Analysis of Discussion Forum Interactions for Different Teaching Modalities based on Temporal Social Networks

Nidia Guadalupe López-Flores[1], María Óskarsdóttir[1] and Anna Sigriður Islind[1]

[1]Department of Computer Science, Reykjavik University, Menntavegur 1, 101 Reykjavik, Iceland

Abstract

The temporal component has been pointed out as an impactful dimension for educational research, as learning is not a static phenomena. In this paper, we investigate how the students' activity and interaction dynamics in an online forum differ among teaching modalities. The analysis is based on three years of forum interaction data in an undergraduate course where the only significant change was the teaching modality, from on-site (2019) to fully online (2021). We build and analyse temporal networks by studying changes in several networks' measures and features. Our preliminary results show changes in the students' and teachers' interaction dynamics as well as changes in the posts' content with fully online teaching. This on-going research is focused on studying the impact of teaching modalities on the discussion forum interactions. Initial results are promising, and other features such as the students' closeness and betweenness, as well as the dynamics changes' relationship with the academic performance are being explored to improve our understanding on teaching modalities and help generalisation of Network Science and Learning Analytics research in wider educational contexts.

Keywords

Temporal networks, Teaching modality, Online forums

1. Introduction

The education provision has over the course of the last two years, since spring 2020, adapted to major changes and accelerated technology adoption. The impact of digitally enhanced learning and teaching, for instance, blended or online formats, has called attention to revision of teaching methods and modalities [1]. Research on teaching modalities focusing on education components, for example, achievement of learning outcomes, and analysis of study patterns, has highlighted the importance of investigating the social element in online teaching settings [2, 3].

Network analysis is one of the research approaches in Learning Analytics (LA) for analysis and modelling of relational data in educational settings [4]. Social relationships are an essential component of learning processes, as they enhance collaboration, skills acquisition, and provide the learner with social, psychological and academic benefits [5]. Similarly to other real networks, social networks representing students' interactions are not static as their connections and participants change over time. For example, when the students start a course or programme,

Proceedings of the NetSciLA22 workshop, March 22, 2022

nidia20@ru.is (N. G. López-Flores); mariaoskars@ru.is (M. Óskarsdóttir); islind@ru.is (A. S. Islind)
0000-0001-8074-4488 (N. G. López-Flores); 0000-0001-5095-5356 (M. Óskarsdóttir); 0000-0002-4563-0001 (A. S. Islind)

they might not know all their peers; at first, they could start interacting with a small group of friends and, as the programme progresses, they could get in contact with other students, because of academic or personal preferences. Similarly, at the end of the course, some connections would have vanished if the students stopped interacting at some point during the programme.

This on-going research focuses on exploring and describing the differences in the usage, content and interaction dynamics in a discussion forum platform commonly used in undergraduate courses while taking into account different teaching modalities. We aim to answer the research question: To what extent does the change in teaching modality impact the usage and interaction patterns in discussion forum platforms? To answer the question, three years of forum data were gathered and analysed to outline the changes in interaction dynamics when the modality of teaching shifted from on-site in 2019, transitioning to emergency remote teaching [6] in the middle of the term in 2020, to the *current* new normality with online teaching in spring 2021. The data were analysed using a Social Network approach, by creating three temporal networks, one per year. All the temporal networks consist of twelve snapshots, each of them representing the weekly interactions among the students taking part in the forum threads. The temporal networks were used to compare the changes in the dynamics and evolution of the group interactions as the term progressed. In addition, the changes in the interaction patterns and posts' content were addressed by comparing several indicators, such as posts' frequency, length and complexity, week of posting, length of answers' waiting time, and posts' answerer [1].

The rest of the paper is organised as follows. In Section 2, we present related research on Social Network Analysis, discussion forums, and temporal networks. In Section 3, we present the data and the method followed in the study. Finally, in Section 4 we present the preliminary results obtained, and in Section 5 the discussion and next steps are outlined.

2. Related work

2.1. Discussion forum interactions and Social Network Analysis

Networks are created based on existing relationships among two or more entities, a social network refers to a collection of people sharing connections within an environment [7]. Social Network Analysis (SNA) has been used to investigate several aspects of education; in most cases the students are represented as nodes in the network, and the edges connecting them represent different kinds of relationships or communication events, either online or face-to-face [4]. Among the educational aspects investigated using SNA, the most common are academic success and dropout [8, 9], the influence of homophily on performance [10, 11], Massive Open Online Courses (MOOCs) [12], study patterns [13], course selection [14, 15], collaborative learning [16], and community detection [14, 17, 18]. Data sources commonly used to construct the networks include self-reports and surveys in face-to-face settings, as well as discussion logs and threads in online events and forums for online networks. In regards of forum interactions, Poquet [19] investigated the effect of enrolment, course design, and individual learner characteristics in performance-base similarity in online discussion forums in higher education. Xu et al. [20] applied SNA to analyse the effect of the content of forum posts on the connections among

[1] In describing this data, an answerer is any participant of the course that provides answers to the forum threads.

students participating in MOOCs. Gitinabard et al. [21] used online discussion forum data to analyse the correlation among network metrics and student performance. Their findings show that students who asked more questions and received more feedback got higher grades. Similarly, Williams-Dobosz et al. [22] applied SNA to examine the connection among learning performance and help seeking. Their studies also concluded that connectivity per se does not have a direct influence on learning outcomes achievement.

2.2. Temporal networks

Temporal networks, (i.e., networks where the nodes' connections change over time) are useful for modelling the dynamic behaviour of many complex real-world systems because their interactions are rarely constant over time [23, 24]. Considering the moment when the connections happen when creating networks has an impact on several properties and measures, such as connectedness, shortest paths, and centrality [25]. Evolution over time is an important dimension in learning processes. Consequently, in the latest years, the attention paid to the temporality and its effects on learning has increased; focusing on varied elements of educational settings [26, 27]. Saqr et al. [28] studied the role of temporal measures for predicting academic performance; their findings underline that the temporal dimension provides essential information on learning patterns and can potentially support instructors in addressing students' performance. Xu et al. [12] implemented SNA and community detection algorithms over data from an 8-week MOOCs course to study the moment when most connections among the students enrolled happen. In their study, most of the connections and communities were created during the first two weeks of the course and evidence of performance homophily was found between the students and their closest friends at the end of the course. Vörös et al. [29] reported on the collection methodology of a longitudinal data set of undergraduate students, collected from 2016 to 2019. In their study, the undergraduate students answered a set of short and long surveys related to social connections, individual background and study behaviour during the three years of their undergraduate studies. Data from social media platforms and two field experiments were also gathered to complement the surveys' answers. Shirvani et al. [30] based their research on online discussion forums in two MOOCs offered by Coursera to analyse the content and social structure dynamics and temporal patterns. Their research shows that activity levels can be predicted one week in advance using the information in the temporal networks.

3. Methods

3.1. Data

This study encompasses the discussion forum interaction data of one undergraduate course from 2019 to 2021. The course is a first-year course for the computer science programme at Reykjavik University and a second year course for the engineering programme. The course is always taught in the spring term, with a high number of students enrolled each year: 215, 233 and 318 students respectively. The main characteristic of this case study is that during the three years included, most of the course's features remained without changes. The course had the same teacher, syllabus, book, assessment structure, and an active online discussion forum available for the students' use. The only component with significant changes was the teaching modality. In 2019, the teaching was delivered on-site at the university premises. In 2020 it started on-site, but due to the meeting restrictions imposed as response to the COVID-19 pandemic outbreak, the teaching modality was suddenly moved to online learning at the beginning of March. Finally, in 2021, the course was planned and delivered completely online.

The data were gathered from Piazza, an online discussion forum platform widely used among undergraduate courses to support and enhance student-student and student-instructor communication. In this course, the students were allowed to post questions, notes, and polls to the forum participants registered in the course section in Piazza. In addition, they were able to post answers, updates and follow-up questions to any public post previously published. The discussion forum was included into the course settings with the intention to: (1) encourage the students to collaborate with their classmates, and (2) provide the students with a direct communication channel with the instructors (teacher and teaching assistants). Neither their registration into the course section nor their participation in the threads were mandatory or included in the assessment structure. In addition, categorical grade placements from A to D were created based on the numerical grades obtained from the Learning Management System. Students with outstanding performance were allocated in A grade placement, whereas students who did not pass the course were allocated in D grade placement.

3.2. Network construction and analysis

To answer our research question, the analysis performed includes two components. The first component, the social network analysis is included with the objective of studying the evolution of communication events along the term and compare it across years and teaching modalities. The second component, related to posts' content features, complements the social networks analysis and is useful to get a better understanding of the differences between teaching modalities. Table 1 displays the features included in the posts' content analysis.

For the network analysis, the networks were created based on the edge list created using the forum threads archive, as described in Figure 1. Each node in the network represents a forum participant (teacher, teaching assistant, or student). The categorical grade placements were included in the network as an attribute of each student node.

As noted above, the interaction network in the discussion forum is dynamic, since not all participants registered are present since the beginning, and more edges are added as the course progresses. Representing dynamic interactions as networks can be addressed in several ways [31]. In this study, following the syllabus structure and the asynchronous dynamic on the online discussion forum; we decided to create twelve static networks per year. Each network represents one week of interactions between the participants (See Fig. 2). The networks were created based on the post's timestamp to analyse their growth and their features' evolution during the term; and compare them among teaching modalities to study their differences. Table 2 displays the network measures that were calculated and analysed for all years and networks created. The data pre-processing and social network analysis were conducted in Python using the NetworkX package.

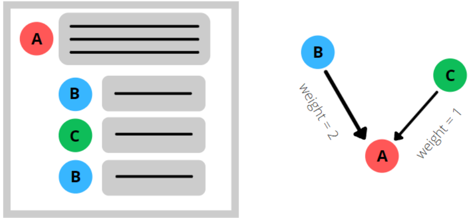


Figure 1: Network creation based on the Piazza threads. Left: Forum discussion structure. Any new thread starts with a question, note or poll published by students or instructors. Any participant can post an answer, follow-up question, or updates. Right: Directed edges in the network are created from participants who posted answers, follow-up questions, or updates (Nodes B and C) to questions, notes or polls posted by any other participant (Node A). The weight of each link is calculated by accumulating the number of times one node (source) answered to another (target).

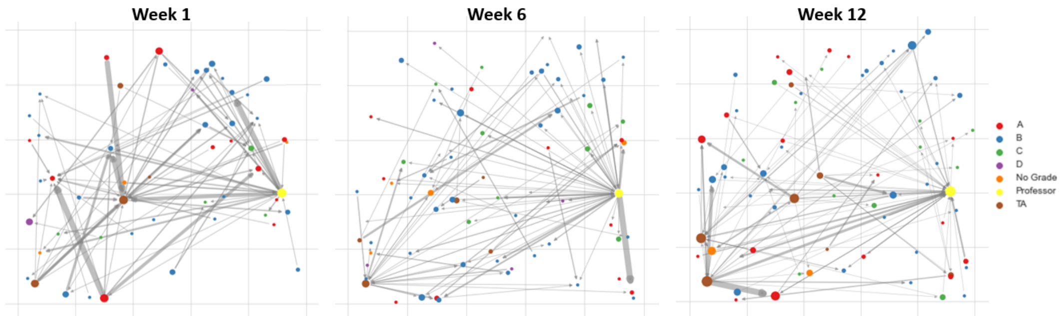


Figure 2: Temporal networks created based on 2021 posts in weeks 1, 6, and 12. The nodes are coloured based on categorical grade placement obtained: A (red), B (blue), C (green), D (purple), No grade (orange). The teacher and teaching assistants are yellow and brown respectively. The nodes are sized according to the number of contributions posted each week.

4. Preliminary results

Figure 3 displays the weekly evolution of the network measures presented in Table 2. The exploratory analysis on the networks' structure reveals differences in the students' and teachers' activity among the modalities of teaching. Those results, combined with the analysis of the posts' features and the platform usage among years, lead to the first insights about the effect of the change in the teaching modality between years. These insights can be split into students, instructors and comments' structure conclusions.

For the students:

1. With fully online teaching, a higher proportion of students registered in the forum platform.
2. The percentage of active students increased from 41% and 52%, to 66% in 2021.
3. Active students had on average grades 35% higher than non-registered and observer students in online teaching, whereas with on-site teaching the increment was 5%, and 15% in the transition in 2020.

Table 1
Set of features analysed in the posts' content component.

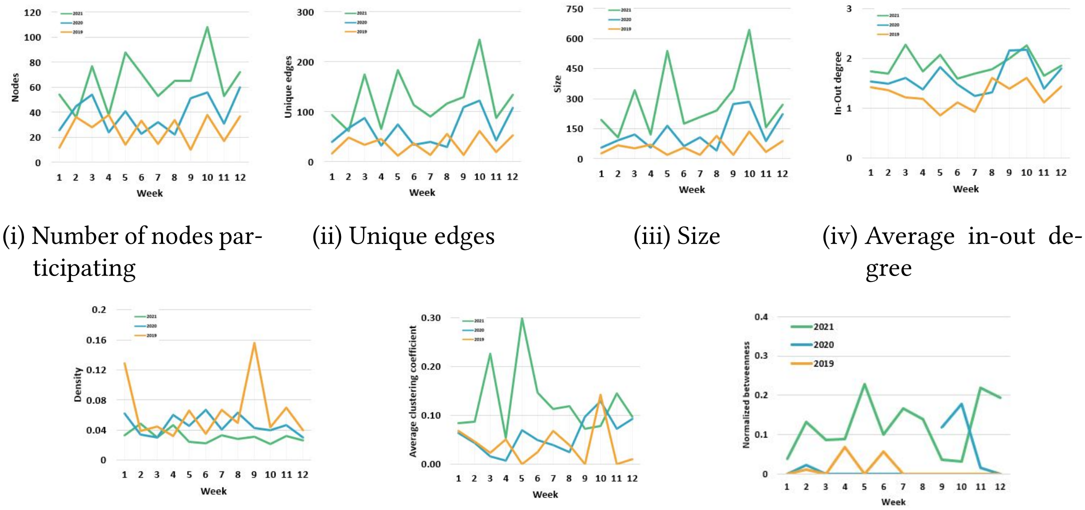
Feature
Length of the first post in each thread
Number of unique words within the post
Whether the post was published anonymously or not
Time in weeks until a student's first post or contribution
Time in hours until the first answer received to each post
Follow-up questions in each thread
Number of answers provided by the teacher, teaching assistants, and students

Table 2
List of network measures computed and analysed in the first component and their meaning in the context of the discussion forum network.

Measure	Meaning in the discussion forum's network context
No. of nodes	Number of students taking part in the forum threads by posting or answering threads.
No. of edges	Number of unique edges as an indicator of how active the participants are by contacting different participants each week.
Size	Weight assigned to the edges as an indicator of how strong the connections are. The more communication events occur between each pair of participants, the higher the size will be.
In-degree	Number of peers answering a student's post as an indicator of the nodes' posts popularity.
Out-degree	Number of peers the student answered to by posting an answer or follow-up question.
Density	Number of posts relative to the number of possible students' connections as an indicator of the network's completeness.
Clustering coefficient	Proportion of a nodes' friends answering to each other, as an indicator of how tightly connected the students are.
Teacher's betweenness	Teacher's influence on the information spreading among the course participants.

4. The networks of fully online teaching grew faster than with the other modalities, as a result of a higher percentage of students participating in the forum each week; the number of unique connections and contributions was also higher.
5. In 2021, with a higher average in-degree, out-degree, and clustering coefficient; the students got in contact with more people and were more connected with their close peers.
6. In fully online teaching, there was no difference in the week when the students started posting questions on the platform, in contrast to previous years, when D students started posting later in the term.
7. The students were more involved in the discussion threads, the number of follow-up questions posted increased in 2021.

In regards to the instructors' activity, in 2021 with fully online teaching:



(i) Number of nodes participating

(ii) Unique edges

(iii) Size

(iv) Average in-out degree

(v) Density

(vi) Average clustering coefficient

(vii) Teacher's betweenness

Figure 3: Weekly evolution of the temporal network' measures displayed on Table 2 from 2019 to 2021.

1. The teacher's betweenness indicates the influence on the spreading of information among the students was stronger, highlighting the importance of the teacher's presence when the teaching modality was fully online (Fig. 3vii).
2. The number of questions increased by 75% and 51% compared to the questions posted in 2019 and 2020 respectively.
3. The instructors responded to a higher percentage of the questions.
4. The answers were provided faster, the median waiting time with fully online teaching was 56 minutes, whereas with on-site teaching the median time was calculated at 4.2 hours.

These results help to dimension the work overload experienced by teachers and teaching assistants since the beginning of the pandemic.

Regarding the comments' structure and usage analysis:

1. The posts were longer and more complex than those posted in previous years.
2. The contributions posted anonymously decreased with fully online teaching. Possible explanations include; the students' want to be identified by their instructors and peers due to the lack of direct contact; as the students never met in person, neither the teacher nor their peers, they could feel less exposed participating in the forum.

5. Discussion and proposed analysis

This paper presents the initial conclusions of an exploratory analysis performed with three years of discussion forum data. Previous research focusing on discussion forum interactions and their evolution usually relies on a specific course or teaching modality. In contrast, this paper combines the descriptive analysis of the networks' structure with the posts' features to compare the impact of the teaching modality choice on the usage of the forum, and over the participants' interactions. This approach allowed us to notice that the teaching modality impacted not only the activity levels, but also the way the students' and instructors' connected with other participants, as well the comments features (e.g. length and anonymity status).

Initial insights from the descriptive analysis of the temporal networks created and their measures indicate that this approach is appropriate for researching the differences in the interactions and usage over time between teaching modalities. However, one limitation of this research is the learning context of the discussion forum in this study; which might prevent our conclusions from being extended to courses with mandatory or graded forum participation.

The subsequent work of this study includes the exploration of temporal network measures for students and their differences across teaching modalities. Furthermore, the relationship between the temporal network measures, their evolution, and the academic performance should be explored in more detail to develop models that account for teaching modality and help to identify students at risk of failing or drop out.

Both the initial and final results obtained from this research will contribute to the LA field by enhancing our understanding on the effect of teaching modality choices in higher education. In addition, this research will help to advance in the still limited research on temporal networks in educational contexts.

References

[1] M. Gaebel, T. Zhang, H. Stoeber, A. Morrisroe, Digitally enhanced learning and teaching in European higher education institutions, Technical Report, European University Association absl., 2021.

[2] J. C. Bahamón, A. Rorrer, Improving student learning outcomes in online courses: An investigation into the effects of multiple teaching modalities, in: Proceedings of the 51st ACM Technical Symposium on Computer Science Education, SIGCSE '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 1179–1185. URL: https://doi.org/10.1145/3328778.3366880. doi:10.1145/3328778.3366880.

[3] M. Lewis, Z. Deng, S. Krause-Levy, A. Salguero, W. G. Griswold, L. Porter, C. Alvarado, Exploring student experiences in early computing courses during emergency remote teaching, in: Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1, ITiCSE '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 88–94. URL: https://doi.org/10.1145/3430665.3456315. doi:10.1145/3430665.3456315.

[4] O. Poquet, M. Saqr, B. Chen, Recommendations for network research in learning analytics: To open a conversation, 2021.

[5] R. Huang, J. M. Spector, J. Yang, Educational Technology, Springer, 2019.

[6] C. Hodges, S. Moore, B. Lockee, T. Trust, A. Bund, The difference between emergency remote teaching and online learning, 2020. URL: https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning.

[7] L. Igual, S. Seguí, Introduction to Data Science, Springer International Publishing, 2017. doi:10.1007/978-3-319-50017-1.

[8] D. Blansky, C. Kavanaugh, C. Boothroyd, B. Benson, J. Gallagher, J. Endress, H. Sayama, Spread of academic success in a high school social network, PLoS ONE 8 (2013). doi:10.1371/journal.pone.0055944.

[9] N. Gitinabard, F. Khoshnevisan, C. F. Lynch, E. Y. Wang, Your actions or your associates? predicting certification and dropout in moocs with behavioral and social features, in: International Educational Data Mining Society, 2018.

[10] B. Rienties, D. Tempelaar, Turning groups inside out: A social network perspective, Journal of the Learning Sciences 27 (2018). doi:10.1080/10508406.2017.1398652.

[11] Q. Nguyen, O. Poquet, C. Brooks, W. Li, Exploring homophily in demographics and academic performance using spatial-temporal student networks, in: A. N. Rafferty, J. Whitehill, C. Romero, V. Cavalli-Sforza (Eds.), Proceedings of the 13th International Conference on Educational Data Mining, EDM 2020, Fully virtual conference, July 10-13, 2020, International Educational Data Mining Society, 2020. URL: https://educationaldatamining.org/files/conferences/EDM2020/papers/paper_170.pdf.

[12] Y. Xu, C. Lynch, T. Barnes, How many friends can you make in a week?: evolving social relationships in moocs over time, in: EDM, 2018.

[13] S. Y. Lee, H. S. Chae, G. Natriello, Identifying user engagement patterns in an online video discussion platform, in: EDM, 2018.

[14] E. G. Sturludóttir, E. Arnardóttir, G. Hjálmtýsson, M. Óskarsdóttir, Gaining insights on student course selection in higher education with community detection, 2021. arXiv:2105.01589.

[15] G. M. Weiss, N. Nguyen, K. Dominguez, D. D. Leeds, Identifying hubs in undergraduate course networks based on scaled co-enrollments: Extended version, 2021. arXiv:2104.14500.

[16] L. Yan, R. Martinez-Maldonado, B. G. Cordoba, J. Deppeler, D. Corrigan, G. F. Nieto, D. Gasevic, Footprints at school: Modelling in-class social dynamics from students' physical positioning traces, in: LAK21: 11th International Learning Analytics and Knowledge Conference, LAK21, Association for Computing Machinery, New York, NY, USA, 2021, p. 43–54. URL: https://doi.org/10.1145/3448139.3448144. doi:10.1145/3448139.3448144.

[17] S. Yassine, S. Kadry, M.-A. Sicilia, Application of community detection algorithms on learning networks. the case of khan academy repository, Computer Applications in Engineering Education 29 (2020) 411–424. doi:10.1002/cae.22212.

[18] N. López Flores, A. S. Islind, M. Oskarsdottir, Exploring study profiles of computer science students with social network analysis, in: The 55th Hawaii International Conference on System Sciences (HICSS), 2022. doi:10.24251/HICSS.2022.214.

[19] O. Poquet, Why Birds of a Feather Flock Together: Factors Triaging Students in Online Forums, Association for Computing Machinery, New York, NY, USA, 2021, p. 469–474. URL: https://doi.org/10.1145/3448139.3448185.

[20] Y. Xu, N. Gitinabard, C. F. Lynch, T. Barnes, What You Say is Relevant to How You Make Friends: Measuring the Effect of Content on Social Connection, in: Proceedings of the 12th International Conference on Educational Data Mining, International Educational Data Mining Society, 2019.

[21] N. Gitinabard, L. Xue, C. Lynch, S. Heckman, T. Barnes, A social network analysis on blended courses, 2017, pp. 22–26.

[22] D. Williams-Dobosz, R. F. L. a. Azevedo, A. Jeng, V. Thakkar, S. Bhat, N. Bosch, M. Perry, A social network analysis of online engagement for college students traditionally underrepresented in stem, in: LAK21: 11th International Learning Analytics and Knowledge Conference, LAK21, Association for Computing Machinery, New York, NY, USA, 2021, p. 207–215. URL: https://doi.org/10.1145/3448139.3448159. doi:10.1145/3448139.3448159.

[23] P. Holme, J. Saramäki, Temporal networks, Physics Reports 519 (2012) 97–125. URL: https://www.sciencedirect.com/science/article/pii/S0370157312000841. doi:https://doi.org/10.1016/j.physrep.2012.03.001, temporal Networks.

[24] V. Nicosia, J. Tang, C. Mascolo, M. Musolesi, G. Russo, V. Latora, Graph Metrics for Temporal Networks, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 15–40. URL: https://doi.org/10.1007/978-3-642-36461-7_2. doi:10.1007/978-3-642-36461-7_2.

[25] V. Nicosia, J. Tang, M. Musolesi, G. Russo, C. Mascolo, V. Latora, Components in time-varying graphs, Chaos: An Interdisciplinary Journal of Nonlinear Science 22 (2012) 023101. URL: https://doi.org/10.1063/1.3697996. doi:10.1063/1.3697996. arXiv:https://doi.org/10.1063/1.3697996.

[26] B. Chen, O. Poquet, Networks in learning analytics: Where theory, methodology, and practice intersect, Journal of Learning Analytics 9 (2022) 1–12. URL: https://learning-analytics.info/index.php/JLA/article/view/7697. doi:10.18608/jla.2022.7697.

[27] M. Saqr, O. Viberg, J. Nouri, S. Oyelere, Multimodal temporal network analysis to improve learner support and teaching, 2020.

[28] M. Saqr, J. Nouri, U. Fors, Time to focus on the temporal dimension of learning. a learning analytics study of the temporal patterns of students' interactions and self-regulation, International Journal of Technology Enhanced Learning 11 (2019). doi:10.1504/IJTEL.2019.10020597.

[29] A. Vörös, Z. Boda, T. Elmer, M. Hoffman, K. Mepham, I. J. Raabe, C. Stadtfeld, The swiss studentlife study: Investigating the emergence of an undergraduate community through dynamic, multidimensional social network data, Social Networks 65 (2021) 71–84. doi:10.1016/j.socnet.2020.11.006.

[30] M. Shirvani Boroujeni, P. Dillenbourg, Discovery and temporal analysis of mooc study patterns, Journal of Learning Analytics 6 (2019) 16–33. doi:10.18608/jla.2019.61.2.

[31] N. Masuda, R. Lambiotte, A Guide to Temporal Networks, volume 06, WORLD SCIENTIFIC (EUROPE), 2020. doi:10.1142/q0268.