



First-Year Engineering Students' Strategies for Taking Exams

Jakub Kuzilek^{1,2} · Zdenek Zdrahal¹ · Jonas Vaclavek¹ · Viktor Fuglik^{1,3} · Jan Skocilas⁴ · Annika Wolff⁵

Accepted: 11 July 2022 / Published online: 2 August 2022
© The Author(s) 2022

Abstract

Student drop-out is one of the most critical issues that higher educational institutions face nowadays. The problem is significant for first-year students. These freshmen are especially at risk of failing due to the transition from different educational settings at high school. Thanks to the massive boom of Information and Communication Technologies, universities have started to collect a vast amount of study- and student-related data. Teachers can use the collected information to support students at risk of failing their studies. At the Faculty of Mechanical Engineering, Czech Technical University in Prague, the situation is no different, and first-year students are a vulnerable group similar to other institutions. The most critical part of the first year is the first exam period. One of the essential skills the student needs to develop is planning for exams. The presented research aims to explore the exam-taking patterns of first-year students. Data of 361 first-year students have been analysed and used to construct “layered” Markov chain probabilistic graphs. The graphs have revealed interesting behavioural patterns within the groups of successful and unsuccessful students.

Keywords First-year students · Markov chains · Exam-taking patterns · Similarity · Retention

This work was supported by a junior grant research project by Czech Science Foundation GACR no. GJ18-04150Y and by the Research Cluster Digitalization, Diversity and Lifelong Learning - Consequences for Higher Education (D^2L^2) of the FernUniversität in Hagen, Germany.

✉ Jakub Kuzilek
jakub.kuzilek@hu-berlin.de

¹ CTU in Prague, CIIRC, Jugoslavskych partyzanu 1580/3, Prague, Czech Republic

² Humboldt-University of Berlin, Unter den Linden 6, Berlin, Germany

³ Computer Science Centre, Charles University, Ovocny trh 560/5, Prague, Czech Republic

⁴ Faculty of Mechanical Engineering, CTU in Prague, Prague, Czech Republic

⁵ Department of Software Engineering, School of Engineering Science, LUT University, Lappeenranta, Finland

Introduction

European Union higher education institutions face student drop-out. Between 20% and 54% of students fail to complete their degrees (Quinn, 2013). The situation in the engineering disciplines is especially grave (Kabra, 2011; Zdrahal et al., 2016). In combination with the increasing education costs, universities face a significant challenge.

Thanks to the increased usage of Information and Communication Technologies (ICT) in learning and management, the increased education costs are partially compensated (Bowen, 2015). Universities use information systems for recording and storing study-related data such as student demographics, course contents, student results, etc. These data open the opportunity to understand underlying educational processes. However, many higher education institutions (HEIs) collect the data in a diverse ecosystem of systems, making it challenging to extract, collect and analyse available data. Still, even in the case of having only student outcomes, it is possible to uncover the relevant information (Pandey & Sharma, 2013).

In that context, HEIs employ various methods, including Educational Data Mining and Learning Analytics research fields. One of the most promising is Artificial Intelligence, which starts influencing all aspects of everyday human life. In education, various attempts to understand educational processes and even predict future outcomes of students using advanced AI techniques have emerged (Papamitsiou & Economides, 2014).

Critical aspects such as instructional styles, faculty expectations, or the learners' behavioural, cognitive, motivational and developmental capabilities have been associated with the first-year student' success (Daempfle, 2003). First-year students need to adapt to the new academic life and learn how to efficiently manage time and study practices to succeed in their studies (Sebesta & Bray Speth, 2017). These study aspects are well described by metacognition and self-regulated learning (SLR) (Zimmerman, 1989). In SRL, the students act as active participants in their learning. Systematic and proactive students are manifesting three aspects of SRL (Sebesta & Bray Speth, 2017): motivation processes (interest in studies, accepting responsibility); metacognitive processes (planning, goal setting, monitoring the learning progress, self-evaluating); and behavioural processes (seeking information and advice, adopting effective study strategies). The student's ability for setting goals and planning/time management is vital metacognitive processes recognised themselves concerning the success in academic career (Winne, 2013; Sebesta & Bray Speth, 2017).

The Faculty of Mechanical Engineering, Czech Technical University in Prague¹ (FME) faces the issue of first-year engineering student retention. The FME, like many HEIs in post-Austrian-Hungary countries such as Austria, Czech and Slovak Republic, etc., follows a typical educational framework (Van der Plank et al., 2012). University students study several courses during the academic year, divided into the winter and summer semesters. Every semester is then followed by the exam period.

¹ <https://www.fs.cvut.cz/>

Students need to learn how to manage and plan their exams within the predefined time to successfully progress through the exam period and follow the study trajectory based on a study plan. Thus, it is essential for them to learn how to organise their time efficiently (“learn how to learn”). To achieve this, providing students with the ways leading to success and showing them how the time is organised by students who have been successful in the past may lead to better adaptation to the new academic life and increase the odds of the student passing. Thus our research focuses on extracting the exam-taking patterns of successful, passing and failing students to provide the students and teachers with insights on how the students should organise the exam period.

Research Questions

Our research focuses on the questions: *How do students organise their exam period?* and *What, if any, are the differences in exam-taking behaviour between successful and failing students?*. To answer them, the method - which transforms the student exam attempt log into a probabilistic graphical representation - will be introduced. The FME use-case data will be divided into three groups based on final student outcomes (successful, passing, and failing students); the Markov chain models will be constructed and compared.

State-of-the-art

The boom of ICT brings new possibilities of measuring, storing, and analysing various processes and phenomena. Higher education is no exception, and universities use different tools to collect and analyse study-related data. These data can uncover essential knowledge about the study process (Papamitsiou & Economides, 2014). The analysis of the data has been primarily aimed at predictive modelling of student performance using data collected from Virtual Learning Environments (VLEs), which has been proved to be the source of helpful information (Arnold & Pistilli, 2012; Kennedy et al., 2015).

Student demographics, combined with the information about their study history without any VLE related data, can be used for the estimation of students’ success in their studies (Sharabiani et al., 2014; Shehata & Arnold, 2015). Howard et al., (2018) deployed the Early Warning System with Bayesian Additive Regressive Trees to predict the student grade after the sixth week of the semester with a mean absolute error of 6.5 percentage points. Romero et al. (2013) predict the students’ performance based on the student participation in discussion forums. The accuracy of the proposed approach varied based on the selected machine learning algorithm for model building around 80%.

In the case that only students’ past performance (grades) is available, it can be still used to model students’ performance (Pandey & Sharma, 2013). The J48 decision tree builds using the performance data achieved an accuracy of 80.15%.

Furthermore, machine learning methods can be used to analyse face-to-face learning. For example, Kent et al. (2017) explored the relationship between student

engagement and learning outcomes in the context of traditional brick and mortar UK university.

The basic idea of student performance modelling can be further extended to analysing student activity (behaviour) within the VLE, focusing on student performance. The modelling of the student activity helps to uncover patterns corresponding to the manifestation of SLR metacognitive processes. The approach can make use of various methods. For example, Fincham et al. (2019) identified the learning behaviour patterns by building Hidden Markov Model and found a significant association between the students' learning strategies and their academic outcomes. Matcha et al. (2019) applied process mining and clustering to extract study behaviour from click-stream data. The detected study strategies were associated with the course performance.

Hlosta et al. (2014) proposed two methods for student activity analysis in VLEs: General Unary Hypothesis Automaton and Markov chains. The researchers extracted data from the online moodle-like platform the Open University (UK) used and analysed how students interacted with the VLE system every week. This analysis uncovered how a student behaves within such online educational systems and the bottlenecks in the course design. This idea was extended by Okubo et al. (2017). The authors deployed the Markov chain-based method at Kyushu University and implemented it as a Moodle analysis module. The developed plugin served the university teachers for adjusting the learning content more appropriately.

More recent work of Davis et al. (2016) employed Markov chains in the analysis of Massive Open Online Course data from edX and Coursera courses with over 100,000 students. They were also interested in the interaction of students with the online learning system and the learning path students took. The research found that failing students tend to review course materials more often.

In 2018 Kuzilek et al. (2018) used Markov chains on the same task as Hlosta et al. (2014) with the students' VLE activities generalised to the higher level of abstraction, leading to uncovering a passive withdrawal pattern within the VLE behavioural data.

Marques and Belo (2011) applied Markov chains on analysis of behavioural profiles at a Portuguese university. They could isolate the most common patterns students follow when browsing through the VLE system containing the educational resources.

To conclude, most of the current methods for analysing students' learning behaviour focus on online environments, which provide rich information about each student's learning path when focusing on the selected learning goal. The approaches based on the online learning data demonstrate the importance of capturing student learning behaviour. However, there exists the evidence (Pandey & Sharma, 2013; Zdrahal et al., 2016) suggesting the use of only sparse student data can be used for modelling student success. Thus, having the information about student exam-taking sequence might enable proper data transformation to uncover underlying meta-cognitive processes of students when managing their time during the exam period. We will also show that the student behaviour captured by exam-taking sequence records can reveal vital information about student time management and identify patterns that lead to success.

Table 1 First semester courses

Exam courses		Non-exam courses	
Name	Credits	Name	Credits
Mathematics I.	6	Engineering Design I.	3
Constructive Geometry	5	Fundamentals of Technology I.	3
Physics I.	5	Computer Support for Study	3
		Management Skills	2
		History of Technology	2
Total credits	16	Total credits	13

Educational Setting at FME

The Czech Technical University in Prague is one of the biggest and oldest technical universities in Europe.² Currently, it has eight faculties and about 18,000 students studying 160 study programmes. One of its faculties is the Faculty of Mechanical Engineering, which provides studies in three bachelor's and six master's study programmes. It has approximately 3,000 students currently pursuing their degree, from which 400 registered for studies in the Theoretical fundamentals of Mechanical Engineering programme, which is the flagship bachelor programme provided by the faculty enrolled by most of the students.

The faculty does not use any form of centralised VLE system. Most online learning platforms serve only to distribute educational materials, and no online activity is recorded.

The typical study year at FME is divided into the winter and summer semesters. Both followed by the exam period and both 14 teaching weeks long. The Winter semester usually starts at the beginning of October and has winter holidays in the last quarter. The summer semester begins after the winter exam period, followed by the exam period, which ends the academic year.

The exam period has six weeks (five regular and one extra for first-year students). Every week, each course usually has one or more exam dates listed. Every student has two attempts to pass the exam, and if failing both, the dean can allow a third attempt.

To simplify the transition from high school, first-year FME students have a prescribed selection of courses they should attend in the winter semester (Table 1). There are three exam courses (as advertised by FME): Mathematics I. (M), Constructive Geometry (C) and Physics I. (P). The other courses are completed by completing the laboratory tasks and seminar work. Exam courses grading is done in line with the EU ECTS credit, and scoring system (Bonjean, 2019), where A-E represents passing mark and F non-passing mark. Each course is awarded points (credits) reflecting students' time and work requirements to pass. To progress into the second semester, students

² <https://www.cvut.cz/en/welcome-to-ctu>

Table 2 Example of exam attempt log

Student id	Date	Course	Result
15456106	2018-01-22	Mathematics I.	F
15456106	2018-02-16	Mathematics I.	E
15456106	2018-02-24	Constructive geometry	D
15456106	2018-02-29	Physics I.	A
13274506	2018-01-26	Mathematics I.	B
...

need to achieve at least 15 ECTS credits, thus succeeding in at least one exam course and all non-exam courses is required to pass the first semester.

Data

There were 361 incoming students registered for the studies in the study programme *Theoretical fundamentals of Mechanical Engineering* in the academic year 2017/2018. From them, we removed those who were allowed a third exam attempt by the dean. Additionally, 24 students who passed all exams but were unregistered from studies due to various external non-study related reasons such as long-term illness or family issues have been excluded from the further investigation - the exclusion results in a dataset of 311 students. The university data warehouse contains dates and outcomes of all student exam attempts in the winter semester. Thus, the extracted data includes three to six pieces of information for every student in the form of an exam attempt log. A randomly generated example of attempt logs is shown in Table 2. The log contains the ids of the students, the date of the exam, course and the exam outcome. The first student in our example took four exams in total, with the first failed attempt in the Mathematics I. exam (mark F). This example will be further used for the explanation of our analytical approach.

Methods

This section describes the process of creating the probabilistic graph model. The method of forming the model is demonstrated on data from FME, and its extension to different settings is straightforward. First, the transformation of the attempt log to the space of student exam states is introduced. Next, the building of the Markov chain model is explained, the measures of graph complexity and the methodology for the testing of similarity between graphs is described.

Transforming Student Attempts to Student Exam States

A week represents a typical preparation time for the exam. Thus the attempt-log is aggregated to the level of weeks using the starting date of the exam period. This results in allocating each recorded attempt to one of the six weeks.

A student can be in four possible “situations” regarding the exam:

0. the student did not attempt the exam so far,
1. the student failed the first attempt,
2. the student failed the second attempt (and failed the course),
3. the student passed the exam (and passed the course).

The students have three exams in the first academic semester, and the resulting state space has $4^3 = 64$ possible states for every week of the exam period. The student weekly exam state (one of possible 64 states) can be transformed into the vector \mathbf{e} of length 3:

$$\vec{e} = (P, C, M), \quad (1)$$

where $P \in \{0, 1, 2, 3\}$, $C \in \{0, 1, 2, 3\}$ and $M \in \{0, 1, 2, 3\}$ represents students’ current “situation” in Physics I. (P), Constructive geometry (C) and Mathematics I. (M). State 0 corresponds to state when the student did not attempt the exam, state 1 to state when the student failed the first attempt, state 2 to the state when the student failed the second attempt (and failed the course), and state 3 to state when student pass the exam (and the course).

The vector \mathbf{e} represents quaternary (base-4) numbers, which can be transformed into a decimal number X :

$$X = 16P + 4C + M. \quad (2)$$

The number $X \in 0, 1, 2, \dots, 63$ represents the student’s weekly exam state, which accumulates their achievements in the exam period so far. $X = 0$ represents the fact that the student did not attempt any exam and $X = 63$ represent the fact that the student passed all three exams. Table 3 shows the complete list of possible states.

The students’ weekly states table is used to construct a probabilistic graphical model, e. g. Markov chain. Every student’s progression through the exam period is represented as the sequence of 6 states. In our example the student 15456106 attempt log is transformed to the weekly exam states containing states $S_{student} = \langle 1, 1, 1, 3, 15, 63 \rangle$. This sequence is illustrated in Fig. 1. In each week, the student’s state accumulates their results so far. Thus, the student has the same state in the first three weeks.

The sequence fully characterizes each student’s exam-taking pattern. It reflects how students progress through the exam period, and it is vital to identify patterns of successful and unsuccessful students. Since each student sequence is unique to that person, it is necessary to construct the general “scheme” of the exam-taking pattern. To construct the general pattern, we employed the technique of Markov chains.

Markov Chains

Markov chains is the statistical technique to create the model for describing the sequence of events, in which the probability of the next event depends only on the previous event. The

Table 3 Complete list of all possible weekly student-exam states

State	P	C	M	State	P	C	M	State	P	C	M	State	P	C	M
0	0	0	0	16	1	0	0	32	2	0	0	48	3	0	0
1	0	0	1	17	1	0	1	33	2	0	1	49	3	0	1
2	0	0	2	18	1	0	2	34	2	0	2	50	3	0	2
3	0	0	3	19	1	0	3	35	2	0	3	51	3	0	3
4	0	1	0	20	1	1	0	36	2	1	0	52	3	1	0
5	0	1	1	21	1	1	1	37	2	1	1	53	3	1	1
6	0	1	2	22	1	1	2	38	2	1	2	54	3	1	2
7	0	1	3	23	1	1	3	39	2	1	3	55	3	1	3
8	0	2	0	24	1	2	0	40	2	2	0	56	3	2	0
9	0	2	1	25	1	2	1	41	2	2	1	57	3	2	1
10	0	2	2	26	1	2	2	42	2	2	2	58	3	2	2
11	0	2	3	27	1	2	3	43	2	2	3	59	3	2	3
12	0	3	0	28	1	3	0	44	2	3	0	60	3	3	0
13	0	3	1	29	1	3	1	45	2	3	1	61	3	3	1
14	0	3	2	30	1	3	2	46	2	3	2	62	3	3	2
15	0	3	3	31	1	3	3	47	2	3	3	63	3	3	3

The column *State* contains the numeric value of the state; columns *P*, *C*, *M* are corresponding states in each course with possible values: 0 = The exam is not attempted, 1 = Failed 1st attempt, 2 = Failed 2nd attempt, 3 = Passed

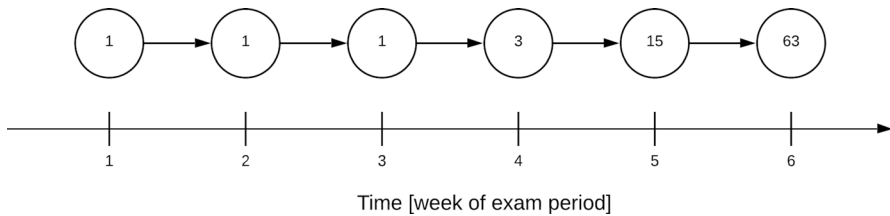


Fig. 1 Student exam sequence for example student 15456106

primary benefit of representing event sequence using the Markov chain is the simplicity of the model building. It is beneficial for modelling discrete-time and discrete space stochastic processes. On the other hand, the assumption of dependency on the previous state (event) means that the model works with no memory and cannot provide more complex information about the process. In our case, the sequence of events is represented by the student exam sequence in six exam period weeks, where each state accumulates the information from previous states. The limitation of Markov chains regarding the memory is partially compensated, keeping the advantage of creating a general model using “simple” transition probabilities between states. The following paragraph will introduce the technique in more detail.

A sequence of random variables forms a discrete-time random process $\{X_n\} = \{X_n : n = 0, 1, 2, \dots\}$, where n is the index. The process X_n takes on values of a

countable set $S \in \{s_1, s_2, \dots, s_N\}$, which is called a state space within which the values s_i represent the individual states. The process X_n is then called a Markov chain if it satisfies the Markov property:

$$P(X_n = j | X_0 = x_0, \dots, X_{n-1} = i) = P(X_n = j | X_{n-1} = i), \quad (3)$$

for all $i, j, x_0, \dots, x_{n-2} \in S$ for all $n = 1, 2, 3, \dots$. This means that the process and its future state is dependent only on the present state and not on past states. For such a process we can define a square transition probability matrix \mathbf{P} with dimension $|S| \times |S|$ and transition probabilities (elements) $p_{i,j}$ such that:

$$p_{i,j} = P(X_{n+1} = j | X_n = i). \quad (4)$$

Transition matrix \mathbf{P} is a stochastic matrix with non-negative elements such that: $\sum_{j \in S} p_{i,j} = 1$, for any $i \in S$. The Markov chain is called homogeneous if its transition probabilities do not depend on the time, i.e.: $P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i)$. Otherwise, it is called non-homogeneous.

Students accumulate their exam results in their exam state, sampled every week of the exam period. The accumulation of results in each exam state means that the transition to the next week's state depends only on the most recent state. This fulfils the Markov property, and a Markov chain probabilistic model can be constructed (Norris & Norris, 1998). The model is specified by the set of all possible states from all exam weeks S of length $6 * 64 = 384$. Using set S , the transition matrix P is constructed. The entry in the i -th row and j -th column of matrix P represents the probability p_{ij} that a student moves from exam state s_i to the exam state s_j . The transition probability p_{ij} represents the proportion of the time that student ends up in state s_j being in state s_i in the previous week. For example, for state 3 in week 2, 45 students end up in state 15 in week 4 out of 60 students in state 3 in week 2. That means that $p_{3,15} = 0.75\%$ between week 3 and 4.

Transition matrix P will be sparse with only non-zero elements representing transition sub-matrices from week to week. The layered structure of the transition matrix suggests that the resulting Markov chain is non-homogeneous. The states in the last layer (the final exam week) are absorbing states (state cannot be left once entered) (Norris & Norris, 1998). The resulting transition matrix is sparse and satisfies the condition of being in the canonical form (Norris & Norris, 1998):

$$P = \begin{pmatrix} Q & R \\ \mathbf{0} & I_r \end{pmatrix}, \quad (5)$$

where Q is t -by- t square matrix containing probabilities of transitioning between t states, R is the t -by- r matrix of probabilities of transitioning into r absorbing states, $\mathbf{0}$ is r -by- t zero matrix and I_r is r -by- r identity matrix. In our case matrix Q represents student transition probabilities of states in weeks 1 to 5 and matrix R transition probabilities from week 5 to week 6 states. When a student progresses through the exam period, he moves between layers (weeks) of the Markov chain starting in the far left layer and ending in the far-right layer.

Student Performance Groups

The students can be divided into three groups according to their success in the winter semester and whether they progressed to the second academic year:

- *Successful students.* Students passed all courses (3 exams) in the winter semester and proceeded to the second academic year.
- *Passing students.* Students passed most of their courses (1 or 2 exams) and proceeded to the second academic year. Students in this category achieved enough ECTS credits to continue their studies. However, they did not achieve a full “score” of 29 credits.
- *Failing students.* Students failed all three exams in the winter semester and did not accommodate enough credits to proceed even though they might pass all non-exam courses.

Estimation of the Graph Complexity

Markov chains are, in principle, directed weighted graphs, which can be examined in several ways. One possibility is to investigate graph complexity, which reflects how many elementary graph transformations are needed to construct the graph (Zinovyev & Mirkes, 2013). In another world, it reflects how “complicated” the graph is. Measuring the complexity of a directed graph can be done using various measures such as DAG-width, directed treewidth, girth or polynomial-based complexity (Dehmer et al., 2019). We adopted the last approach, which is based on constructing graph polynomials based on the out- and in- degrees of the graph. The technique does not suffer from high degeneracy (Dehmer et al., 2019). We selected average complexity measure \hat{I} calculated as a mean value between out- and in- zeros δ of corresponding graph polynomial on the interval (0, 1) (see Dehmer et al. (2019) for details). The computed measure is in the interval (0, 1), where values close to 1 represents more complex directed graphs and values close to 0 less complex graphs.

Comparison of Graphs

Markov chains are probabilistic graphs (Norris & Norris, 1998) represented by a weighted adjacency (transition) matrix P of dimension $n \times n$, where n is the number of possible states. In our case $n = 384$. The transformation of the graph into the high-dimensional space using the so-called kernel function (Samatova et al., 2013) has been used to compare graphs. Using the kernel function, one can construct a positive semi-definite transformation kernel matrix, which represents a special case of Mercer’s theorem (Mercer, 1909). Thus, the kernel function serves as a valid similarity measure (Samatova et al., 2013). The shortest-path kernel method (Borgwardt & Kriegel, 2005) measures similarity by finding the all-pairs shortest paths using a direct product graph. The algorithm produces the similarity matrix M containing numbers equal to the number of shared shortest paths between graphs for every pair of graphs on the algorithm input. The matrix M is normalized using:

Table 4 Number of students and their proportion to the whole cohort in performance groups

Group	Number of students	Proportion
Successfull	153	49.2 %
Passing	63	20.3 %
Failing	95	30.5 %

$$M_{ij} = \frac{M_{ij}}{\sqrt{M_{ii}M_{jj}}}, \quad (6)$$

where M is the similarity matrix and M_{ij} is the element of matrix M on i -th row and j -th column. The resulting matrix has elements ranging from 0 to 1, where 1 represents equality and 0 complete dissimilarity of two graphs.

Extracting Exam Taking Patterns

The resulting Markov chain graph is complex and reflects reality by tracking all exam paths taken by students in the group. However, the complex Markov chain is not suitable for analysing the patterns carried out by the majority of students, which reflects the “behaviour” of students in a more general way. For this purpose, the absolute values transition matrix has been extracted in addition to the transition matrix. It contains the same information as the transition matrix, but instead of probability, it shows the absolute values of how many students have taken this path. Using the absolute values, we extracted the Markov chain graph edges, taken by more than 10% of the students in the performance group. The application of this extraction approach enabled us to extract the predominant patterns of exam behaviour within different groups of students.

Results & Discussion

This section presents the results and discussion of student exam behaviour modelling using Markov chains and computed similarities between patterns of students in different performance groups using the shortest path kernel method.

The student cohort has been divided into performance groups according to the methodology presented in “[Student Performance Groups](#)”. The number of students in each performance group is shown in Table 4. We can observe that approximately 50 % of students pass all requirements for completing the first academic year and continue their studies without any additional workload in the following year. Additionally, 20.3 % of the student cohort do not complete all courses successfully yet pass the minimum requirements to proceed to the next academic year. However, they need to invest additional effort to finish the failed first-year courses in the next academic years. Finally, 30.5 % of the student cohort do not fulfil the minimum requirements, and they are de-registered from the studies by the FME.

Table 5 Markov chain statistics for each performance group of students

Measure	Successful	Passing	Failing
Nodes	48	84	114
Edges	83	134	206
Complexity	0.1844	0.2896	0.7474

For each student performance group, the Markov chain model has been constructed. The model represents the probabilities of students taking the path from state s_i in one week to state s_j in the second week. The graph is represented by the transition matrix. The constructed Markov chains can be found in figures in Appendix 1. Nodes correspond to weekly student exam states, and an edge represents the probability that a student in state s_i proceeds to state s_j . Each week exam states are positioned into layers to help visualise the constructed graphs. Each layer then represents one week of the exam period. For each performance group, two graphs have been built. The first one has the labelling of each node corresponding to the exam state number the second one then has labels corresponding to the number of students in the exam state depicted in the first graph.

For each constructed Markov chain, the number of nodes, edges and the complexity \hat{I} has been computed, and the result is presented in the Table 5. We can observe that the number of nodes is the lowest for the students in the Successful group and the highest in the performance group of Failing students. The same observation can be made for the graph edges and the complexity \hat{I} . This suggests that successful students follow a limited number of exam-taking paths during the exam period and can pass the exam on the first attempt (this can be verified by observation of Markov chain in Fig. 7). The limited number of exam-taking patterns in the group of successful students suggest that specific order of exams might influence the winter exam period outcome, thus influencing the result of the first academic year.

Applying the extraction technique described in “[Extracting Exam Taking Patterns](#)” to the constructed Markov chains for the three performance groups resulted in the significant simplification of Markov chains and led to extracting pathways with the nodes “containing” the most important proportion of student cohort. The following sections will describe these extracted pathways enriched with findings from the original “large” Markov chains presented in Appendix 1.

Comparison Between Performance Groups

The similarity between successful and passing students is higher than between successful and failing students. The similarity between passing and failing students is higher than the similarity between successful and failing and lower than between successful and passing. The expected “ordering” of student groups shows that the minor differences between graphs make a significant difference in the student outcomes.

The most notable observation, which can explain the difference between student performance groups, is the pace of taking exams. The most successful students start

Table 6 Similarities between Markov chains for different performance groups

	Successful	Passing	Failing
Successful	1	0.997	0.954
Passing	0.997	1	0.973
Failing	0.954	0.973	1

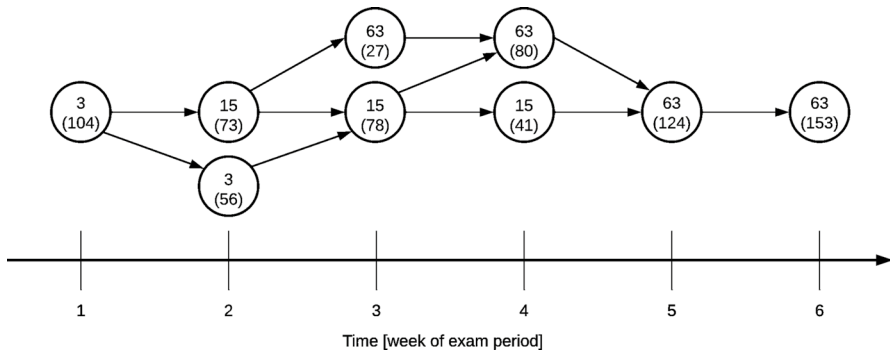


Fig. 2 Simplified Markov chain for the group of successful students in the winter exam period of academic year 2017/2018

at the beginning of the exam period and progress through the exams faster with minimum delays. The failing students tend to have a slower attempt ratio, leaving them less time for the proper exam preparation.

The computed similarities suggest that a student’s transition from the low-performance group to the high-performance group is possible since the difference in graph similarities is small.

As the last step of our analysis, we compared the normalized similarities between Markov chains of student performance groups (Table 6). The computed similarities are quite high (> 0.95). The high similarity is because all the behaviours share common patterns related to the strict pace of the exam period. However, we can observe that the speed at which students progress through the exam period is essential.

Successful Students

Figure 2 shows the extracted Markov chain for the group of successful students. The nodes are labelled with the state numbers, and the value in brackets represents the number of students in a corresponding state. Since successful students are selected as a group of those who finished all the exams successfully, there is only one absorbing (end) state 63, which reflect the situation when the student passed all the exams. The most common path in the sense of student absolute numbers and probabilities is Mathematics I. (state 3 in week 1), Constructive geometry (state 15 in week 2), a week or two pause (states 15 in weeks 3 and 4) and Physics I. (state 63 in week 4 or 5). This path suggests that the Physics I. exam requires more preparation time than

the Constructive geometry, which is usually taken one week after the Mathematics I. exam.

No one was able to complete the exam period in under three weeks. Only 27 students (approx. 18%) were able to finish all the exams within three weeks (state 63 in the 3rd week). So that most of the successful student cohort make at least a small “pause” between two consecutive exams. Based on the numbers, the pause is usually between the Constructive geometry and Physics I. exam.

Students who passed Mathematics I. in the first week (state 3) are divided into two groups. The first took a one-week pause (state 3 in the second week), and the second took the Constructive geometry exam (state 15 in the second week). There is a minority of students who need the second attempt to pass the exam. This pattern is not demonstrated in Fig. 2 but can be observed in the full-scale Markov chain in Appendix 5.

No student within the group failed the Mathematics I. exam in the first week. When exploring this anomaly, it has been uncovered that students with good performance during the semester has been offered to pass the course without the exam. Approximately $2/3$ of the successful students passed the exam like this. The Mathematics I. course was the only one offering the pass based on the semester performance.

The observed patterns suggest that successful students tend to schedule the exams in the first weeks of the exam period. The observed pattern of finishing Mathematics I.; Constructive geometry; and Physics I. suggests that the Constructive geometry course is considered the “simplest” one from all exam courses in the winter exam period. This is also supported by the fact that approximately half the students took the Constructive geometry exam week after the Mathematics I. exam.

Passing Students

The resulting extracted Markov chain for a group of passing students can be found in Fig. 3. At first, it is worth noting that there is no clear pattern in taking the exams by the students compared to the group of successful students. The Markov chain shows a significant portion of students who did nothing (state 0) in the first exam week (41 out of 63 students). Most of the students active in that first week took and passed Mathematics I. (state 3). However, from students who finished Mathematics I. in the first week, the majority make a “break” in the following week. The same pattern is observed in the group of successful students. Following this pattern in the complete Markov chain in Appendix 1 shows that most students also took “pause” the following week before taking any other exam. This suggests two possible scenarios: 1) students underestimate the time required to prepare for the next exam, or 2) students are required to finish requirements for the non-exam courses. The second reason suggests the importance of focusing on finishing the non-exam courses before the exam period.

Most students did not succeed in Mathematics I. and Constructive geometry within the first two weeks. However, 15 students finished both exams in week 3.

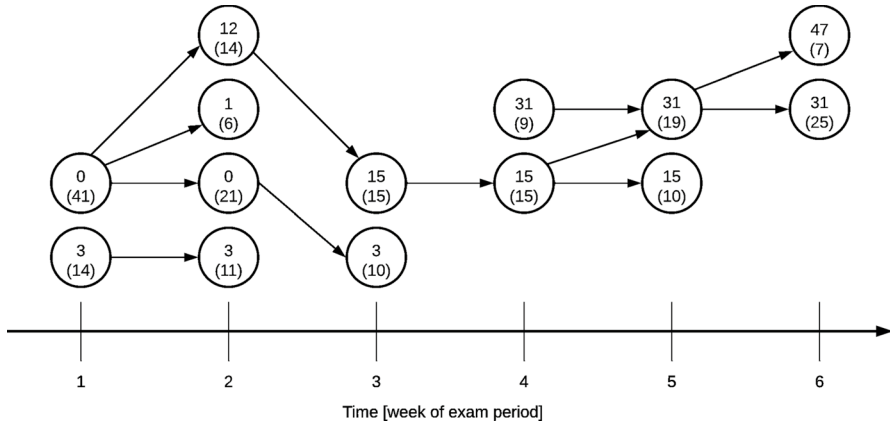


Fig. 3 Simplified Markov chain for the group of passing students in winter exam period of academic year 2017/2018

The state (and preceding pattern) suggests an overlap with the group of successful students. We can expect that the difference between passing and successful students is in the individual requirements for preparing for the exams. This also suggests that teachers should focus on students who did not finish the combination of Mathematics I. and Constructive geometry exams before week 3 to support them in preparation for the Physics I. exam. The state 31 in weeks four to six shows that a significant portion of passing students fail to pass the Physics I. exam, suggesting that the time requirements for exam preparation are higher than most students estimate.

Most students do not take the second attempt on the Physics I. exam. This suggests that students are moving their last attempt to the summer semester or the next academic year, making them vulnerable to failure in the next period of their academic life due to increased workload.

The students usually pass Mathematics I. in the first three weeks of the exam period, followed by the successful attempt in Constructive geometry. This pattern of passing Mathematics I. followed by successful completion of the Constructive geometry is shared with the successful students, suggesting that this is the most easily achievable outcome from the first semester. Thus, teachers should recommend that students who did not achieve both exams focus on them to continue their studies.

There are 21 students (approx. 33% of students in a group) who did not take any exam attempt for the first two weeks of the exam period. Those students might either struggle with preparation for the exam; or focus on finishing the non-exam courses they did not manage to finish before the exam period. Passing students' performance group is manifested by starting the exams later in the exam period. This suggests that they are more under time pressure and increased workload during the exam period. Thus, one possible solution is to recommend that they start with the exams earlier and increase the teacher support before the exam period to help students finish non-exam courses before the end of the semester.

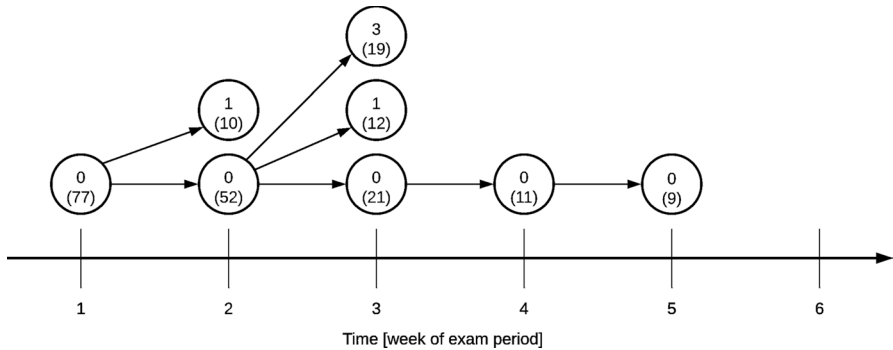


Fig. 4 Simplified Markov chain for the group of failing students in winter exam period of academic year 2017/2018

When observing the full Markov chain in Appendix 1, it can be observed that there are more possible end states for the group of passing students. This is because students in this performance group usually do not finish all the required exams, and they end with one or more unsuccessful attempts in one or two exams. The majority of students in this group fail to pass the Physics I. exam, suggesting that this exam is the most problematic from the student’s point of view. It might be recommended to examine the reasons in more detail to adjust the course, course content or the examination more appropriately.

Failing Students

Figure 4 presents the simplified Markov chain for the performance group of failing students. It can be observed that students in this performance group predominantly did not attempt any exam in the first week, and approximately half of the students did nothing in the second week. This suggests that this group might have problems with the preparation for the exams or have difficulties finishing non-exam courses during the semester, requiring them to allocate the time during the exam period. This observation is also made in the previous group of passing students, suggesting that in detail analysis is required by the teaching personnel to uncover necessary details to assist the students appropriately.

The 19 students passed the Mathematics I. exam in the third week. Still, they did not proceed to the second academic year. This suggests that they either: 1) did not finish all non-exam courses in the winter semester (leaving them with not enough ECTS credits and being de-registered from the studies), 2) or did not achieve all requirements at the end of the first academic year (and again being de-registered). This suggests that not finishing all courses in the scheduled semester creates an additional workload, which might lead to undesired outcomes from the studies.

All students of the failing performance group attempted at least one exam (as seen in the complete Markov chain in Appendix 1). This suggests that all the students are partially involved in the learning, and the group of “passive withdrawal” students is missing.

The observation of long student periods of inactivity suggests that the workload for the exam preparation for all exams during the end of the exam period might be high. Also, there might be a problem with properly scheduling the exams since there are enough exam dates, but the number of possible combinations is reduced; thus, flexibility is reduced.

The observed patterns suggest that teachers increase students' awareness to adequately plan their activities during the exam period, try to finish non-exam courses at the end of the semester, and seek guidance when needed.

Summary of Findings

The evaluation of the produced Markov chain and deeper analysis of extracted “simplified” Markov chains uncovered interesting student exam-taking patterns. The following summarises the main findings:

- **Students starts with the Mathematics I. exam.** This pattern is presented within all performance groups. The first reason is the fact mentioned above that part of the students was offered the possibility to pass the course without an exam based on their performance in the semester. This also suggests that the students are the most prepared for the exam before the exam period starts. It also indicates this exam should be taken as soon as possible to accommodate enough time for the other two exams. Students in less performing groups finish the exam later than the successful students' performance group. This suggests that the students might need more time to complete non-exam courses during the exam period. This phenomenon needs to be examined in more detail.
- **There is usually a week gap before taking Physics I. attempt.** The week gap suggests that students need more time for the preparation for the exam indicating the requirements for passing the course are either higher, or the course needs more different knowledge than Mathematics I. and Constructive geometry exams, thus needs more time to accommodate the required knowledge.
- **Passing and failing students start taking exams later.** This pattern is in connection with previous observation. Again it suggests that the students either need to finish the non-exam courses during the exam period, or it is harder for them to gain enough knowledge/confidence before making an exam attempt. Still, the observed pattern suggests that the teachers should provide timely intervention to support students with proper suggestions regarding their study performance.
- **Failing students are inactive in the first two weeks of the exam period.** The pattern of starting the exams later is mainly manifested in the performance group of failing students suggesting these students have some study-related issues either with non-exam courses; or with preparation for the exams. This might trigger necessary intervention to support the students who did not attempt any of the exams.
- **Significant portion of passing students fail in Physics I. exam.** The observation suggested that the Physics I. exam is considered as the “hardest” by the students. It indicates that the course might need to make several adjustments to

address student needs better. With the observation of successful students' taking pattern of passing the exam, the last suggests that the course might require deeper understanding or comprehensive knowledge different from other courses.

The ability for setting goals and planning/time management is recognised by students themselves as a critical metacognitive process for study success (Sebesta & Bray Speth, 2017). Observed patterns can confirm this within all three performance groups and measured complexity of the constructed Markov chains. The successful students Markov chain has the lowest complexity suggesting that this group follows the defined, planned sequence of exams, fulfilling all the requirements to proceed to the second academic year. Also, the increased complexity of lower-performing groups suggests that students' metacognitive processes in these groups are not properly set, and students require more guidance from the teachers to succeed. This is vital, especially in the first weeks of the exam period, where passing and failing students probably need to also focus on non-exam courses, leaving them less time to prepare for the exam courses. This suggests that the teaching personnel should intervene during this vulnerable period, helping students with the proper planning and suggestions regarding the sequence and pace of the exams. Students of the successful group finish the exams usually in 4 to 5 weeks suggesting they are exhibiting some aspects of SRL (Zimmerman, 1989).

Markov chains are the technique for constructing a simple statistical model in which the state depends only on the previous state. However, by accumulating the student "achievements" in the student exam state, the Markov chain uncovers the sequence of exams taken by the majority of students in three different performance groups. The uncovered structure and complexity of constructed graphs suggest that students within low performing groups require more guidance. This guidance should occur in the first weeks of the exam period, where the difference between performance groups is manifested by the later attempt in the Mathematics I. exam. Based on the observations presented in this paper and our previous research, which results align with the presented findings, FME implemented several steps to support its students.

Intervention Framework

The analysis presented in this paper and published previously (Zdrahal et al., 2016) serve as a basis for the interventions carried out by FME staff. Interventions are phased and contain several steps, which identify at-risk students based on the achieved ECTS credits.

The first step takes place during the information campaign at the FME Open Day. Potential students are informed about the study, and they receive the first notification about the problematic study parts in the first study year. Next, during the student enrollment, the student gets the study information materials in the form of so-called White book³ with study-related information including contacts, a list of accredited study programmes, study plans and time requirements for each course. In addition,

³ <https://bit.ly/2WIMMES>

they are informed that there exists the Learning Analytics framework for identifying at-risk students and their support.

Three weeks before the winter exam period, students are invited to the optional lecture of Vice-Dean for Education, where the progression rules and the study statistics from previous years. They are informed about the best exam-taking strategies and collected experiences from last academic year students obtained from surveys. Usually, around 70 % of the first-year students attend the lecture. The video and slides from the talk are then available online for students to use.

In the first week of the exam period, help with the planning and preparation for passing the exams is offered to students with low credit scores (approximately 90 students). About 10 % of the invited students respond.

Two weeks before the end of the winter exam period, students with less than 10 ECTS credits earned are again offered help to plan the exam schedule for the last two weeks of the exam period. The intervention is carried out individually and anonymously. Usually, there are only a few students who responded to this call. These students are provided with help similar to those responding in the first week, focusing on their current circumstances.

In addition to these “active” steps to intervene at-risk students vice-dean is available for consultations at least once a week for the whole semester. The consultation meetings can also be anonymous and individually focused if students express their wishes.

FME organizes the voluntary faculty survey after the semester to improve the learning outcomes, where students evaluate every course they attended and every lecturer they met. Additionally, first-year students receive an additional survey in which they assess the difficulty of the first-year courses.

Qualitative Research

In parallel to the research of the student exam taking, the qualitative pedagogical investigation has been carried out. It focused on course content analysis, including lectures, seminars, and laboratory exercises accompanied by the guided semi-structured interview with lecturers and students.

The interviews helped to understand the underlying behaviour of students during the semester and exam period, including exam planning. The findings confirmed that failing students with delayed exams in the exam period has been “waiting” for the later attempt dates. The reasons for the “waiting” were procrastination, waiting for the peers to get insights on the exam requirements and more extended adaptation to the new environment due to the move to the capital city.

Conclusions

In this paper, we addressed our research questions: *How do students organise their exam period?* and *What, if any, are the differences in exam-taking behaviour between successful and failing students?* via analysing student exam data from the winter semester taken at traditional university. At first, the student exam log

has been transferred to the student exam state sequence. The resulting student sequences were then used for constructing Markov chains, which represented the behaviour of the selected cohort of students. Three different Markov chains have been built, reflecting other performance groups regarding student success in the winter semester and first academic year. Constructed graphs demonstrated different levels of complexity, with the lowest complexity presented within the group of successful students. The lowest complexity suggests that the successful students mastered the ability to plan and manage their time (Sebesta & Bray Speth, 2017). However, a more detailed analysis on complexity values within the graphs with different sets of states is required to understand better the underlying exam-taking phenomena. Such analysis would require simulation study with pre-defined conditions to enable the needed complexity analysis, and it is out of the scope of the presented research.

In contrast to this observation, another student's performance has more complex graphs suggesting that they might struggle with the planning issue leading them to a situation where either they need to focus on non-exam courses; or cannot prepare for the exams properly. The uncovered patterns suggested that the FME should intervene on the students who did not start taking exams in the first weeks of the exam period, helping them plan and acquire the required knowledge. It is also worth considering revising the Physics I. course regarding the required ability and difficulty. The successful students manifested a clear exam-taking pattern, suggesting that this is a manifestation of SRL metacognitive processes every student needs to acquire to succeed in the academic career. In addition, the first round of interventions and qualitative pedagogical research has been carried out. The interventions provided students with additional help, and there are first indications that the interventions improved the retention of students. However, a detailed analysis of the intervention framework efficiency has not been carried out so far. Such a study has been planned for years after the completion of this study. The qualitative research included the guided interviews with lecturers and students confirmed the observations exam-taking patterns and provided more insights into the reasons behind the practices leading to the failure.

Appendix 1

This appendix presents the full scale Markov chains of three different performance groups (Figures 5, 6, 7, 8, 9 and 10). The label of the node represents the name of the state. Each colourful layer of nodes represents one week of the winter exam period. Every edge aims from the left node s_i in week n to the right node s_j in the week $n + 1$. The colour of the edge represents the transition probability from one node to another. Lighter colours denote a lower probability, darker a higher one.

In addition to the transition graph, the same graph with different node labels has been constructed. Each label represents the absolute number of students in the state. The assembled “re-labeled” graphs are shown on Figs. 6, 8 and 10.

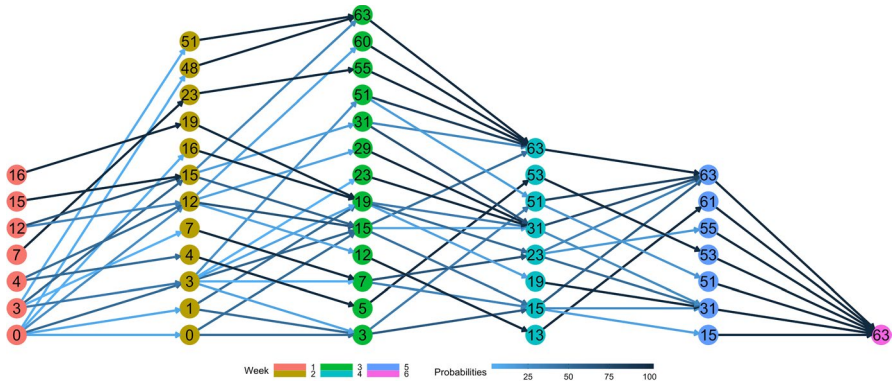


Fig. 5 Markov chain for the group of successful students in winter exam period of academic year 2017/2018

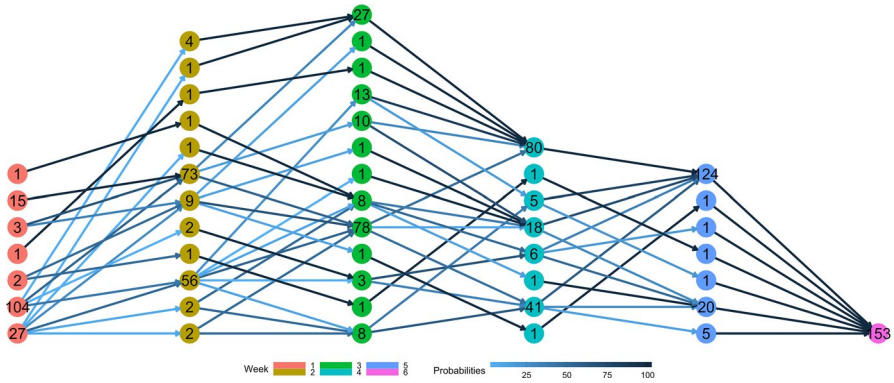


Fig. 6 Markov chain with labels representing number of students for the group of successful students in winter exam period of academic year 2017/2018

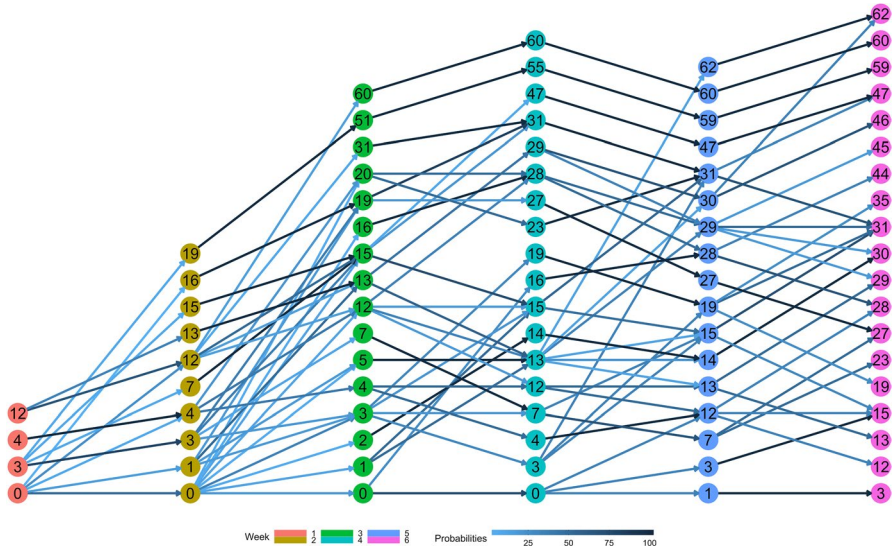


Fig. 7 Markov chain for the group of passing students in winter exam period of academic year 2017/2018

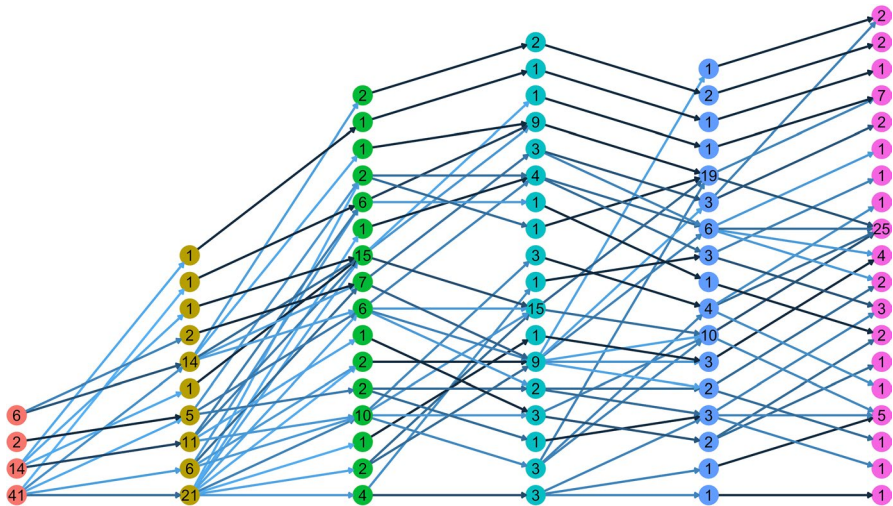


Fig. 8 Markov chain with labels representing number of students for the group of passing students in winter exam period of academic year 2017/2018

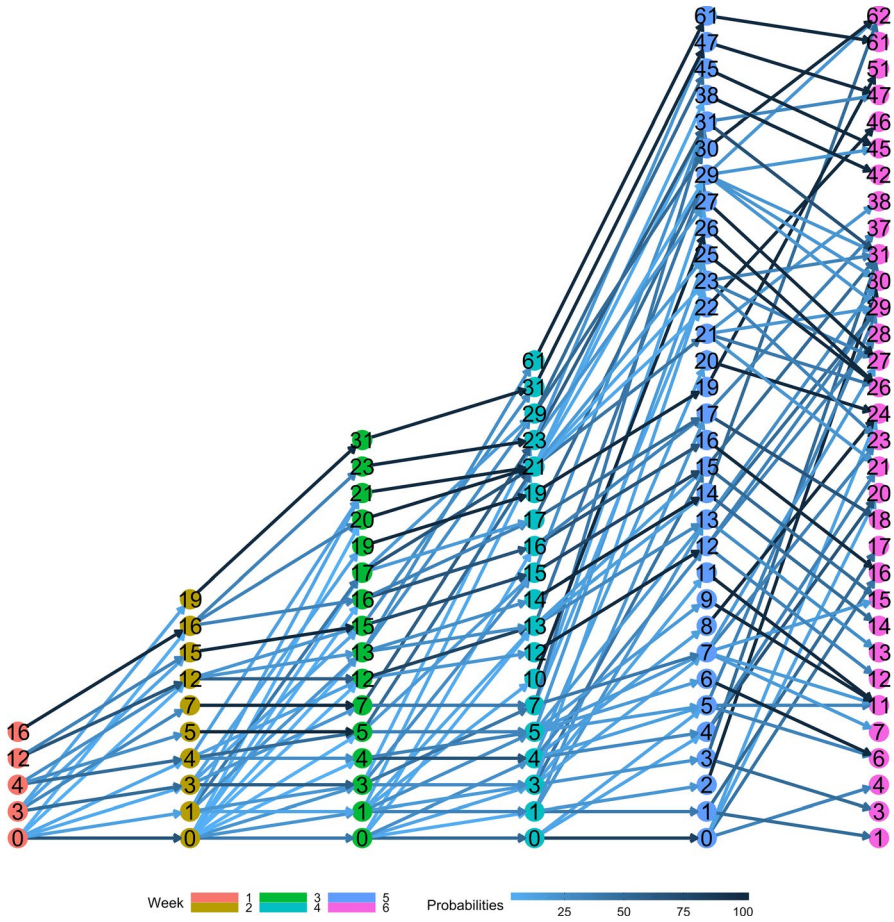


Fig. 9 Markov chain for the group of failed students in winter exam period of academic year 2017/2018

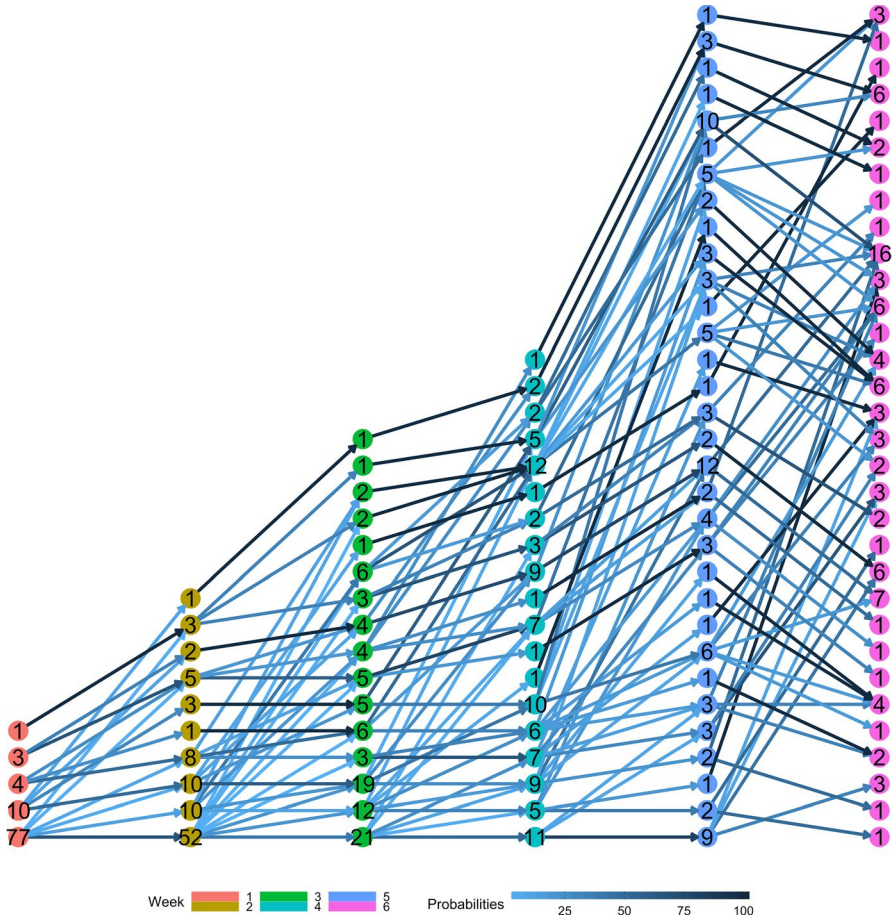


Fig. 10 Markov chain with labels representing number of students for the group of failed students in winter exam period of academic year 2017/2018

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arnold, K.E., & Pistilli, M.D. (2012). Course signals at purdue: Using learning analytics to increase student success. *Proceedings of the 2nd international conference on learning analytics and knowledge* (May), 267–270. <https://doi.org/10.1145/2330601.2330666>
- Bonjean, D. (2019). European credit transfer and accumulation system (ects). https://ec.europa.eu/education/resources-and-tools/european-credit-transfer-and-accumulation-system-ects_en
- Borgwardt, K.M., & Kriegel, H.P. (2005). Shortest-path kernels on graphs. In: *Fifth IEEE international conference on data mining (ICDM'05)* (pp. 8–pp). IEEE
- Bowen, W. (2015). The log-on degree; technology and universities. <https://www.economist.com/united-states/2015/03/12/the-log-on-degree>. Accessed 05/2015
- Daempfle, P. A. (2003). An analysis of the high attrition rates among first year college science, math, and engineering majors. *Journal of College Student Retention: Research, Theory & Practice*, 5(1), 37–52.
- Davis, D., Chen, G., Hauff, C., & Houben, G.J. (2016). Gauging mooc learners' adherence to the designed learning path. *International Educational Data Mining Society*
- Dehmer, M., Chen, Z., Emmert-Streib, F., Tripathi, S., Mowshowitz, A., Levitchi, A., et al. (2019). Measuring the complexity of directed graphs: A polynomial-based approach. *PLoS One*, 14(11), 1–19. <https://doi.org/10.1371/journal.pone.0223745>.
- Fincham, E., Gašević, D., Jovanović, J., & Pardo, A. (2019). From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies*, 12(1), 59–72. <https://doi.org/10.1109/TLT.2018.2823317>.
- Hlosta, M., Herrmannova, D., Vachova, L., Kuzilek, J., Zdrahal, Z., & Wolff, A. (2014). Modelling student online behaviour in a virtual learning environment. In *Machine Learning and Learning Analytics workshop at The 4th International Conference on Learning Analytics and Knowledge (LAK14)*, 24–28 Mar 2014, Indianapolis, Indiana, USA.
- Howard, E., Meehan, M., & Parnell, A. (2018). Contrasting prediction methods for early warning systems at undergraduate level. *The Internet and Higher Education*, 37, 66–75.
- Kabra R R, B.R.S. (2011). Performance prediction of engineering students using decision trees. *International Journal of Computer Applications* (0975 - 8887) Volume 36- No.11, December 2011, 36(11)
- Kennedy, G., Coffrin, C., de Barba, P., & Corrin, L. (2015). Predicting success: how learners' prior knowledge, skills and activities predict MOOC performance. *Proceedings of the fifth international conference on learning analytics and knowledge - LAK '15* (pp. 136–140). <https://doi.org/10.1145/2723576.2723593>.
- Kent, C., Boulton, C.A., Williams, H. (2017). Towards measurement of the relationship between student engagement and learning outcomes at a bricks-and-mortar University. *Joint Proceedings of the Sixth Multimodal Learning Analytics (MMLA) Workshop and the Second Cross-LAK Workshop*

- (MMLA-CrossLAK) (1828), 4–14. https://ore.exeter.ac.uk/repository/bitstream/handle/10871/28538/invited_paper_2.pdf?sequence=1&isAllowed=y%0Ahttp://ceur-ws.org/Vol-1828/#paper-02
- Kuzilek, J., Vaclavek, J., Fuglik, V., & Zdrahal, Z. (2018). Student drop-out modelling using virtual learning environment behaviour data. In: *European conference on technology enhanced learning* (pp. 166–171). Springer
- Marques, A., & Belo, O. (2011). Discovering student web usage profiles using markov chains. *Electronic Journal of e-Learning*, 9(1), 63–74.
- Matcha, W., Gašević, D., Uzir, N.A., Jovanović, J., & Pardo, A. (2019). Analytics of learning strategies: Associations with academic performance and feedback. In: *Proceedings of the 9th international conference on learning analytics & knowledge, LAK19* (pp. 461–470). Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3303772.3303787>.
- Mercer, J. (1909). Functions of positive and negative type, and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society, London*, 209, 415–446.
- Norris, J. R., & Norris, J. R. (1998). *Markov chains*. 2. Cambridge: Cambridge University Press.
- Okubo, F., Shimada, A., & Taniguchi, Y. (2017). A visualization system for predicting learning activities using state transition graphs. *14th International conference on cognition and exploratory learning in the digital age (CELDA 2017)* (Celda) (pp. 173–180)
- Pandey, M., & Sharma, V.K. (2013). A decision tree algorithm pertaining to the student performance analysis and prediction. *International Journal of Computer Applications*, 61(13)
- Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society*, 17(4), 49–64.
- Quinn, J. (2013). Drop-out and completion in higher education in Europe. *European Union*
- Romero, C., López, M. I., Luna, J. M., & Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. *Computers and Education*, 68, 458–472. <https://doi.org/10.1016/j.compedu.2013.06.009>.
- Samatova, N. F., Hendrix, W., Jenkins, J., Padmanabhan, K., & Chakraborty, A. (2013). *Practical graph mining with R*. Boca Raton: CRC Press.
- Sebesta, A. J., & Bray Speth, E. (2017). How should i study for the exam? self-regulated learning strategies and achievement in introductory biology. *CBE—Life Sciences Education*, 16(2), ar30.
- Sharabiani, A., Karim, F., Sharabiani, A., Atanasov, M., & Darabi, H. (2014). An enhanced bayesian network model for prediction of students' academic performance in engineering programs. In: *2014 IEEE global engineering education conference (EDUCON)* (pp. 832–837). <https://doi.org/10.1109/EDUCON.2014.6826192>. <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6826192>
- Shehata, S., & Arnold, K.E. (2015). Measuring student success using predictive engine. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15* (pp. 416–417). <https://doi.org/10.1145/2723576.2723661>.
- Van der Plank, P., et al. (2012). Effects of habsburg educational policies measured by census statistics. *Jezikoslovlje*, 13(2), 373–393.
- Winne, P. (2013). Learning strategies, study skills, and self-regulated learning in postsecondary education
- Zdrahal, Z., Hlosta, M., & Kuzilek, J. (2016). Analysing performance of first year engineering students. In: *Learning analytics and knowledge: Data literacy for learning analytics workshop*
- Zdrahal, Z., Hlosta, M., & Kuzilek, J. (2016). Analysing performance of first year engineering students. In: *Learning analytics and knowledge: Data literacy for learning analytics workshop*. <http://oro.open.ac.uk/58597/>
- Zimmerman, B.J. (1989). Models of self-regulated learning and academic achievement. In: *Self-regulated learning and academic achievement* (pp. 1–25). Springer
- Zinoyev, A., & Mirkes, E. (2013). Data complexity measured by principal graphs. *Computers & Mathematics with Applications*, 65(10), 1471–1482.

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