Visual Analysis of Contact Patterns in School Environments

Jean R. PONCIANO[1], Claudio D. G. LINHARES[1], Sara L. MELO[2], Luciano V. LIMA[2], Bruno A. N. TRAVENÇOLO[1]

[1]*Faculty of Computing, Federal University of Uberlândia, Brazil*
[2]*Faculty of Electrical Engineering, Federal University of Uberlândia, Brazil*
e-mail: {jeanrobertop, claudiodgl, saraluziamelo, lucianovieiralimaster, travencolo}@gmail.com

Abstract. Information Visualisation strategies can be applied in a variety of domains. In the context of temporal networks, i.e., networks in which interactions between individuals occur throughout time, efforts have been conducted to develop visual approaches that allow finding interaction patterns, anomalies, and other behaviours not previously perceived in the data. This paper presents two case studies involving real-world education networks from a primary school and a high school. For this purpose, we used the *Massive Sequence View (MSV)* layout with the *Community-based Node Ordering (CNO)* method, two well established approaches for visual analysis of temporal networks. Our results show that the identified patterns involving students/students and students/ teachers represent important information to benefit and support decision making about school management and teaching strategies, especially those related to strategic group formation.

Keywords: temporal networks, collaborative learning, visualisation in education, learning analytics, dynamic networks, network communities, node reordering, massive sequence view.

1. Introduction

Several researchers have been focusing in recent years on data modeling involving data sets that represent any kind of interaction among elements. A structure highly adopted for such representation is the so-called complex network (Albert and Barabási, 2002). In a complex network, each element is called *node* and the interaction that links two nodes is called *edge* (or *connection*). This type of network can be seen in several knowledge areas, as, for example, computer science (e.g., web pages connected by hyperlinks), sociology (e.g., people connected due to a collaboration), and biology (e.g., networks of proteins with similar functions) (Estrada, 2015).

Only the information of nodes and edges in a network, however, may not be enough depending on the desired type of analysis, as occurs, for example, in the study of social

network evolution. In such situations, it is important to know *when* the interactions occur, in addition to nodes and edges. Networks that offer these three dimensions (node, edge, and time) are called *temporal networks* (Holme and Saramäki, 2012).

A temporal network can be analyzed by using different approaches. Statistical analyses represent a common approach and are useful to identify specific trends and patterns in the data. Only the numeric output, however, may represent a "black-box" to the user, thus impairing pattern comprehension. Another approach involves Information Visualisation, whose strategies help in data analysis by providing interactive and graphical computational tools, thus including the user in the entire process of exploration and validation (Ware, 2012). An adequate Information Visualisation strategy allows a visual analysis that is as much intuitive as possible and also helps the user in finding unexpected patterns, anomalies, and other behaviours in the data.

In the educational context, the visual analysis of temporal networks may be useful to facilitate the full understanding of how the interactions involving students and teachers occur over time, as well as to allow the identification of patterns that influence the learning process. In this way, a school class could be considered as a small network composed by students and teachers interacting among themselves over time. By performing a visual analysis in this type of network, teachers may identify factors that impair and/or assist in the teaching and learning process.

In this paper, we present case studies involving two real-world education networks that represent a primary and a high school. Both networks were analyzed using well established temporal network visualisation strategies. By using these strategies, we analyzed the interactions among students/students and students/teachers, as well as the associated metadata. Although none of these networks provide metadata about individual or group performances on tests and exams, we show that the analysis of both network structure and interaction dynamics allows the identification of meaningful patterns that can be used to support and optimize the teacher's decision making, especially those related to strategic group formation and student performance analysis.

This paper is organized as follows. Section 2 presents studies related to the use of Information Visualisation strategies applied to educational scenarios. Section 3 presents the strategies used for the visual analysis in the case studies. Section 4 describes the two analyzed temporal networks and the performed experiments. Section 5 discusses conclusion and future work.

2. Information Visualisation in Educational Context

Data sets analysis using Information Visualisation inserts the user in the exploration process through graphical and interactive computational approaches (Ware, 2012). By doing so, data understanding and consequently the identification of patterns, anomalies, trends and other behaviours is facilitated, resulting in faster and more reliable decisions (Linhares *et al.*, 2019).

Visualisation techniques have been widely used in recent years to analyze data from various domains, e.g., in healthcare applications (Shneiderman *et al.*, 2013), corporate

environments (van den Elzen *et al.*, 2013; Zhao *et al.*, 2018) and social networks (Linhares *et al.*, 2019). In educational context, Information Visualisation strategies can be used to identify patterns and behaviours involving students and teachers (Ruipérez-Valiente *et al.*, 2015; Caligaris *et al.*, 2015; McNeil, 2015), improving and optimizing the teaching and learning process (Vieira *et al.*, 2018).

In (Chen *et al.*, 2016), a visual analysis system called *DropoutSeer* was developed to evaluate the relationship between online learning activity and student dropout behaviour. Heterogeneous data were extracted from three different types of student activity logs: clickstream, forum posts, and attribution records. A case study was developed and one of the results showed a positive correlation in the performance of the students who have watched the course videos (learning objects).

The authors of (Mu *et al.*, 2019) presented a visual analysis system, named *MOOCad*, that allows exploring anomalies in learning process patterns in open online courses. They propose an approach to interpret individual learning sequences and anomalies in the activities of each student according to temporal patterns. As outcome, it is possible to identify the anomalous groups based on the learning sequence data. In the same way, other researchers (e.g., (Sung *et al.*, 2017; Shin *et al.*, 2018)) have used visual analysis in educational tools to verify the impact of instructional media and to interpret both comments and interactions among students.

Considering the aforementioned efforts, Information Visualisation strategies are employed to assist the identification of patterns involving students in computational environments rather than classroom teaching environments. The visual analysis helps in the interpretation of students and teachers behaviours (nodes in a temporal network). In this context, this paper aims to use visual analysis techniques to study interaction evolution involving students/students and students/teachers and analyze how this information can affect decision making about teaching strategies, especially those related to strategic group formation, and school management.

3. Temporal Network Visualisation

Several Information Visualisation strategies can be employed in the analysis of temporal networks, as, for example, space-time cubes (Bach, 2016), node-link diagrams (also known as structural layouts) (Linhares *et al.*, 2019), circular approaches (v. d. Elzen *et al.*, 2014), and *Massive Sequence View* (MSV) (Holten *et al.*, 2007; van den Elzen *et al.*, 2013). Among these, the MSV layout is the most appropriate when the task is to analyze the interaction distribution throughout time (Linhares *et al.*, 2019).

Fig. 1 presents a visualisation of a small network using the MSV layout. In Fig. 1(a), the network is shown as tabular data, with the first two columns representing the nodes involved in the interactions and with the third column representing the timestamp in which the interactions occurred. Fig. 1(b) illustrates the MSV layout generated from the tabular data. In MSV, the vertical and horizontal axes represent nodes (from top to bottom) and timestamps (from left to right), respectively. For every interaction between two nodes, a vertical line (an edge) is drawn linking them.
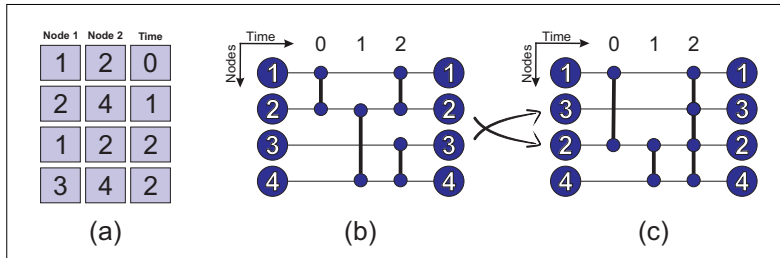
Fig. 1. Representation of a temporal network. (a) Tabular data. (b) MSV layout, in which the x-axis represents the time and the y-axis represents the nodes, respectively. (c) The same visual representation of (b) but with a different node ordering. As it can be seen, the node ordering influences the generated layout.

An important property of MSV layout is node positioning. By comparing the layouts from Fig. 1(b-c), one can see changes in the edge lengths due to differences in node ordering. In real-world networks, node ordering reflects on visual clutter of the layout, and thus in the perception of patterns. There are several node ordering methods described in literature, as, for example, *optimized MSV* (v. d. Elzen *et al.*, 2014), *recurrent neighbors* (Linhares *et al.*, 2017), and *community-based node ordering* (CNO) (Linhares *et al.*, 2019).

CNO is a visual scalable method that uses as base for its ordering strategy the community concept, i.e., groups of nodes that interact more among themselves than to nodes from other groups (Lancichinetti and Fortunato, 2009). As CNO is hierarchical, it subdivides each group into smaller ones at each hierarchical level, as exemplified in Fig. 2(a-c). In educational context, level 1 may contain a community that represents a school class. This class is then decomposed at each new level, so level 2 could have communities representing different groups of students that were gathered to perform a task proposed by their teacher. Finally, CNO allows analyzing only those interactions that occur inside each community of the desired level (by showing only intra-community
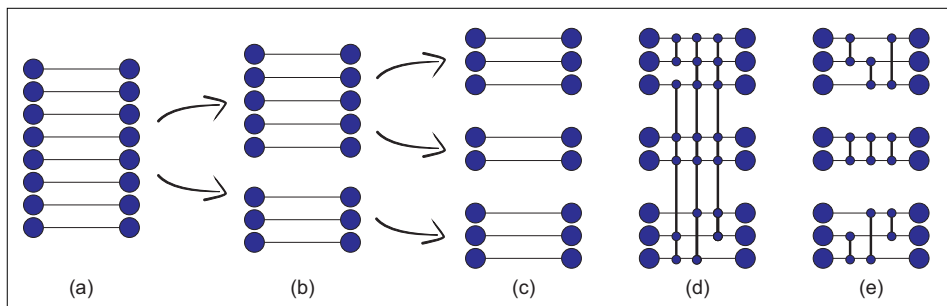

Fig. 2. Different CNO levels applied to the MSV layout. (a-c) The CNO hierarchical approach subdivides the groups at each level; (d-e) CNO allows filtering in only intra-community edges, which facilitates the analysis of specific communities.

edges). Filtering out edges between communities facilitates the comprehension of interaction evolution inside each community (Fig. 2(d-e)).

Network community detection can be seen as a clustering task, highly used in data mining scenarios, but applied to complex networks (Guidotti and Coscia, 2017). In this way, traditional clustering methods, such as DBSCAN (Ester *et al.*, 1996), can be adapted to the community detection task (Gialampoukidis *et al.*, 2016; Linhares *et al.*, 2020). Among the existent community detection algorithms, *Louvain* (Blondel *et al.*, 2008) and *Infomap* (Rosvall and Bergstrom, 2008) represent two of the most recommended approaches due to their performances and low computational complexity (Fortunato and Hric, 2016). As the underlying basis of CNO is community detection, the produced layout is somehow related to hidden patterns that can emerge when visualised in the layout, facilitating identification of visual patterns in the network, accelerating and supporting decision making based on them. In this work, the terms "communities" and "groups" will be used indistinctly from now on.

The CNO ordering process involves three steps (Linhares *et al.*, 2019). The first one is related to community detection, and so any non-overlapping detection algorithm (e.g., *Louvain* or *Infomap*) can be adopted. The other two steps refer to the inter- and intra-community reordering processes. Each of them employs a node ordering method such as *recurrent neighbors* or *optimized MSV*.

4. Case Studies

This section presents two case studies involving real-world education networks from different scenarios, *Primary School* and *High School*. For this purpose, we initially present information related to the two networks and then discuss the analysis using CNO.

4.1. *Education Networks*

The first network, named *Primary School*, contains information about contacts between teachers and students of a primary school in Lyon, France (Stehlé *et al.*, 2011; Gemmetto *et al.*, 2014). The data refer to contacts from Oct. 1st to Oct. 2nd, 2009. The network contains 242 nodes and 125,773 connections distributed in 5,846 timestamps (each timestamp refers to a 20-second interval) and represents contacts from the first to fifth grades, each of them having two classes (A and B). In this network, most contacts involve students in the same class and each class has an assigned teacher. Besides the connections, gender information of the students is also provided. A school day comprises morning and afternoon, with a break for lunch.

The *High School* network contains contacts among students of a high school in Marseille, France (Mastrandrea *et al.*, 2015). The network contains 327 nodes and 188,508 connections distributed in 18,179 timestamps (a 20-second interval each). The data was collected during five days, between Dec. 2nd and 6th, 2013. Nine classes are represented by the network, each of them associated with a type of specialization: PC and PC* are

focused on chemistry and physics; MP, MP*1, and MP*2 on mathematics and physics; 2BIO1, 2BIO2, and 2BIO3 in biology; and PSI on engineering. Besides the connections, gender information is provided along with the information of who is a friend of whom on Facebook.

For both networks, each connection is obtained using RFID sensors (Cattuto *et al.*, 2010). Initially, each person receives a badge with the sensor. Whenever there is a face-to-face approximation between two individuals (between 1 to 1.5 meters), and at least one information exchange between their sensors in a 20-second interval, the contact (connection) is recorded. In this way, the two individuals become nodes in the network and their contact becomes a connection (edge) linking them in the respective timestamp (Cattuto *et al.*, 2010).

4.2. *Visual Analysis Using CNO*

All visual analyses were performed using the software *Dynamic Network Visualisation (DyNetVis)* (Linhares *et al.*, 2017), freely available at `www.dynetvis.com`. The adopted CNO configuration uses *Louvain* as community detection algorithm (CNO step 1) and recurrent neighbors as node ordering strategy (CNO steps 2 and 3). For the analyzed networks, CNO presents a hierarchical approach that allows detecting groups of individuals inside each school class. Ideally, each group of the CNO first level contains all students of a specific class, which is then divided into smaller groups at each subsequent level.

Fig. 3 presents the two days of the *Primary School* network using the MSV layout with hidden edges and two node ordering strategies. The first one is a naive strategy in which nodes are placed in rows uniformly chosen at random (random node ordering, Fig. 3(a)). The second one is CNO level 1 (Fig. 3(b)). To facilitate visual analysis, we associate different colors to different classes. The layout generated by CNO presents at least three patterns not visible in the layout from the random ordering: (i) the size of each class, i.e., the number of students in relation to the total number of students in the school (perceived by the grouping of the same class nodes in the layout); (ii) for some classes in the first day, there are time intervals without any student interaction (e.g., classes 2B, 1A, and 2A) – probably due to sports activities, when contact information was not gathered (Stehlé *et al.*, 2011) (see arrow at class 2B on the first day of Fig. 3(b)); (iii) at the end of the second day, the students from the fourth grade (classes 4A and 4B) left the network earlier than the other students (indicated in Fig. 3(b) by the arrows on the second day). This event may be related to the absence of teachers and the consequent dismissal of the students. From this example, it can be noted how the quality of visual approaches can improve or impair the perception of patterns in the *Primary School*.

Fig. 4 shows the occurrence of similar patterns in the *High School* by presenting two of the five days of the network (days 3 and 4) using CNO level 1. The visual analysis using MSV and CNO facilitates the perception that the students of class PC leave the network before the others on day 3. In the same way, it is easy to note that class PSI* students join the network after the others on day 4. Another pattern involves class 2BIO3 (in red) on day 3: the number of interactions involving this class students is higher in the

Fig. 3. Visual analysis of the *Primary School* network by using different node ordering approaches for the MSV layout. (a) MSV layout with hidden edges and random node ordering; (b) MSV layout with hidden edges and CNO node ordering. Note that an efficient node ordering facilitates pattern identification.

first two-thirds of day 3 (approximately) than in the last one (see the absence of several nodes in the end of the day). When correlated with the factors that motivated these pat-
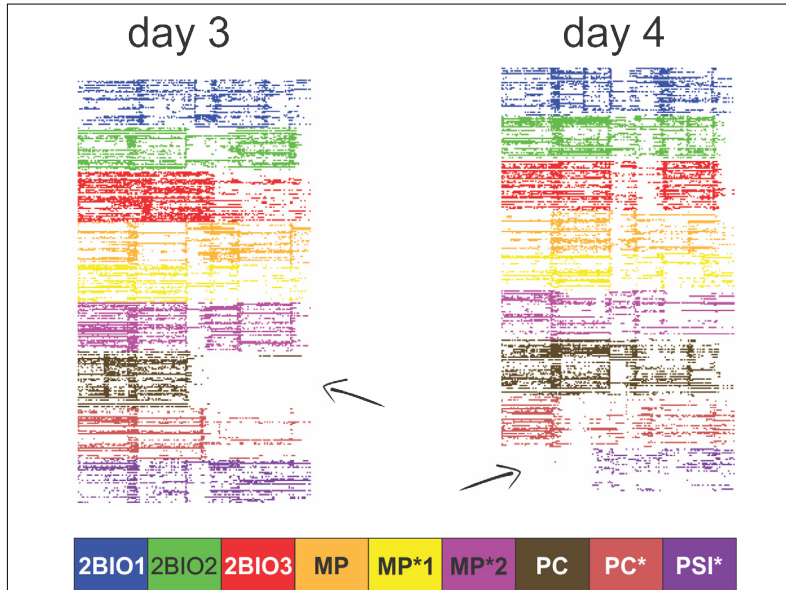
Fig. 4. Layout MSV with hidden edges showing two of the five days of the High School network. The students from class PC left the school earlier than the other students in day 3. In the same way, the students from class PSI* joined the school later than the other students in day 4.

terns, this can represent a useful information to improve both school management decisions and the teaching and learning process.

Another useful resource is related to the user interaction process on the MSV layout in a way that the identification of patterns and trends in the data is facilitated. An important interaction is node selection, which highlights the nodes and edges that will be analyzed by applying an opacity degree on the others. This resource facilitates local analyses in regions of interest.

Fig. 5 illustrates the groups detected on class 5B of the *Primary School* network. The group from level 1 is composed of all students of such class and the respective teacher. At each new level, groups of students with more interaction between themselves are identified. The last level (level 4) presents some small groups, composed of only 2 or 3 individuals. Although Fig. 5 presents only class 5B, this behaviour can be generalized to all other classes of the *Primary School*. This identification is important because the teacher can choose between encourage these formed groups (ideal groups of students) or change them (if the teacher considers these groups as a bad formation – with students that disrupt the class, for instance). Using CNO levels, the teacher can also define the number of students for each group according to a specific need of a learning activity.

Another analysis involves the temporal relation between students and teachers. Fig. 5 shows that one group of CNO level 4 contains only the teacher and a single student. Fig. 6 illustrates how the interactions between this student and the teacher occur in the two days of the *Primary School* network. On the first day, they have several interactions
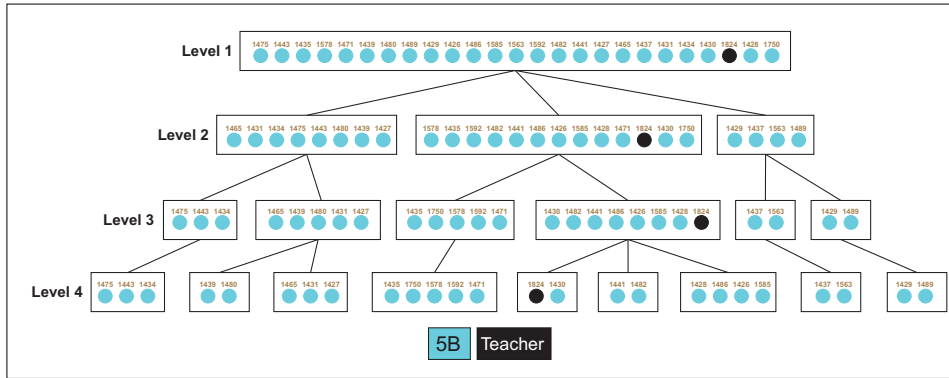
Fig. 5. Hierarchy created by CNO method in class 5B of the *Primary School* network. This hierarchical structure allows choosing groups with different sizes. Moreover, the teacher can choose between encourage or change the CNO group formation in the class.
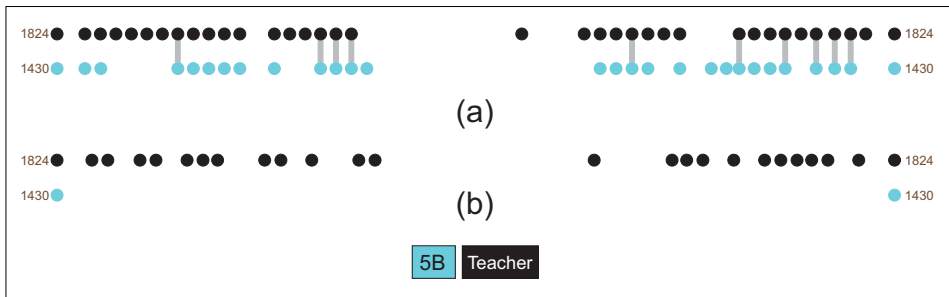

Fig. 6. Group that contains only the teacher of the class 5B and one of this class's student. (a) Several interactions between themselves on the first day of the network; (b) The student is not present in the network on the second day.

between each other (Fig. 6(a)), which may have been motivated by the teacher's perception that this student had more difficulty with the subject/activity. On the second day, this student was not present in the network – which can indicate that the student missed the class this day. Although Fig. 6 shows only the interactions involving the teacher and this particular student, the teacher also interacted with other students.

Besides the analysis involving individuals of the same class, the interactions among individuals of different classes can also be analyzed. Fig. 7 presents third grade students (classes 3A and 3B) and their teachers on the first day of the *Primary School* network (CNO level 1). Fig. 7(a) shows the interactions among all students and teachers. The lunch period (represented by the red rectangle in the figure) represents the period in which there is more interaction involving students from different classes. However, a pattern that is not visible in the layout from Fig. 7(a) can be identified when interacting with the layout by using node selection to highlight the interactions involving class 3B students and class 3A teacher (Fig. 7(b)): the class 3A teacher interacts a lot with the class 3B students in several different moments along the first day of the network,
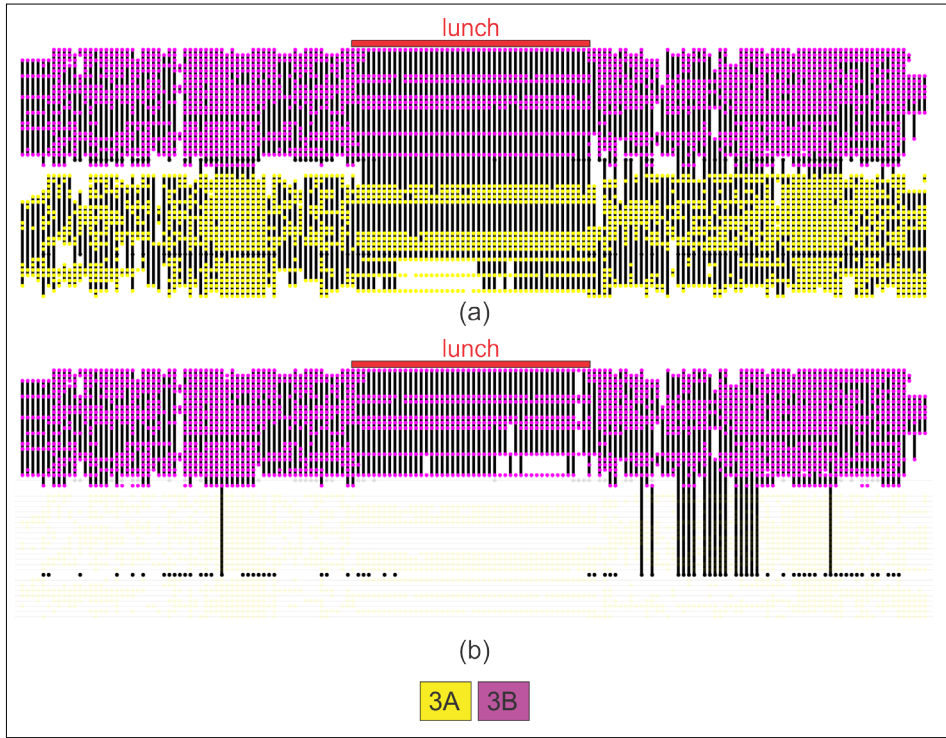
Fig. 7. Visual analysis of the interactions involving students of the third grade (classes 3A and 3B) and their teachers on the first day of the P*rimary School* network. (a) Interactions involving all students and teachers of the third grade; (b) Class 3A teacher has many interactions with the class 3B students, especially after lunch period. The identification of the factors that motivate this behaviour can be useful to support decision making related to the school management.

especially after lunch. The identification of the factors that motivate this behaviour, as, for example, "Where was the class 3B teacher in this period?" or "The absence of class 3B teacher was previously planned or not?" can be useful to support decision making related to school management.

As it could be noted from the previous analyses, it is important to investigate the amount of interactions involving teachers and students during the teaching and learning process. Such analyses can support the development of adequate pedagogical strategies in which, for instance, the teacher identify students with low performance (when relating the groups of students with the individual performances), as well as make it possible to observe whether little attention was given to a student with low performance (when analyzing the interaction frequency between the teacher and this student). Such analyses are also relevant for the school management point of view since it is possible to observe the relationship among teachers and students. It is possible, for example, to check whether a teacher is interacting mainly with specific groups of students, and so adopt administrative decisions that correct or encourage this behaviour.
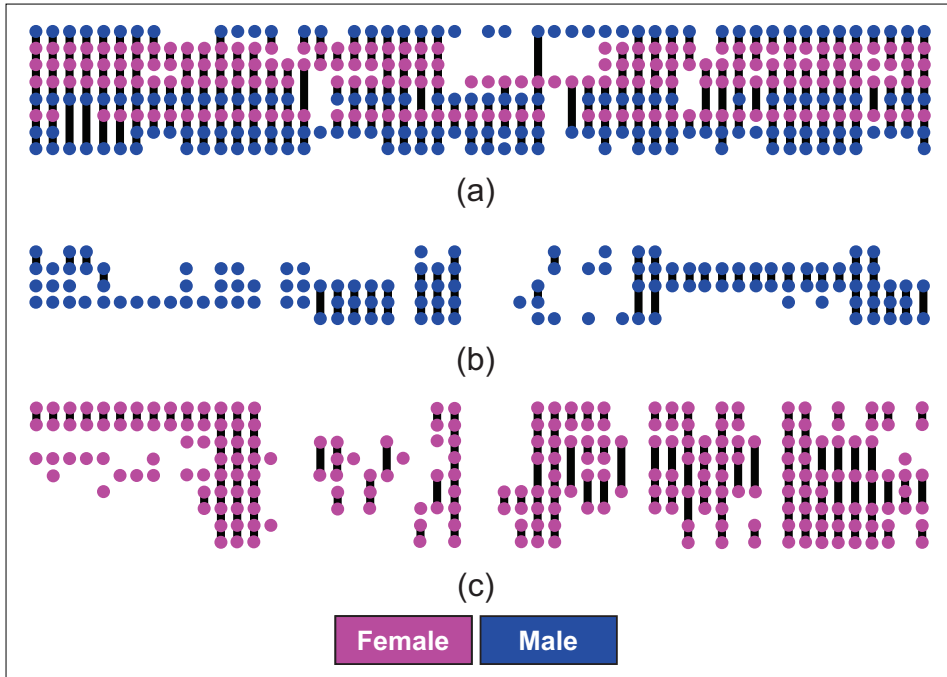
Fig. 8. Groups of students in the *High School* network detected by CNO level 4 and categorized according to gender information. (a) Heterogeneous group, composed by both men and women; (b) Homogeneous group composed only by men; (c) Homogeneous group composed only by women. Gender information may also provide a new perspective to the teacher concerning strategic group formation.

Besides student interaction, there are other information that may influence group formation. An example of complementary information is the students' genders. Fig. 8 presents the three types of groups detected by CNO level 4 for the *High School* network according to this information. There are heterogeneous (mixed-gender) groups (Fig. 8(a)) and homogeneous (single-gender) groups, composed only by men (Fig. 8(b)) or only by women (Fig. 8(c)). Taking into account gender information provides a new perspective to the teacher, as, for example, the prioritization of a specific group formation. Homogeneous groups are recommended when teaching unfamiliar tasks in math and science, while heterogeneous ones can benefit the study of familiar materials (Chennabathni and Rejskind 1998). The teacher can take into account that each group formation has its own characteristics and handle it properly. Examples include:

(i) Balanced-gender heterogeneous groups and homogeneous female groups tend to outperform the others (Zhan *et al.*, 2015; Cen *et al.*, 2014).

(ii) Male students in heterogeneous groups are more confident than those in homogeneous groups (Zhan *et al.*, 2015).

(iii) Female homogeneous groups tend to present a better distribution of member contributions (Cen *et al.*, 2014).

Some individual aspects, such as depressive symptoms and self-esteem, are differentially stratified by gender (Pachucki *et al.*, 2015) and thus should be carefully observed by the teacher as they also affect individual and group performances. As an example, girls with more depressive symptoms present social inhibition while girls with high self-esteem tend to have more interaction partners – in contrast, there is no consistent pattern among boys regarding these two behaviours (Pachucki *et al.*, 2015). On the other hand, dyslexia appears to be more present in boys than girls (Rutter *et al.*, 2004).

Table 1 shows the group distribution according to gender information for both networks using CNO level 4. One can see a high discrepancy between the amount of heterogeneous and homogeneous groups in the *Primary School* network, where 72.2% of groups are heterogeneous. Such proportion is much smaller in the *High School* network: only 42.34% of the groups are heterogeneous (47 of 111 groups). This behaviour may be related to the presence of gender homophily (tendency of having more contacts with same-gender individuals), which tends to increase with age (Stehlé *et al.*, 2013; Pachucki *et al.*, 2015). Moreover, when analyzing different classes in the *High School*, it is possible to identify characteristics that also help to explain this behaviour. On PSI class (focused on engineering), for instance, only 1 of 11 groups is composed only by women, while 5 of them contain only men and 5 are heterogeneous. In this class, there are 10 women and 24 men, which may explain the presence of a single female group. On the other hand, class 2BIO1 (biology) has 28 women and 9 men, which is in agreement with the presence of 8 female groups, 4 heterogeneous ones and only 1 composed only by men.

Following (Stehlé *et al.*, 2013), we analyzed the strength of the ties existent in different groups of students of both *Primary* and *High* schools. *Strong ties* comprehend pairs of students whose contacts have a cumulated duration of at least five minutes over the two days of the *Primary School* (arbitrary threshold also used in (Stehlé *et al.*, 2013) for this same network), and more than 12.5 minutes over the five days of the *High School* (the same proportion as adopted for the *Primary School*). Pairs of students who interact less than these thresholds represent *weak ties* in the networks.

Fig. 9 illustrates three different compositions of heterogeneous groups in both networks perceived with CNO level 4. In all groups that have more than one male (M) stu-

Table 1

Types of groups according to gender information and their distribution over the two networks using CNO level 4. The abbreviations `F', `M' and `U' along with the network names refer to "Females", "Males", and "Unknown" (without gender information), respectively. Groups with at least one U member are considered as heterogeneous

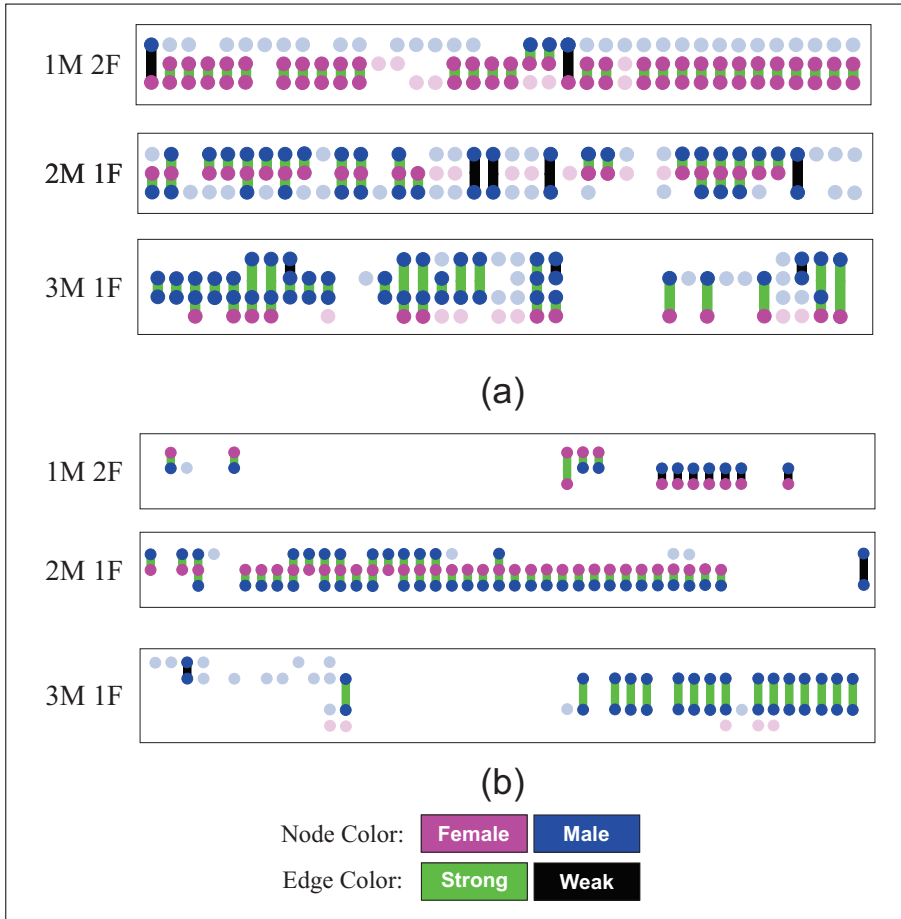| # Groups	Primary Sc. (112 F, 115 M, 15 U)	High Sc. (146 F, 176 M, 7 U)
Only Females	11 (size 2.45 ± 0.65)	28 (size 2.5 ± 0.68)
Only Males	9 (size 3.22 ± 1.13)	36 (size 2.39 ± 0.68)
Heterogeneous	52 (size 3.58 ± 1.02)	47 (size 3.47 ± 1.11)
Overall	72 (size 3.36 ± 1.07)	111 (size 2.87 ± 1.02)

Fig. 9. Different heterogeneous group compositions perceived with CNO level 4 for both networks. (a) *Primary School*. (b) *High School*. The abbreviations `F' and `M' along with the numbers refer to number of "Females" and "Males" in each group.

dent (all groups but 1M2F in both networks), no weak ties involve female (F) students. All weak ties occur between male students probably because female students are better at communication (González-Gómez *et al.*, 2012) and more willing to express their ideas (Zhan *et al.*, 2015; Guntermann and Tovar 1987). The analysis of the groups (*Primary School*, 1M2F and 2M1F – Fig. 9(a)) and (*High School*, 2M1F – Fig. 9(b)) supports this claim: in these three groups, most interactions in the presented time interval involve female students. In all groups of Fig. 9, except (*Primary School*, 1M2F) and one male student from (*High School*, 3M1F), the male students interacted a lot with most of the other group members. This behaviour may be related to the higher confidence that male students tend to have when they are in heterogeneous groups (Zhan *et al.*, 2015). Male students in female-majority groups might experience bad feelings about their positions, impairing their attitudes and performances (Zhan *et al.*, 2015). This aspect may help to
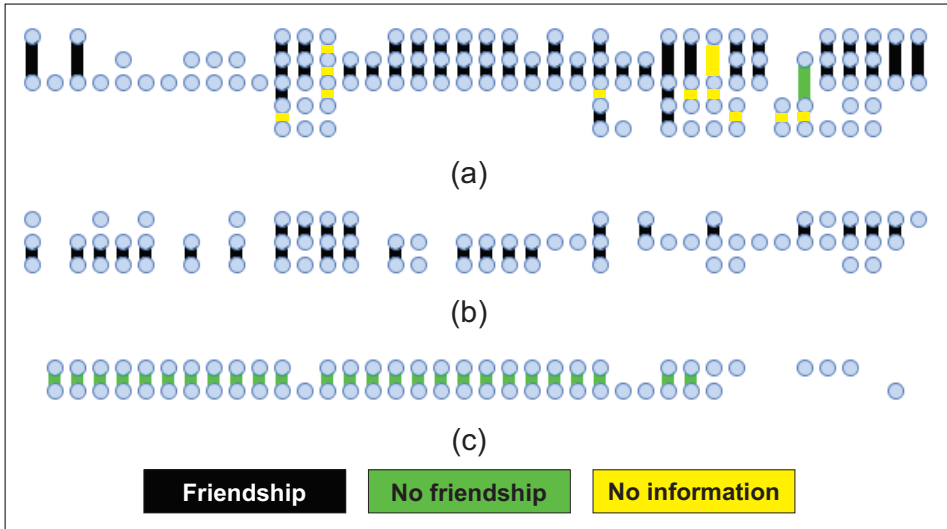
Fig. 10. Groups of students in the *High School* network detected by CNO level 4 and categorized according to the existence or not of friendship relation on Facebook. (a) Heterogeneous group, in which only some students are friends on Facebook; (b) Homogeneous group composed only by friends on Facebook; (c) The unique homogeneous group in the network in which there is not even a single friendship relation on Facebook. The use of complementary information creates a better understanding and represents a useful resource for collaborative learning and knowledge construction.

explain why (*Primary School*, 1M2F) was an exception – the male member interacts with the other group members only a few times in the presented time interval. In fact, male students tend to prefer to be in the majority in a heterogeneous group, while female students prefer to be in the minority (Zhan *et al.*, 2015).

Information related to the existence of interactions among students outside school environment can also influence group formation in school. Recall that *High School* network provides the information about which students have a friendship relation between themselves on the social network Facebook. This information can be associated with groups in class. Fig. 10 presents three types of CNO level 4 detected groups. Similarly to the groups obtained through the gender information, there are heterogeneous groups (Fig. 10(a)), in which only some students are friends on Facebook (for the others there is no friendship and/or no information is provided); and homogeneous groups, composed only by friends on Facebook (Fig. 10(b)) or by students without Facebook friendship (Fig. 10(c)). Heterogeneous groups are expected to be common in schools. However, groups in which there is not even a single friendship relation is not so common. In fact, the group illustrated in Fig. 10(c) is the only occurrence of this type in the *High School* network. Complementary information such as the existence of friendships on social networks – especially in educational networks (e.g., edmodo.com) – creates a better understanding and represents a useful resource for collaborative learning and knowledge construction.

5. Conclusion

The employment of adequate visualisation strategies represents a useful approach to facilitate pattern recognition processes in network data, accelerating and making more reliable strategic decision making. This paper presented two case studies that analyzed two real-world education networks from different scenarios: a primary and a high school. For this purpose, we applied the *Massive Sequence View (MSV)* layout along with the *Community-based Node Ordering (CNO)* node ordering method.

Even without having metadata information about individual or group performances on tests and exams (which would certainly improve our analyses and bring other perspectives), we have shown the importance of visual analyses to assist decision making about teaching strategies and school management. The layouts allowed analyses related to strategic group formation that can take into account different criteria, such as the evaluated day-by-day interaction, friendship on social networks, and gender information. These and other factors may influence the group formation process and thus represent relevant information to support the teacher's decision. In the same way, they provide different data perspectives that could be useful for school management decisions.

As future work, we intend to analyze the school performance of each student in group activity and with groups composed according to the studied types of interactions. Moreover, we can analyze the level of student interaction in groups made by the teacher through visual analysis, and validate the performance and behaviour of the students during the teaching and learning process. In this way, we aim to establish the ideal group formation for a specific class or discipline.

Acknowledgments

References

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):47–97.

Bach, B. (2016). Unfolding Dynamic Networks for Visual Exploration. *IEEE Comp. Graph.*, 36(2):74–82.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech-Theory E.*, 2008(10):P10008.

Caligaris, M., Rodríguez, G., and Laugero, L. (2015). Learning styles and visualization in numerical analysis. *Procedia – Social and Behavioral Sciences*, 174:3696–3701. International Conference on New Horizons in Education, INTE 2014, 25–27 June 2014, Paris, France.

Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J.-F., and Vespignani, A. (2010). Dynamics of person-to-person interactions from distributed rfid sensor networks. *PloS one*, 5(7):e11596.

Cen, L., Ruta, D., Powell, L., and Ng, J. (2014). Does gender matter for collaborative learning? In: *2014 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*, pages 433–440. IEEE.

Chen, Y., Chen, Q., Zhao, M., Boyer, S., Veeramachaneni, K., and Qu, H. (2016). Dropoutseer: Visualizing learning patterns in massive open online courses for dropout reasoning and prediction. In: *2016 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 111–120. IEEE.

Chennabathni, R. and Rejskind, G. (1998). Gender issues in collaborative learning. *Canadian Woman Studies*, 17(4).

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A Densitybased Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, pages 226–231. AAAI Press.

Estrada, E. (2015). Introduction to Complex Networks: Structure and Dynamics. *Lect. Notes Math.*, 2126:93–131.

Fortunato, S. and Hric, D. (2016). Community detection in networks: A user guide. *Phys. Rep.*, 659:1–44.

Gemmetto, V., Barrat, A., and Cattuto, C. (2014). Mitigation of infectious disease at school: Targeted class closure vs school closure. *BMC Infectious Diseases*, 14(1):695.

Gialampoukidis, I., Tsikrika, T., Vrochidis, S., and Kompatsiaris, I. (2016). Community detection in complex networks based on DBSCAN* and a Martingale process. In: *2016 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, pages 1–6.

González-Gómez, F., Guardiola, J., Rodríguez, _ O. M., and Alonso, M. _A. M. (2012). Gender differences in e-learning satisfaction. *Computers & Education*, 58(1):283–290.

Guidotti, R. and Coscia, M. (2017). On the equivalence between community discovery and clustering. In: *International Conference on Smart Objects and Technologies for Social Good*, pages 342–352. Springer.

Guntermann, E. and Tovar, M. (1987). Collaborative problem-solving with logo: Effects of group size and group composition. *Journal of Educational Computing Research*, 3(3):313–334.

Holme, P. and Saramäki, J. (2012). Temporal networks. *Physics Reports*, 519(3):97–125.

Holten, D., Cornelissen, B., and van Wijk, J. J. (2007). Trace Visualization Using Hierarchical Edge Bundles and Massive Sequence Views. In: *2007 4th IEEE International Workshop on Visualizing Software for Understanding and Analysis*, pages 47–54.

Lancichinetti, A. and Fortunato, S. (2009). Community detection algorithms: A comparative analysis. *Phys. Rev. E*, 80.

Linhares, C. D., Ponciano, J. R., Pereira, F. S., Rocha, L. E., Paiva, J. G. S., and Traven_colo, B. A. (2020). Visual analysis for evaluation of community detection algorithms. *Multimedia Tools and Applications*, pages 1–23.

Linhares, C. D. G., Ponciano, J. R., Pereira, F. S. F., Rocha, L. E. C., Paiva, J. G. S., and Travençolo, B. A. N. (2019). A scalable node ordering strategy based on community structure for enhanced temporal network visualization. *Computers & Graphics*, 84:185–198.

Linhares, C. D. G., Travençolo, B. A. N., Paiva, J. G. S., and Rocha, L. E. C. (2017). Dynetvis: A system for visualization of dynamic networks. In: *Proceedings of the Symposium on Applied Computing*, SAC '17, pages 187–194.

Mastrandrea, R., Fournet, J., and Barrat, A. (2015). Contact patterns in a high school: A comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLOS ONE*, 10(9):1–26.

McNeil, S. (2015). Visualizing mental models: understanding cognitive change to support teaching and learning of multimedia design and development. *Educational Technology Research and Development*, 63(1):73–96.

Mu, X., Xu, K., Chen, Q., Du, F., Wang, Y., and Qu, H. (2019). Moocad: Visual analysis of anomalous learning activities in massive open online courses.

Pachucki, M. C., Ozer, E. J., Barrat, A., and Cattuto, C. (2015). Mental health and social networks in early ado-

lescence: A dynamic study of objectively-measured social interaction behaviors. *Social Science & Medicine*, 125:40–50. Special Issue: Social Networks, Health and Mental Health.

Rosvall, M. and Bergstrom, C. T. (2008). Maps of random walks on complex networks reveal community structure. *P. Natl. A. Sci.*, 105(4):1118–1123.

Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., and Kloos, C. D. (2015). Alas-ka: A learning analytics extension for better understanding the learning process in the khan academy platform. *Computers in Human Behavior*, 47:139–148. Learning Analytics, Educational Data Mining and data-driven Educational Decision Making.

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., Moffitt, T. E., Meltzer, H., and Carroll, J. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies. *Jama*, 291(16):2007–2012.

Shin, H., Ko, E.-Y., Williams, J. J., and Kim, J. (2018). Understanding the effect of in-video prompting on learners and instructors. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 319. ACM.

Shneiderman, B., Plaisant, C., and Hesse, B. W. (2013). Improving healthcare with interactive visualization. *Computer*, 46(5):58–66.

Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J., Quaggiotto, M., Van den Broeck, W., Régis, C., Lina, B., and Vanhems, P. (2011). High-resolution measurements of face-to-face contact patterns in a primary school. *PLOS ONE*, 6(8):e23176.

Stehlé, J., Charbonnier, F., Picard, T., Cattuto, C., and Barrat, A. (2013). Gender homophily from spatial behavior in a primary school: A sociometric study. *Social Networks*, 35(4):604–613.

Sung, C.-Y., Huang, X.-Y., Shen, Y., Cherng, F.-Y., Lin, W.-C., and Wang, H.-C. (2017). Exploring online learners' interactive dynamics by visually analyzing their time-anchored comments. In: *Computer Graphics Forum*, volume 36, pages 145–155. Wiley Online Library.

v. d. Elzen, S., Holten, D., Blaas, J., and van Wijk, J. J. (2014). Dynamic network visualization with extended massive sequence views. *IEEE Transactions on Visualization and Computer Graphics*, 20(8):1087–1099.

van den Elzen, S., Holten, D., Blaas, J., and van Wijk, J. J. (2013). Reordering massive sequence views: Enabling temporal and structural analysis of dynamic networks. In *2013 IEEE Pacific Visualization Symposium (PacificVis)*, pages 33–40.

Vieira, C., Parsons, P., and Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education*, 122:119–135.

Ware, C. (2012). *Information Visualization: Perception for Design*. Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann, San Francisco, CA, USA, 3 edition.

Zhan, Z., Fong, P. S., Mei, H., and Liang, T. (2015). E_ects of gender grouping on students' group performance, individual achievements and attitudes in computer-supported collaborative learning. *Comput. Hum. Behav.*, 48(C):587–596.

Zhao, Y., She, Y., Chen, W., Lu, Y., Xia, J., Chen, W., Liu, J., and Zhou, F. (2018). Eod edge sampling for visualizing dynamic network via massive sequence view. *IEEE Access*, 6:53006–53018.

J. R. Ponciano received the BSc and MSc degrees in computer science from the Federal University of Uberlândia, Uberlândia /MG, Brazil. He is currently a PhD student at the same university. His research interests include Information Visualization, especially complex network visualization and visualization of streaming data.

C. D. G. Linhares received the BSc, MSc and PhD degrees in computer science from the Federal University of Uberlândia, Uberlândia /MG, Brazil. His research interests include Information Visualization, especially complex network visualization and visual scalability, Human Computer Interaction and Dynamic processes.

S. L. Melo graduated in Information Systems (2006) and has a MSc degree in Computer Science (2016). She is currently a PhD student in Electrical Engineering at the Federal University of Uberlândia – Brazil. Her research interests include Intelligent Tutoring Systems, Cognitive Science, Affective Computing, and Artificial Intelligence applied to education.

L. V. Lima received a BSc degree in Electrical Engineering (1983), MSc degree in Electrical Engineering (1988), and a PhD degree in Electrical Engineering (1998). He is currently a full professor at the Faculty of Electrical Engineering, Federal University of Uberlândia, Uberlândia, Brazil. His research interests include Artificial Intelligence, Neuro Science, Computer Music, and Adaptative Systems.

B. A. N. Travençolo received a BSc degree in Computer Science (2003) and a PhD degree in Bioinformatics (2007), all from University of São Paulo. He is currently associate professor at the Faculty of Computing, Federal University of Uberlândia, Uberlândia, Brazil. His research interests include Image Processing, Information Visualization, Scientific Visualization and Complex Networks.