

Identificação de Trajetórias de Aprendizagem em um Curso de Graduação e sua relação com a Evasão Escolar

Érica Peters do Carmo, Isabela Gasparini¹, Elaine Harada Teixeira de Oliveira²

¹Depto de Ciência da Computação - Universidade do Estado de Santa Catarina (UDESC)

²Instituto de Computação - Universidade Federal do Amazonas (UFAM)

ericapetersc@gmail.com, isabela.gasparini@udesc.br, elaine@icomp.ufam.edu.br

Abstract. *School dropout is one of the main problems that affect higher education. To prevent dropout, students must fulfill a series of requirements presented in the form of courses. Thus, the path followed by students throughout their academic life can be represented by the sequence of courses taken, named learning path. This paper presents two approaches to modeling these paths in an undergraduate degree and uses them to investigate patterns related to dropout. The results show that the first semester's courses, with a high rate of failure, end up acting as a barriers to the advancement of students in the degree, leading them to abandon their studies.*

Resumo. *A evasão escolar é um dos principais problemas que afetam o ensino superior. Para que ela não ocorra, o estudante deve cumprir uma série de requisitos apresentados em forma de disciplinas. Assim, o caminho traçado pelos estudantes durante a sua vida acadêmica pode ser representado pela sequência de disciplinas cursadas, denominada trajetória de aprendizagem. Este trabalho apresenta duas abordagens para a modelagem dessas trajetórias em um curso de graduação, e as utiliza para investigar os padrões relacionados à evasão. Os resultados obtidos mostram que as disciplinas do primeiro semestre, com alto índice de reprovação, acabam atuando como barreiras para o avanço dos estudantes no curso, levando-os ao abandono dos estudos.*

1. Introdução

A evasão escolar é um problema que afeta diversos níveis de ensino de maneira globalizada [de Sousa et al. 2021]. Apesar de seu conceito estar diretamente relacionado com o contexto de ensino em que ocorre, de maneira abrangente, pode-se definir a evasão como o abandono voluntário dos estudos por parte do estudante antes da sua finalização.

No contexto do ensino superior, desde os primeiros estudos e modelos teóricos propostos na década de 70, entende-se que a evasão não deve ser vista como um fenômeno abrupto, e sim como um processo gradual acarretado por um conjunto de fatores acadêmicos, psicológicos e sociais [Bean and Metzner 1985, Spady 1970, Tinto 1975]. Ainda, os efeitos desse processo afetam tanto os estudantes quanto as instituições de ensino, tendo impacto em aspectos sociais, acadêmicos e econômicos [Santos Baggi and Lopes 2011].

Na perspectiva de um curso de graduação, entende-se que o oposto da ação de evadir é a obtenção do grau acadêmico. Para isso, o estudante deve completar uma série

de disciplinas que representam diferentes requisitos para a finalização do curso. Assim, o caminho percorrido por este estudante pode ser descrito através da sequência de disciplinas cursadas durante sua jornada acadêmica. A esse caminho dá-se o nome de trajetória de aprendizagem.

Ao observar um conjunto de estudantes inseridos em um mesmo curso de graduação, a análise das trajetórias de aprendizagem torna possível a descoberta dos padrões de comportamento dos estudantes no ambiente de ensino. A partir disso, diferentes trabalhos propõem aplicações para o uso dessas trajetórias, como a avaliação da eficiência curricular [Wigdahl et al. 2014], a análise do desempenho acadêmico [Almatrafi et al. 2016] e o entendimento dos padrões de transferências entre cursos [Wang 2016].

Nesse sentido, o presente trabalho tem como objetivo investigar a relação entre a evasão e as trajetórias de aprendizagem dos estudantes de graduação de um curso de Bacharelado em Ciência da Computação ofertado na Universidade do Estado de Santa Catarina (UDESC). Para isso, o trabalho apresenta a aplicação de duas abordagens distintas, estruturada e visual, para a modelagem das trajetórias. A modelagem estruturada, que é utilizada para representar a trajetória individual de cada estudante, é submetida a uma técnica de mineração de dados para a extração dos padrões de trajetórias mais frequentes. Já a modelagem visual, permite a representação agregada das trajetórias dos estudantes divididos em dois grupos, concluintes e evadidos.

Com base na aplicação das modelagens propostas nos dados do curso analisado, realizou-se o estudo dos padrões comportamentais dos estudantes, observando as disciplinas, o semestre letivo em que foram cursadas e os resultados obtidos em cada uma delas. Ao fim, o estudo buscou compreender quais padrões estão relacionados à evasão, destacando as diferenças entre os caminhos traçados pelos estudantes concluintes e os estudantes evadidos.

2. Trajetórias de Aprendizagem

As trajetórias de aprendizagem estão relacionadas aos caminhos percorridos pelos estudantes através de um conjunto de objetos e recursos de aprendizagem ao progredirem em direção a um objetivo educacional. Na literatura, essas trajetórias são definidas e abordadas de diferentes formas, estando sempre caracterizadas de acordo com o contexto educacional em que estão inseridas e os objetos educacionais que as compõem.

O conceito proposto por Clements e Sarama (2004) descreve a trajetória de aprendizagem como um objeto formado por três componentes: **(i)** o objetivo; **(ii)** a progressão do desenvolvimento (representação do conhecimento como uma rede de conceitos e habilidades interconectadas); e **(iii)** as tarefas instrucionais (tarefas voltadas para a aprendizagem de um conceito ou habilidade específica).

No contexto do ensino superior, esses componentes podem ser observados de forma que os estudantes percorrem níveis sucessivos da grade curricular (progressão do desenvolvimento), vencendo disciplinas pré-estabelecidas (tarefas instrucionais), para a obtenção do conhecimento e habilidades necessárias à obtenção do grau (objetivo) [Silva Garcia and Salcedo Gomes 2020].

Neste sentido, a análise das trajetórias de aprendizagem pode ser utilizada como

ferramenta para o desenvolvimento e otimização da grade curricular, e para o entendimento dos padrões comportamentais dos estudantes ao percorrerem essa grade. Ainda, faz-se possível a adição de outras dimensões a essa análise, como por exemplo, o desempenho acadêmico, a evasão escolar e a relação dos estudantes com seus pares. A partir disso, pode-se observar não só os caminhos percorridos pelos estudantes, como também as suas causas e o seu impacto.

3. Abordagens Propostas

A definição de trajetória de aprendizagem utilizada neste trabalho é a sequência de disciplinas percorridas por estudantes durante o curso de graduação em que estão matriculados. Com o intuito de permitir o entendimento e a visualização dessas trajetórias, propõem-se duas abordagens de representação distintas: (i) abordagem estruturada sequencialmente e (ii) abordagem visual através de grafos acíclicos.

3.1. Abordagem Estruturada

A abordagem estruturada busca representar a trajetória de aprendizagem individual dos estudantes. Para isso, observa-se cada um de maneira longitudinal, ou seja, de forma a mapear o caminho (i.e., sequência de disciplinas) percorrido por ele ao longo dos semestres letivos. Assim, define-se a seguinte modelagem para a trajetória de aprendizagem de um estudante: *um conjunto de eventos únicos (disciplinas cursadas) organizados sequencialmente de acordo com o momento em que ocorreram (semestre letivo)*.

De maneira simplificada, cada estudante pode ser representado por uma sequência $s = A, B, C, D...$ tal que A, B, C e D sejam as disciplinas cursadas pelo estudante e estejam ordenadas tanto pela ordem de semestres letivos quanto alfabeticamente. A sequência s_1 presente na Tabela 1 refere-se à trajetória de aprendizagem padrão de um estudante com base na grade curricular do curso analisado.

Apesar dessa sequência apresentar as disciplinas realizadas pelo estudante, ela não explicita a relação entre disciplinas e semestres. Em outras palavras, não é possível visualizar quais disciplinas foram cursadas juntas ou em momentos subsequentes. Assim, pode-se adicionar outra dimensão a cada evento da sequência, um número indicando o momento (semestre letivo) em que o evento (disciplina) ocorreu. A sequência s_2 na Tabela 1 apresenta a trajetória de aprendizagem padrão do curso com a adição das informações relacionadas aos semestres.

Nota-se ainda que a ordem em que os estudantes realizam as disciplinas está diretamente relacionada ao seu status no final de cada uma delas. Isto porque a grade curricular define uma relação de pré-requisitos entre disciplinas distintas, ou seja, se uma disciplina A é pré-requisito para a disciplina B , o estudante só pode cursar B após finalizar A com sucesso. Nesse sentido, para que as trajetórias sejam interpretadas assertivamente, torna-se importante identificar o status do estudante ao final de cada disciplina cursada.

Com base nisso, adiciona-se outra dimensão à representação das trajetórias, um código ao final de cada evento (i.e. disciplina) da sequência indicando o status do estudante no momento em que o evento chegou ao fim, sendo os códigos possíveis: (A) Aprovado, (N) Reprovado por Nota e (F) Reprovado por Frequência. A sequência s_3 presente na Tabela 1 exemplifica a trajetória de aprendizagem de um estudante que cursou todas as disciplinas com sucesso conforme a grade curricular do curso.

Tabela 1. Trajetória de aprendizagem padrão (conforme a grade curricular) representada pela abordagem estruturada

Elementos	Sequência Padrão
Apenas disciplinas	$s_1 = \{AGT, ALG1, CDII, LMA, MCI, TGS, ALG2, CDI2, EST, FCC, LPG, MDI, SNA, ANN, AOC, EDA, LFA, MEP, POO, PPR, AMS, BANI, COM, PRA, REC, SOFT, SOP, REC, BAN2, CAL, CGR, PAP, PES, SDI, TEC, ACT, EMI, MFO, IAR, IHC, OPI, OPII, OPIII, OPIV, PIM, TCCI, ATC007, ETI, OPV, OPVI, TCC2\}$
Disciplinas e semestre	$s_2 \{IAGT, IALG1, ICDII, ILMA, IMCI, ITGS, 2ALG2, 2CDI2, 2EST, 2FCC, 2LPG, 2MDI, 2SNA, 3ANN, 3AOC, 3EDA, 3LFA, 3MEP, 3POO, 3PPR, 4AMS, 4BANI, 4COM, 4PRA, 4REC, 4SOFT, 4SOP, 4REC, 5BAN2, 5CAL, 5CGR, 5PAP, 5PES, 5SDI, 5TEC, 6ACT, 6EMI, 6MFO, 6IAR, 6IHC, 6OPI, 7OPII, 7OPIII, 7OPIV, 7PIM, 7TCCI, 8ATC007, 8ETI, 8OPV, 8OPVI, 8TCC2\}$
Disciplinas, semestre e resultado	$s_3 \{IAGT-A, IALG1-A, ICDII-A, ILMA-A, IMCI-A, ITGS-A, 2ALG2-A, 2CDI2-A, 2EST-A, 2FCC-A, 2LPG-A, 2MDI-A, 2SNA-A, 3ANN-A, 3AOC-A, 3EDA-A, 3LFA-A, 3MEP-A, 3POO-A, 3PPR-A, 4AMS-A, 4BANI-A, 4COM-A, 4PRA-A, 4REC-A, 4SOFT-A, 4SOP-A, 4REC-A, 5BAN2-A, 5CAL-A, 5CGR-A, 5PAP-A, 5PES-A, 5SDI-A, 5TEC-A, 6ACT-A, 6EMI-A, 6MFO-A, 6IAR-A, 6IHC-A, 6OPI-A, 7OPII-A, 7OPIII-A, 7OPIV-A, 7PIM-A, 7TCCI-A, 8ATC007-A, 8ETI-A, 8OPV-A, 8OPVI-A, 8TCC2-A\}$

3.2. Abordagem Visual

A representação visual tem como objetivo principal permitir a visualização agregada das trajetórias de aprendizagem de diferentes estudantes inseridos em um mesmo curso. A análise das trajetórias de maneira conjunta torna-se importante quando é preciso comparar grupos distintos de estudantes (e.g. estudantes com alto e baixo desempenho, estudantes concluintes e evadidos, etc) de forma a encontrar comportamentos comuns a cada um desses grupos.

No trabalho de Ramos et al. (2015), os autores argumentam que a forma mais comum de representar trajetórias de aprendizagem visualmente é dada através de grafos, onde os nós representam os objetos educacionais com os quais os estudantes interagem, e as arestas representam as relações entre estes objetos. A partir disso, o presente trabalho apresenta dois modelos diferentes de grafos para representar as trajetórias dos estudantes do curso analisado.

Com base nos trabalhos de Almatrafi et al. (2016) e Ramos et al. (2017), o primeiro tipo de grafo apresenta em uma mesma estruturas as trajetórias agregadas dos estudantes do curso. A estrutura do grafo $G_1 = \langle N, A \rangle$, exemplificada na Figura 1a, é definida da seguinte forma:

- N : conjunto de nós $\{n_1, n_2, n_3, \dots, n_i\}$ em que cada nó $n_k \in N$ representa uma disciplina do curso e está associado a um peso $w(n_k)$ que indica o número de estudantes que cursaram a disciplina;
- A : conjunto de arestas $\{a_1, a_2, a_3, \dots, a_j\}$ em que cada aresta $a_k \in A$ que liga dois nós $n_1, n_2 \in N$ representa uma relação r entre as disciplinas n_1 e n_2 e tem peso $w(a_k)$ indicando a quantidade de estudantes distintos que cursaram ambas as disciplinas de acordo com a relação r . Essa relação pode ser representada de uma das seguintes formas:
 - Aresta não-direcionada: a aresta a_k que liga os nós n_1 e n_2 denota que as disciplinas referentes a estes nós foram cursadas em um mesmo semestre letivo;

- Aresta direcionada: a aresta a_k com origem em n_1 e destino em n_2 denota que as disciplinas referentes a esses nós foram cursadas em semestres subsequentes, sendo n_2 cursada no período imediatamente posterior ao período em que n_1 foi cursada. Ainda, observa-se um caso especial quando $n_1 = n_2$, ou seja, a aresta a_k tem origem e destino no mesmo nó. Isso ocorre quando um estudante cursa uma disciplina em um semestre, não é aprovado, e cursa a mesma disciplina no semestre seguinte.

Quanto à disposição visual, os nós dos grafos são organizados horizontalmente (da esquerda para a direita) de acordo com os semestres previstos para as disciplinas na grade curricular. Ou seja, se o nó n_1 está à esquerda do nó n_2 , a ementa do curso prevê que a disciplina representada por n_1 deve ser cursada antes da disciplina representada por n_2 . Ainda, cores distintas foram atribuídas a nós cujas disciplinas estão presentes na ementa em semestres distintos.

Os pesos dos nós e arestas são representados pelo seu tamanho e espessura, respectivamente. Assim, quanto maior o nó, maior o número de estudantes que cursaram a disciplina representado por ele, e quanto mais espessa a aresta que liga n_1 e n_2 , maior a quantidade de estudantes que cursaram ambas as disciplinas no mesmo semestre (aresta não direcionada) ou em semestres subsequentes (aresta direcionada).

O segundo tipo de grafo é utilizado para representar a organização das disciplinas de acordo com a grade curricular e o desempenho acadêmico dos estudantes em cada uma delas. A estrutura do grafo $G_2 = \langle N, A \rangle$, exemplificada na Figura 1b, é definida por:

- N : conjunto de nós $\{n_1, n_2, n_3, \dots, n_i\}$ em que cada nó $n_k \in N$ está associada a uma disciplina e é representado por um gráfico de setores dividido em três partes: porcentagem de aprovações (cor verde), porcentagem de reprovações por nota (cor azul) e porcentagem de reprovações por frequência (cor amarela). As porcentagens consideram a razão entre o total de vezes que os estudantes cursaram uma disciplina e o status dos estudantes ao fim desta.
- A : conjunto de arestas $\{a_1, a_2, a_3, \dots, a_j\}$ em que cada aresta $a_k \in A$ com origem em n_1 e destino em n_2 indica que n_1 é pré-requisito de n_2 .

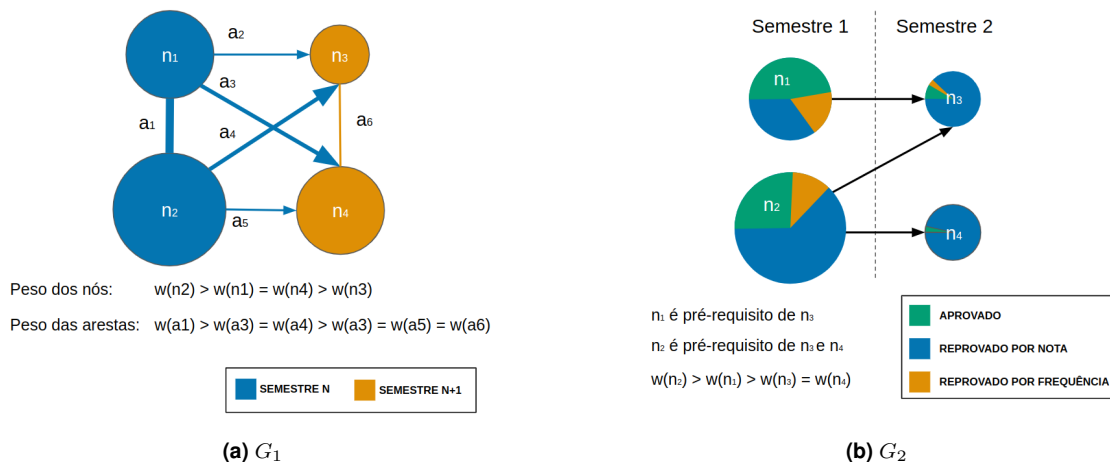


Figura 1. Exemplo da estrutura dos grafos G_1 e G_2

Diferentemente do primeiro tipo de grafo, os nós do grafo G_2 têm como peso o seu próprio grau de saída. Visualmente, isso significa que o tamanho de um nó está associado ao número de arestas com origem neste nó. A interpretação consequente disso é que, se um nó n_1 é maior que um nó n_2 , pode-se dizer que a disciplina representada por n_1 é pré-requisito de um número maior de disciplinas do que aquela representada por n_2 .

4. Aplicação

As abordagens propostas na seção anterior foram aplicadas nos dados dos estudantes do curso de Bacharelado em Ciência da Computação da Universidade do Estado de Santa Catarina (UDESC), tendo como objetivo principal a investigação da relação entre as trajetórias de aprendizagem dos estudantes e a evasão no curso.

Os dados utilizados foram devidamente anonimizados e disponibilizados em planilhas pela Secretaria de Ensino da universidade, de acordo com os critérios da Lei Geral de Proteção de Dados (LGPD). As planilhas estavam divididas da seguinte forma: **(i)** dados demográficos e de ingresso de cada estudante; **(ii)** créditos hora/aula cumpridos por estudante; e **(iii)** histórico das disciplinas contendo o status final de cada estudante.

Os registros presentes no conjunto de dados continham as informações acadêmicas de 635 estudantes, matriculados entre 2009 e 2020 no curso analisado. Contudo, devido às mudanças na grade curricular do curso, optou-se por selecionar apenas os ingressantes a partir do segundo semestre de 2012. Ainda, como o objetivo do estudo era a análise das trajetórias dos estudantes concluintes e evadidos, todos os estudantes ainda ativos no curso foram excluídos da base. O conjunto de dados final possuía o registro de 218 estudantes, sendo 50 estudantes concluintes e 168 estudantes evadidos.

Após a seleção dos dados, realizou-se o devido tratamento e modelagem de acordo com as abordagens propostas para a representação das trajetórias. As subseções a seguir apresentam o processo de desenvolvimento e os resultados obtidos em cada uma das abordagens.

4.1. Abordagem Estruturada

A partir dos dados de entrada, os registros acadêmicos dos estudantes foram transformados nas suas sequências individuais através de um *script* desenvolvido em Python, tendo como saída os arquivos de texto contendo as sequências de todos os estudantes, apenas dos estudantes evadidos e apenas dos estudantes concluintes.

Os arquivos gerados foram formatados para a importação no Weka [Hall et al. 2009], um software de código-aberto que possui a implementação de diversos algoritmos de aprendizado de máquina e mineração de dados. Com o objetivo de identificar os padrões de trajetória comuns a cada grupo de estudantes, concluintes e evadidos, o Weka foi utilizado para a execução do algoritmo GSP (*Generalized Sequential Pattern*) [Srikant and Agrawal 1996], uma evolução dos algoritmos de mineração de itens que considera a sua ordenação, e portanto, realiza a mineração de sequências.

Para o conjunto de sequências de cada grupo de estudantes, o GSP foi executado com o valor suporte de 0.7, de forma a identificar os padrões de trajetória comuns a 70% dos estudantes de cada um dos grupos. A Tabela 2 contém os maiores padrões de sequência comuns aos estudantes evadidos e concluintes. Optou-se por apresentar apenas

as maiores sequências (i.e. sequências com maior número de elementos) pois as demais são apenas subsequências destas.

Tabela 2. Maiores padrões de sequência comuns a cada grupo de estudantes

Dimensões	Grupo de Estudantes	Maior Sequência
Disciplina, semestre e status	Concluintes	{1AGT-A}{1LMA-A}{1MCI-A}{1TGS-A}{2LPG-A} (40)
	Evadidos	-
Disciplinas e semestre	Concluintes	{1AGT}{1ALG1}{1CDI1}{1LMA}{1MCI}{1TGS}{2LPG} (42)
	Evadidos	{1AGT}{1ALG1}{1CDI1}{1LMA}{1MCI}{1TGS} (138)
Disciplinas e status	Concluintes	-
	Evadidos	{ALG1-R}{CDI1-R} (139)

A sequência de disciplinas comuns aos estudantes concluintes composta pelas três dimensões (disciplina, semestre e status), apresentada na segunda linha da Tabela 2, refere-se ao conjunto de disciplinas do primeiro semestre letivo, com exceção da disciplina LPG (Linguagem de Programação) que é do segundo semestre, sendo que todas essas disciplinas estão atreladas ao status de aprovação. A única disciplina do primeiro semestre que não aparece na sequência é CDI1 (Cálculo Diferencial e Integral I). Assim, conclui-se que pelo menos 70% dos estudantes concluintes cursam e são aprovados em 5 das 6 disciplinas do primeiro semestre.

Nota-se ainda que não foram encontrados padrões para os estudantes evadidos quando observam-se as sequências formadas pelas três dimensões. Isso pode estar relacionado a dois fatores distintos. O primeiro é a variação no status final dos estudantes evadidos, ou seja, apesar de cursarem as mesmas disciplinas, eles obtêm resultados distintos o que impede a extração de padrões comuns. O segundo fator é o momento em que cada disciplina é cursada, em outras palavras, a falta de padrões pode estar atrelada à uma grande diferença na ordem em que os estudantes evadidos cursam cada disciplina.

Para a investigação de ambos os fatores mencionados, optou-se por rodar novamente o algoritmo GSP olhando para as sequências formadas por apenas duas dimensões: **i)** disciplinas e semestres; e **ii)** disciplinas e status.

A extração das sequências formadas pelas disciplinas e os semestres em que foram cursadas, na linha 3 da Tabela 2, retornaram resultados semelhantes para os grupos de estudantes evadidos e concluintes. A maior sequência de ambos os grupos indicam que os estudantes cursam no primeiro semestre as 6 disciplinas propostas pela grade curricular do curso para essa fase. Esse é um resultado esperado já que essas são as disciplinas comuns a todos os estudantes que ingressam no curso, sendo exceção apenas aqueles que fizeram a transferência de créditos. Contudo, como as disciplinas do primeiro semestre representam a maior sequência comum aos estudantes, pode-se concluir que a partir do fim do primeiro período letivo, os caminhos traçados pelos estudantes começam a divergir, independente do grupo que está sendo observado.

Em seguida, o algoritmo GSP foi executado para as sequências formadas pelas disciplinas e o status. Os resultados obtidos, representados na linha 4 da Tabela 2, mostram que a maior parte dos estudantes evadidos reprovam nas disciplinas ALG1 (Álgebra

Linear e Geometria Analítica I) e CDII em algum momento da sua jornada acadêmica. Ambas as disciplinas pertencem ao primeiro semestre, conforme a grade curricular proposta, e são pré-requisitos para diversas outras disciplinas ao longo do curso. Logo, reprovar em qualquer uma delas acaba impedindo o estudante de avançar no curso, já que ele não consegue se matricular em muitas das disciplinas dos semestres seguintes.

Por fim, a ferramenta utilizada não foi capaz de computar os padrões de sequência dos estudantes concluintes formados pelas disciplinas e os status com o valor suporte de 0,7. Assim, uma nova execução foi feita aumentando o valor para 0,8. Nesta execução foram obtidos 84962 padrões sequenciais. O motivo para o grande número de padrões está ligado ao fato dos estudantes concluintes terem cursado e sido aprovados em todas as disciplinas obrigatórias do curso. Logo, os padrões gerados são todas as combinações possíveis dessas disciplinas, e por este motivo, estão omitidas desta análise.

4.2. Abordagem Visual

Da mesma forma que a transformação de dados feita para a abordagem estruturada, desenvolveu-se um *script* em Python para transformar os registros acadêmicos fornecidos pela universidade em arquivos *csv* contendo a estrutura dos grafos para cada grupo de estudantes, concluintes e evadidos. Os arquivos gerados foram importados para o Gephi [Bastian et al. 2009], um software de código-aberto para a visualização e manipulação de grafos.

O grafo gerado para o conjunto de trajetórias dos estudantes evadidos, representado na Figura 2, mostra uma grande dispersão nas arestas, indicando o fato dos estudantes cursarem as disciplinas em diferentes momentos da sua vida acadêmica. Ainda, nota-se como o tamanho dos nós diminui da esquerda para a direita. Isto mostra como grande parte dos estudantes evade nos semestres iniciais, cursando apenas algumas poucas disciplinas além das 6 que pertencem ao primeiro semestre. É possível observar também que a maior parte dos estudantes evadidos cursou apenas essas 6 disciplinas, ou seja, acabaram evadindo do curso antes do início do segundo período letivo.

Deve-se olhar com atenção para as disciplinas não pertencentes ao primeiro semestre mas que possuem uma quantidade maior de estudantes quando comparadas com as demais. Entre elas estão LPG (Linguagem de Programação), SNA (Sistemas de

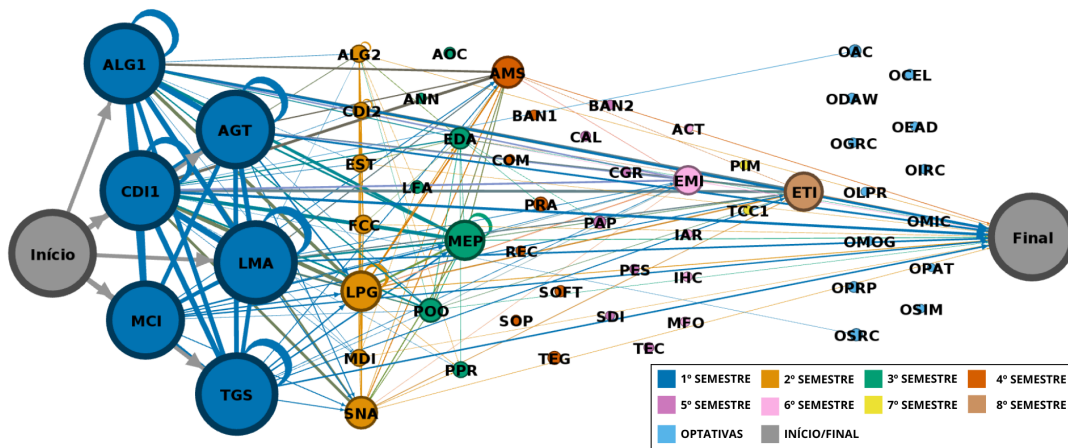


Figura 2. Grafo G_1 dos estudantes evadidos.

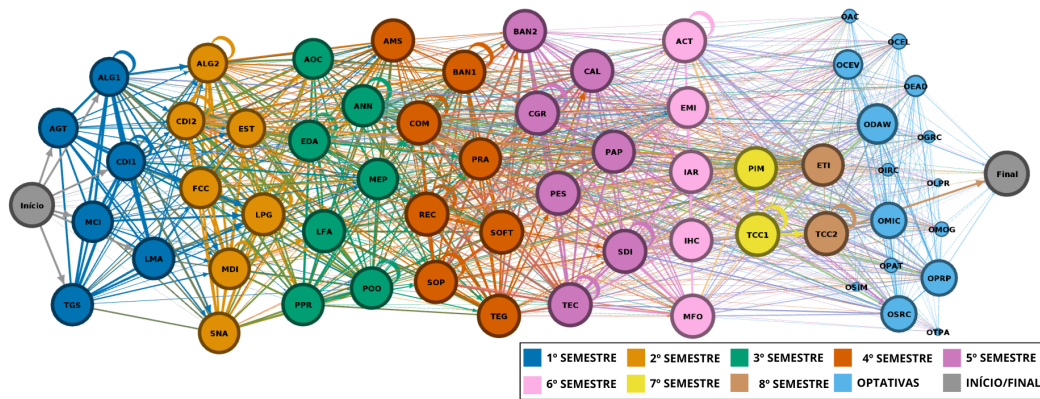


Figura 3. Grafo G_1 dos estudantes concluintes

Numeração e Álgebra de Boole), MEP (Metodologia da Pesquisa), AMS (Análise e Modelagem de Sistemas), EMI (Empreendedorismo em Informática) e ETI (Ética em Informática). O fator comum a essas disciplinas é que elas possuem poucas ou nenhuma outra disciplina agindo como pré-requisito. O significado disso fica mais claro quando observamos as arestas cíclicas nas disciplinas do primeiro semestre. Entende-se, então, que ao reprovarem em disciplinas no início do curso, os estudantes acabam não sendo capazes de avançar devido ao fato de muitas dessas disciplinas serem pré-requisitos para as demais. Assim, enquanto cursam as disciplinas reprovadas novamente, esses estudantes acabam se matriculando nas demais disciplinas que não exigem pré-requisitos.

Ao comparar o grafo das trajetórias dos estudantes evadidos com o grafo das trajetórias dos estudantes concluintes, presente na Figura 3, as diferenças entre as trajetórias de ambos os grupos de estudantes fica mais evidente. Sobre este segundo grafo, nota-se que a maior parte dos nós possui o mesmo tamanho. Isso porque são nós referentes às disciplinas obrigatórias, ou seja, disciplinas que foram cursadas pelo menos uma vez por todos os estudantes concluintes. No entanto, apesar das trajetórias desses estudantes serem formadas em grande parte pelo mesmo conjunto de disciplinas, estas são cursadas em momentos distintos da sua jornada acadêmica, o que pode ser observado pela dispersão das arestas.

A análise final tem como base o grafo G_2 , ilustrado na Figura 4, que representa o status dos estudantes de ambos os grupos em cada disciplina cursada e a maneira com que estas estão organizadas na grade curricular. É a partir desse grafo que se pode concluir o principal fator relacionado à evasão. As disciplinas que agem como pré-requisito para a maior parte das demais disciplinas da grade (i.e. disciplinas representadas pelos nós de maior peso) possuem alta taxa de reprovação, tanto por nota quanto por frequência. Ainda, essas disciplinas estão localizadas principalmente no primeiro e segundo semestres.

O efeito prático disso é que, devido ao alto nível de dificuldade atrelado às taxas de reprovação, essas disciplinas acabam agindo como barreiras para alguns estudantes, impedindo-os de avançar no curso e gerando um possível desengajamento que culmina na decisão de evadir. Observando, por exemplo, as disciplinas ALG1 e CDI1, percebe-se que o total de vezes em que os estudantes reprovaram por nota e frequência supera o total de vezes em que os estudantes foram aprovados. Isso reflete o padrão sequencial apresentado anteriormente na Tabela 2.

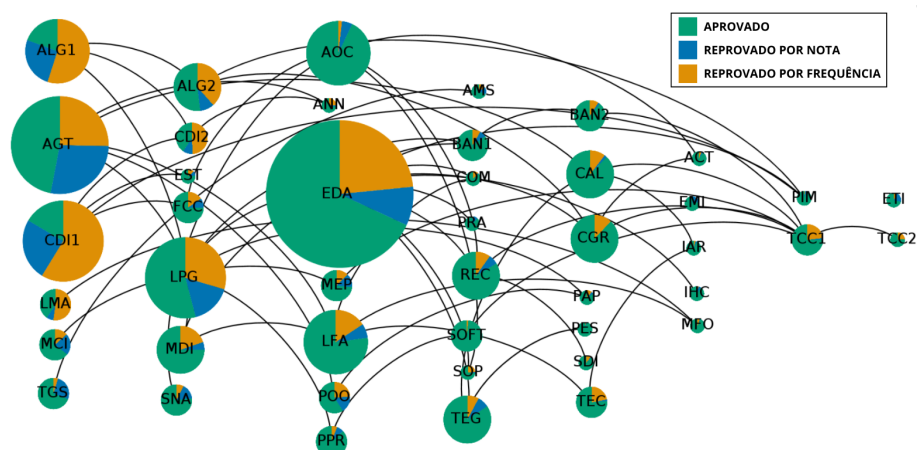


Figura 4. Grafo G_2 dos estudantes evadidos e concluintes

5. Considerações Finais

O objetivo principal do presente trabalho foi a análise da relação entre as trajetórias de aprendizagem e a evasão dos estudantes em cursos de nível superior. Para isto, duas abordagens de representação das trajetórias foram propostas, e posteriormente, aplicadas nos dados dos estudantes de um curso de Bacharelado em Ciência da Computação. Uma vez modelados os dados, realizou-se a mineração de padrões comuns às trajetórias de aprendizagem, bem como a comparação visual destas, para os estudantes evadidos e concluintes.

Entre os resultados obtidos, destacam-se duas principais descobertas em relação aos caminhos traçados pelos estudantes. A primeira diz respeito ao fato de ambos os grupos percorrem as disciplinas de maneira diversa, não se atendo à grade curricular proposta. Já a segunda mostra que grande parte dos estudantes evade ao fim do primeiro semestre, após reprovarem em algumas das disciplinas iniciais. Os estudantes que decidem continuar mesmo após as reprovações têm dificuldade em avançar no curso já que as disciplinas iniciais são pré-requisito para as demais.

Neste contexto, ressalta-se a necessidade das instituições de ensino pensarem em iniciativas capazes de mitigar a evasão e que tenham como foco principal os primeiros semestres da jornada acadêmica dos estudantes. Entre as iniciativas sugeridas estão o reforço nas disciplinas da primeira fase e a revisão da grade curricular para verificar se os pré-requisitos de cada disciplina estão condizentes com a sua ementa.

Por fim, sugere-se como trabalho futuro a análise da evasão com base em trajetórias de aprendizagem que contenham outras dimensões relacionadas aos estudantes, como por exemplo, informações socioeconômicas e demográficas. Isto porque entende-se que o fenômeno da evasão está relacionado não só ao ambiente de ensino, mas também aos diferentes aspectos que envolvem o contexto individual dos estudantes.

Agradecimentos

Agradecemos o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) 308513/2020-7 e 308395/2020-4 e a FAPESC Edital nº027/2020 TO nº2021TR795.

Referências

- Almatrafi, O., Johri, A., Rangwala, H., e Lester, J. (2016). Identifying course trajectories of high achieving engineering students through data analytics. In *2016 ASEE Annual Conference & Exposition*, Nova Orleans, Louisiana. ASEE Conferences.
- Bastian, M., Heymann, S., e Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks.
- Bean, J. P. e Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research*, 55(4):485–540.
- Clements, D. H. e Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning*, 6(2):81–89.
- de Sousa, L. R., de Carvalho, V. O., Penteado, B. E., e Affonso, F. J. (2021). A systematic mapping on the use of data mining for the face-to-face school dropout problem.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., e Witten, I. H. (2009). The weka data mining software: An update. *SIGKDD Explorations*, 11(1).
- Ramos, D., de Oliveira, E. H. T., Monteverde, I., e Oliveira, K. (2015). Trilhas de aprendizagem em ambientes virtuais de ensino-aprendizagem: Uma revisão sistemática da literatura. *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação - SBIE)*, 26(1).
- Ramos, D. B., Ramos, I. M. M., do Nascimento, P. B., de Souza Amaral, G., e de Oliveira, E. H. T. (2017). Um modelo para trilhas de aprendizagem em um ambiente virtual de aprendizagem. In *Simpósio Brasileiro de Informática na Educação* 1407–1416.
- Santos Baggi, C. A. D. e Lopes, D. A. (2011). Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 16:355 – 374.
- Silva Garcia, L. M. L. d. e Salcedo Gomes, R. (2020). Visualização e análise da trajetória de aprendizagem realizada no currículo no ensino superior. *Simpósio Brasileiro de Informática na Educação (SBIE 2020)* 1593–1602.
- Spady, W. G. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, 1:64 – 85.
- Srikant, R. e Agrawal, R. (1996). Mining sequential patterns: Generalizations and performance improvements. In *Advances in Database Technology — EDBT '96* 1–17. Springer Berlin Heidelberg.
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1):89–125.
- Wang, X. (2016). Course-taking patterns of community college students beginning in stem: Using data mining techniques to reveal viable stem transfer pathways. *Research in Higher Education*, (57):544–569.
- Wigdahl, J., Heileman, G., Slim, A., e Abdallah, C. (2014). Curricular efficiency: What role does it play in student success? *ASEE Annual Conference and Exposition, Conference Proceedings*.