however, we argue that similar insights could be obtained from different educational datasets / contexts.

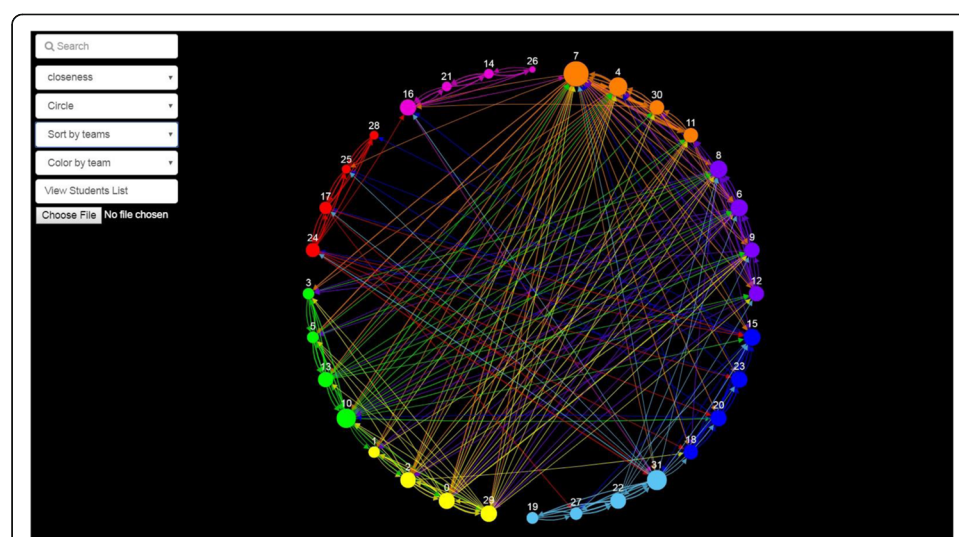


Fig. 6 Visualization provided by StudentViz using *Circular* layout. Nodes' diameters are proportional to *PageRank* metric, while their colors depict affiliation to a specific team. Through this image, the instructor can observe the status of each team in relation with other teams, as well as the status of each learner in relation with his/her team peers. This figure supports VN2 and VN3, as well as P3

As mentioned in section “StudentViz Tool Description”, StudentViz provides a *Data view*, which is a tabular representation of SNA metrics values for each student / team (see Fig. 7). The main utility of this view is the ability to compare and rank students / teams according to the SNA metrics considered. Hence, the teacher can observe a student’s performance in comparison with others and perceive the magnitude of the differences through the computed values. This is a good starting point for an in-depth understanding of students’ collaboration patterns; insights can be discovered by observing some collaboration particularities in *Data view* and then confirming them using *Main view* or vice versa.

Thus the teacher could compare the performance of team members by simply ranking the students in *Data view* by *Team*. She/he could subsequently take remedial actions for the underperforming teams. Such teams could be further analyzed from the graphs provided in the *Main view* of StudentViz. For example, the instructor can visualize a graph that depicts communication between teams (aggregated communication among students of each team), as illustrated in Fig. 5. As the diameter of the nodes is proportional to the *PageRank* metric, the larger the node the higher the collaboration degree of the respective team.

Another interesting opportunity for the teacher is to analyze the evolution of students’ collaboration patterns over the course of the semester. As mentioned in section “StudentViz Tool Description”, with the help of the DtoG module, StudentViz is able to provide the instructor with visualizations and data not only for the entire semester, but also for specific time periods. Figures 8 and 9 illustrate this capability, by providing collaboration visualizations for two time windows: weeks 5&6 and weeks 11&12 respectively. This functionality can provide valuable insights on the development and evolution of the learning/collaboration network. By looking at the inter-team collaboration patterns that form higher order communities, we notice that teams involved in

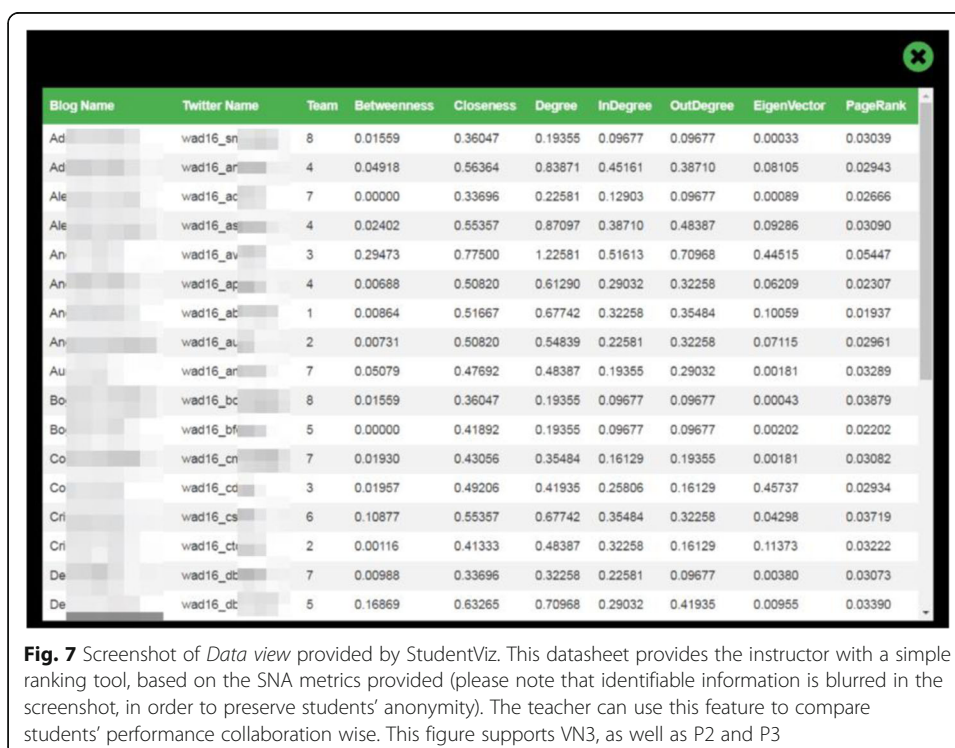
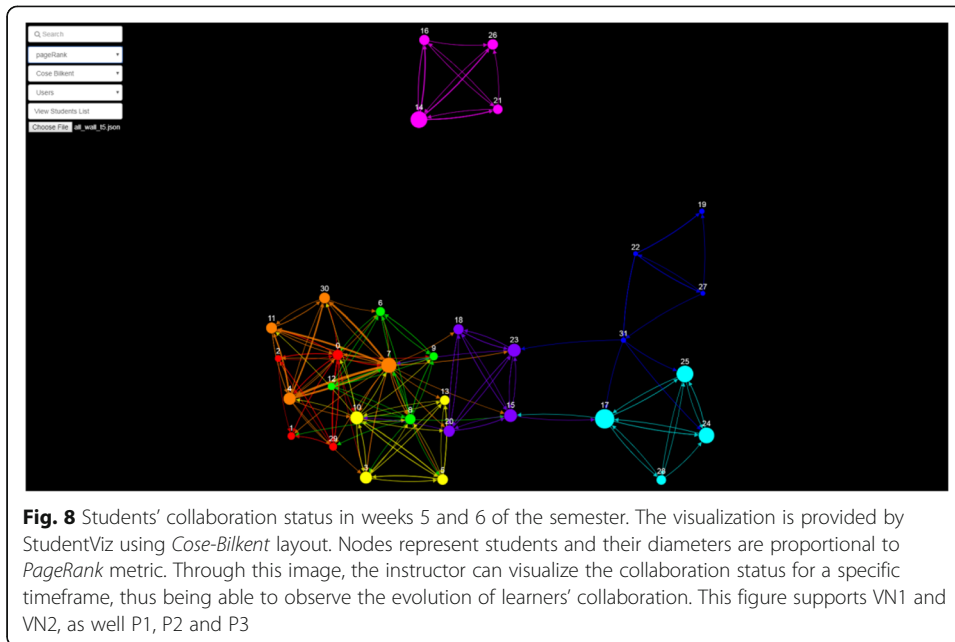


Fig. 7 Screenshot of *Data view* provided by StudentViz. This datasheet provides the instructor with a simple ranking tool, based on the SNA metrics provided (please note that identifiable information is blurred in the screenshot, in order to preserve students’ anonymity). The teacher can use this feature to compare students’ performance collaboration wise. This figure supports VN3, as well as P2 and P3



stable communities (over time) have more intense social interactions. Hence, collaboration among teams appears more beneficial when it is relatively stable/lasting.

In order to further analyze the collaboration patterns, we created a tabular representation of the community formation in our learning scenario, as shown in Table 1. The table was generated by examining the collaboration graphs shown by StudentViz for weeks 5&6 (timeframe T1), 7&8 (T2), 9&10 (T3), 11&12 (T4). We can thus visualize the communities of students that were formed in each specific timeframe. The choice of time windows is justified by the fact that students had to make intermediary project presentations every two weeks, so there is a milestone presentation in each timeframe.

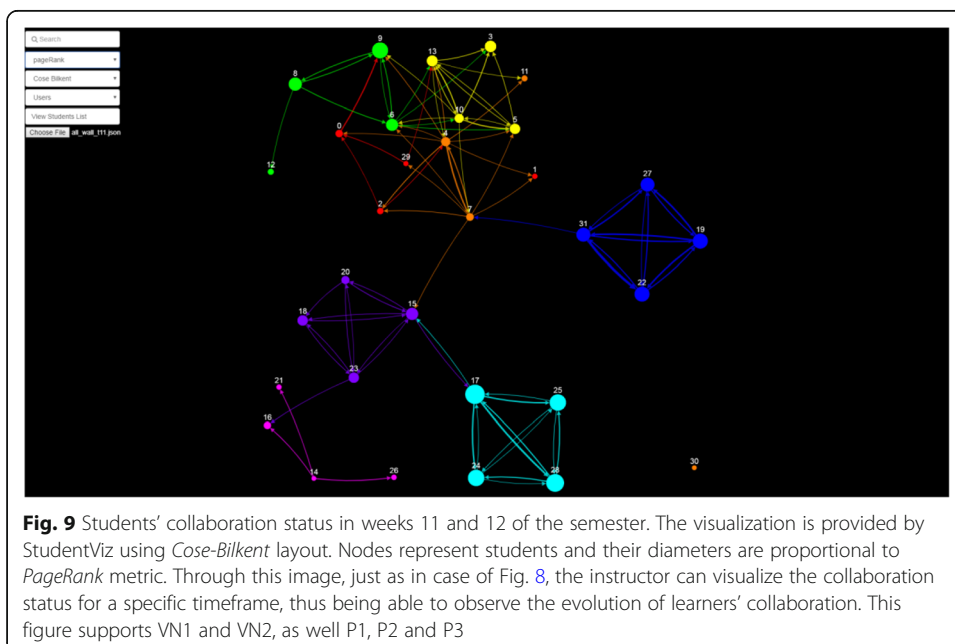


Table 1 Tabular representation of the communities arising from collaborations among teams

	PAGERANK	Timeframe T1	Timeframe T2	Timeframe T3	Timeframe T4
TEAM 1	15	a	a	a	a
TEAM 2	14	b	a	b	b
TEAM 3	21	a	a	a	a
TEAM 4	16	c	c	a	3a, 1b
TEAM 5	8	d	d	d	d
TEAM 6	13	e	d	g	e
TEAM 7	10	d	f	g	h
TEAM 8	3	h	h	h	h

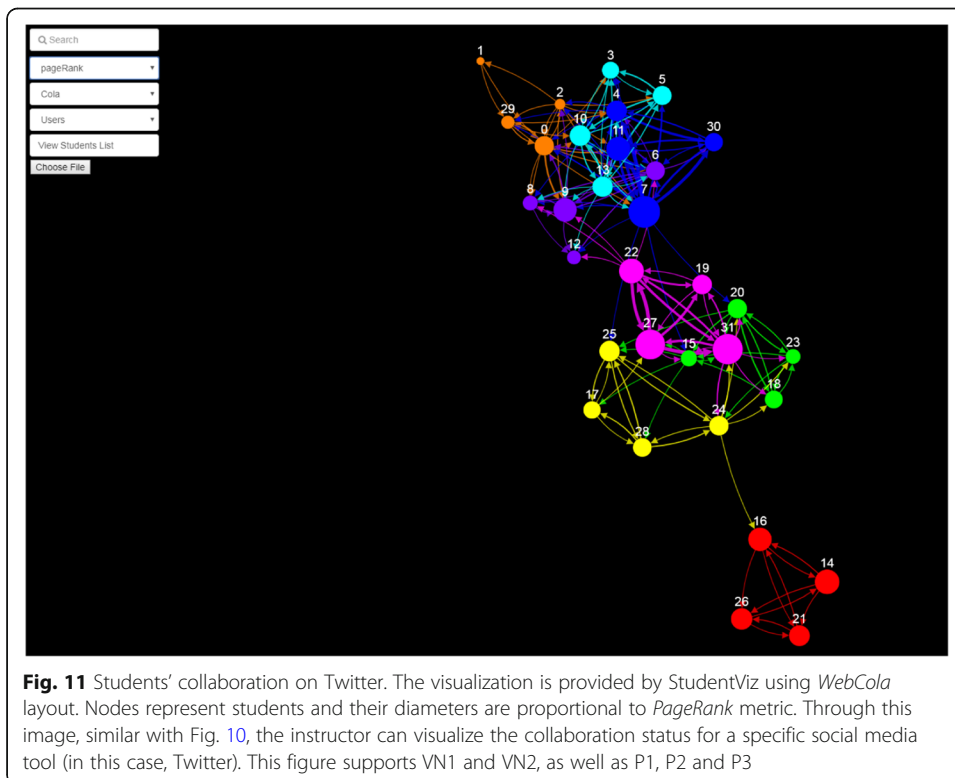
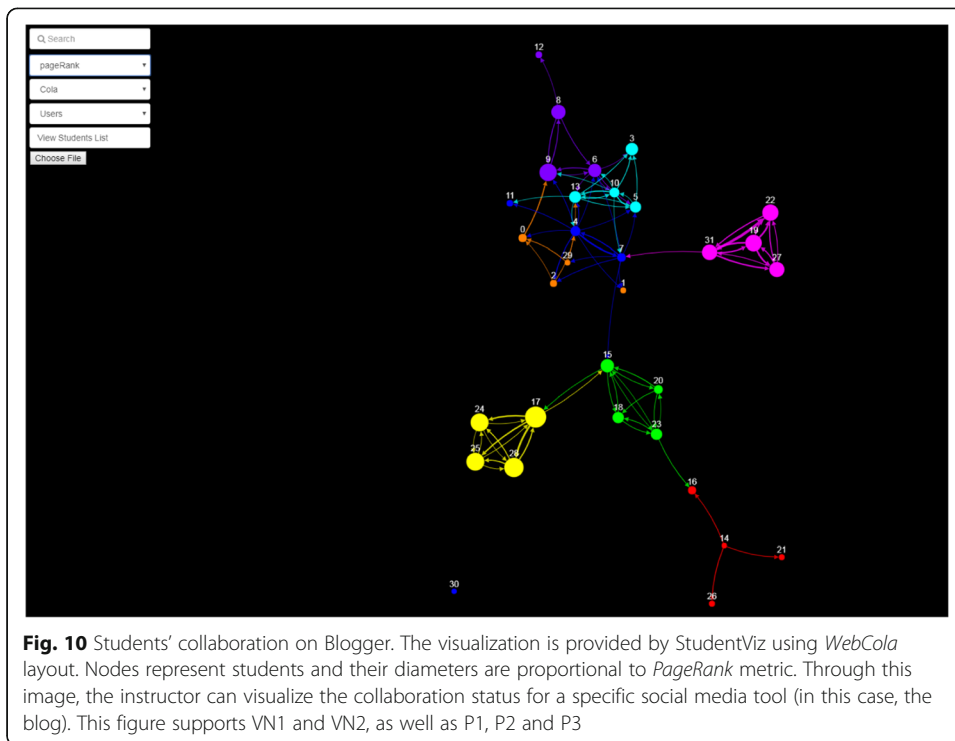
Lines represent the teams, while colored columns represent 4 consecutive timeframes. The *PageRank* column depicts the value obtained by each specific team in the teams graph; for readability reasons, the value of the *PageRank* has been multiplied by 100 and rounded. Cell color depicts affiliation of at least 3 members of a team to a specific community. The letters in each cell bring additional information about community membership: in case a cell has only one letter, then all the members of that respective team are part of the same community (e.g.: in *timeframe T1*, all members of *Team 1* are part of community *a*). Otherwise, a number in front of each letter represents the count of team members that are part of the community specified by that letter (e.g.: in *timeframe T4*, 3 members of *Team 4* belong to community *a* and 1 member to community *b*).

By analyzing Table 1, we can see that teams 1 & 3 form a community (*backbone*) from start to finish, hence they collaborate in a stable manner. This community is joined in timeframe T3 by team 4, which is sustained also in T4, with the exception of one student. Team 2 is generally forming its own community with the exception of timeframe T2, when it is part of the *backbone* community. Hence, we can argue that collaboration among teams 1–4 is relatively stable, at least compared with teams 5–8. We can also note that their collaboration is substantial (as can be seen in Figs. 3, 4 and 5) and that they have the highest *PageRank* values (as can be seen in the *Data view*).

As far as the other teams are concerned, we can notice some similarities between the inter-team collaboration patterns of teams 5 and 8: both try to collaborate with other teams but no collaboration is sustained for more than one timeframe. However, there are also substantial differences: team 5 is depicted in the majority of plots as sustaining a closer collaboration with teams 1–4 than teams 6–8, hence being the most collaborative among teams 5–8. On the contrary, team 8 is at the opposite end of the collaboration spectrum, with very few collaborations established outside the team. Similarly, teams 6 & 7 try various collaborations with other teams, however no stable communities are formed (lasting at least two consecutive timeframes).

Overall, three general collaboration patterns can be identified for our learning scenario: teams that collaborate in a stable manner (teams 1, 3, 4), teams that collaborate but not in a stable manner (teams 2, 5, 6, 7) and teams that barely collaborate (team 8).

Another analysis that can be performed with the help of StudentViz is the comparison between collaboration status on each social media tool. The teacher can visualize both the overall (aggregated) collaborations (as we have seen so far) and the specific collaborations taking place on Blogger and Twitter, as can be seen in Figs. 10 and 11 respectively. This comparison is important, as the tools play distinct roles in the project-based learning scenario: Twitter supports faster and shorter exchanges of notifications, news and links, while Blogger is used for longer, more crafted posts with questions, full-fledged solutions and progress reports. It is therefore expected that the number of tweets and overall collaboration on Twitter (predominantly a social space) has a higher level than on Blogger (predominantly a content space). Nevertheless, a



balanced collaboration on both tools is desirable, as can be noted for some of the teams in Figs. 10 and 11. These visualizations provide instructors with an additional insight into students' collaboration patterns.

Conclusions and future work

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; also, studies can be conducted by filtering the collaborations by media source and/or timeframe. Hence, we consider that all basic visualization needs are successfully addressed, and an adequate support for gaining valuable insight is provided.

The main advantage of StudentViz compared with generic network visualization platforms (as surveyed in section "Related Work") is its ability to provide simple and pedagogically relevant SNA visualizations, while being easy to use by the average teacher. This is in line with the principles behind similar tools (Dawson et al., 2010; Rabbany et al., 2014), but applied in novel educational settings; students' collaboration is supported by social media tools (interaction data being retrieved from eMUSE social learning environment), rather than by classic discussion forums (data being retrieved from traditional learning management systems).

As future work, we plan to extend StudentViz with more visualizations, such as three-dimensional plotting methods, time-based approaches and additional representations, like the one depicted in Table 1. Another desirable functionality is the visualization of two graphs side by side, with the nodes having a mirrored position, to provide an easier comparison. A graphical selection tool could also be developed, that will allow teachers to extract sub-graphs for further analysis. The *Data view* may be extended with advanced filtering methods, that will give instructors more options to focus on particular parts and traits of the graphs. A module where teachers can easily define their personalized metrics may also be included. Adding a content analysis dimension to StudentViz, as in (Chen et al., 2018) or (Rabbany et al., 2014), would be an interesting research direction. This could provide further information to the instructor, that could be used to offer personalized feedback and interventions to the students.

Finally, StudentViz could 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. Consequently, our goal is to perform also more experimental studies, in different settings and with a larger number of students and teachers; collecting and analyzing instructors' subjective experience and opinions on StudentViz would provide additional validation for our platform. Furthermore, investigating students' collaboration patterns over several semesters and courses could provide a valuable insight into the social learning process.

Endnotes

¹<http://cuttlefish.sourceforge.net>

²<http://www.visone.info/html/about.html>

³<http://marvl.infotech.monash.edu/webcola>

Funding

This work was supported by QFORIT programme, University of Craiova, 2017.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

AB contributed to the conception, design and implementation of StudentViz platform, performed analysis and interpretation of data and was involved in drafting and revising the manuscript. AC contributed to the design and implementation of StudentViz platform and was involved in drafting the manuscript. EP contributed to the conception and design of the StudentViz platform, performed the experimental study, collected the data and was involved in drafting and revising the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 13 June 2018 Accepted: 27 August 2018

Published online: 18 September 2018

References

- K. Ala-Mutka, *Review of Learning in ICT-Enabled Networks and Communities* (JRC-IPTS, 2009) <http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=2721>
- D. Auber, in *Graph Drawing Software, Mathematics and Visualization*. Tulip - a huge graph visualization framework (Springer, 2004), pp. 105–126
- A. Bandura, *Social Learning Theory* (Prentice Hall, Englewood Cliffs, 1977)
- A. Becheru, E. Popescu, in *Proc. ICCS 2017*. Design of a conceptual knowledge extraction framework for a social learning environment based on social network analysis methods (2017a), pp. 177–182
- A. Becheru, E. Popescu, in *Proc. ICSTCC 2017*. Using social network analysis to investigate students' collaboration patterns in eMUSE platform (2017b), pp. 266–271
- P. Boldi, S. Vigna, Axioms for centrality. *Internet Math.* **10**(3–4), 222–262 (2014)
- S. Bull, A. Duncan, B. Ginon, M. Kickmeier-Rust, *Educational Data Visualisation Approaches and Open Learner Modelling*. Project report (2015) <http://css-kmi.tugraz.at/mkrwww/leas-box/downloads/D4.1.pdf>
- B. Chen, Y.H. Chang, F. Ouyang, W. Zhou, Fostering student engagement in online discussion through social learning analytics. *Internet High. Educ.* **37**, 21–30 (2018)
- P.T. Crespo, C. Antunes, Predicting teamwork results from social network analysis. *Expert. Syst.* **32**(2), 312–325 (2015)
- S. Dawson, A. Bakharia, E. Heathcote, in *In: Proc. 7th International Conference on Networked Learning 2010*. SNAPP: Realising the affordances of real-time SNA within networked learning environments (2010), pp. 125–133
- M. De Laat, V. Lally, L. Lipponen, R.J. Simons, Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis. *Int. J. Comput.-Support. Collab. Learn.* **2**, 87–103 (2007)
- D. Desale, (2015). *Top 30 Social Network Analysis and Visualization Tools*. Available at: www.kdnuggets.com/2015/06/top-30-social-network-analysis-visualization-tools.html
- U. Dogrusoz, E. Giral, A. Cetintas, A. Civril, E. Demir, A layout algorithm for undirected compound graphs. *Inf. Sci.* **179**(7), 980–994 (2009)
- J. Dron, T. Anderson, *Teaching Crowds. Learning and Social Media*. (AU press, Athabasca University, 2014)
- T.M. Fruchterman, E.M. Reingold, Graph drawing by force-directed placement. *Software: Practice and Experience* **21**(11), 1129–1164 (1991)
- A. Hagberg, D. Schult, P. Swart, in *Proc. SciPy2008 (7th Python in Science Conference)*. Exploring network structure, dynamics, and function using NetworkX (2008), pp. 11–15.
- S. Hain, A. Back, Personal learning journal - course design for using weblogs in higher education. *The Electronic Journal of e-Learning* **6**(3), 189–196 (2008)
- C. Haythornthwaite, M. de Laat, in *In: Proc. 7th International Conference on Networked Learning 2010*. Social networks and learning networks: Using social network perspectives to understand social learning (2010), pp. 183–190
- M. Jacomy, T. Venturini, S. Heymann, M. Bastian, ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* **9**(6), e98679 (2014)
- J. Klerkx, K. Verbert, E. Duval, in *Handbook of Learning Analytics*. Learning Analytics Dashboards (2017) <https://doi.org/10.18608/hla17>, pp. 143–150
- R. Lambiotte, J.C. Delvenne, M. Barahona, *Laplacian Dynamics and Multiscale Modular Structure in Networks*. arXiv:0812.1770 (2008)
- C. Lumby, N. Anderson, S. Hugman, Apres Le deluge: Social media in learning and teaching. *The Journal of International Communication* **20**(2), 119–132 (2014)
- L.P. Macfadyen, S. Dawson, Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Comput. Educ.* **54**, 588–599 (2010)
- S. Maglajlic, C. Gutl, in *Proc. ICL 2012 (15th International Conference on Interactive Collaborative Learning)*. Efficiency in e-learning: Can learning outcomes be improved by using social networks of trainees and tutors? (2012)
- L. Page, S. Brin, R. Motwani, T. Winograd, *The PageRank Citation Ranking: Bringing Order to the Web*. Technical Report, Stanford InfoLab (1999)
- E. Popescu, Providing collaborative learning support with social media in an integrated environment. *World Wide Web - Internet and Web Information Systems* **17**(2), 199–212 (2014)

- E. Popescu, L.M. Petrosanu, in *Innovations in Smart Learning, Lecture Notes in Educational Technology*. Integrating a peer evaluation module in a social learning platform (Springer, 2016), pp. 141-150
- R. Rabbany, S. Elatia, M. Takaffoli, O.R. Zaiane, in *Educational Data Mining, Applications and Trends*. Collaborative learning of students in online discussion forums: A social network analysis perspective (Springer, 2014), pp. 441–466
- C. Romero, S. Ventura, Educational data mining: A review of the state-of-the-art. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **40**(6), 601–618 (2010)
- P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003)
- J.S. Yi, Y.A. Kang, J.T. Stasko, J.A. Jacko, in *Proc. BELIV'08*. Understanding and characterizing insights: How do people gain insights using information visualization? (ACM, 2008), article no. 4

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
