A Study on Curriculum Planning and Its Relationship with Graduation GPA and Time To Degree

Sara Morsy
University of Minnesota
morsy@cs.umn.edu

George Karypis
University of Minnesota
karypis@cs.umn.edu

ABSTRACT

In recent years, several data-driven methods have been developed to help undergraduate students during course selection and sequencing. These methods tend to utilize the whole set of past course registration data, regardless of the past students' graduation GPA and time to degree (TTD). Though some previous work has shown through the results of their developed models that students of different GPA tend to take courses in different sequence, the actual analysis of the degree plans and how/if they relate to the students' graduation GPA and time-to-degree has not received much attention. This study analyzes how the student's academic level when they take different courses, as well as the pairwise degree similarity between pairs of students relate to the students' graduation GPA and TTD. Our study uses a large-scale dataset that contains 25 majors from different colleges at the University of Minnesota and spans 16 years. The analysis shows that TTD is highly correlated with both the timing and ordering of courses that students follow in their degree plans, while the correlation between graduation GPA and the course timing and ordering is not as high. We also perform a case study that uses course timing and ordering features to predict whether the student at each semester will graduate on-time or over-time. The results show that careful curriculum planning is needed to improve graduation rates in universities.

CCS CONCEPTS

• **Information systems** → **Data mining**; • **Applied computing** → **Education**;

KEYWORDS

undergraduate education, degree planning, course sequencing, course timing, curriculum planning, degree similarity, academic performance, GPA, time to degree, time to degree prediction

LAK19, March 4–8, 2019, Tempe, AZ, USA

ACM ISBN 978-1-4503-6256-6/19/03. . . $15.00
https://doi.org/10.1145/3303772.3303783

ACM Reference Format:
Sara Morsy and George Karypis. 2019. A Study on Curriculum Planning and Its Relationship with Graduation GPA and Time To Degree. In *The 9th International Learning Analytics & Knowledge Conference (LAK19), March 4–8, 2019, Tempe, AZ, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3303772.3303783

1 INTRODUCTION AND BACKGROUND

Many academic degree programs offer flexible plans, each containing a set of required (core) courses that need to be taken by all students, as well as a set of elective or liberal art courses, from which each student can choose a subset to take that interests them. Students often seek advice regarding their course selection and curriculum planning process from their advisers as well as fellow students. They end up having different degree plans based on the information they get from their limited sources, which may or may not be optimal for their academic success and on-time graduation.

With the increasing amount of available data about undergraduate students and their registration information, researchers have been able to develop machine learning and data-driven methods to further help students with their course selection process. These methods are based on: association rule mining [3], student-based collaborative filtering [7], group popularity ranking [7], content-based recommendation [8], and matrix factorization [4, 7]. Other methods focused on recommending the whole sequence of courses that satisfy the degree requirements [10–12]. These previous methods train their models on all of the past students' registration data, regardless of their graduation Grade Point Average [1] (GPA) and Time To Degree (TTD). A few other studies that developed course recommendation methods have shown through the analysis of their developed methods' results that different GPA-based groups of students tend to follow different sequencing for courses. For instance, Cucuringu *et al* [6] applied multiple rank aggregation methods, such as PageRank and SVD-Rank, on Math major students at their university to obtain global course sequences that are most consistent with the given data. Their results showed that different GPA-based groups of students tend to follow different course sequencing.

To illustrate the differences in the students' academic outcomes, we plotted the distribution of the students' graduation GPA and TTD at the University of Minnesota, which is shown in Figure 1. As shown in the figure, there is a large variability in the graduation GPA and TTD of students, with

[1] https://www.edglossary.org/grade-point-average/

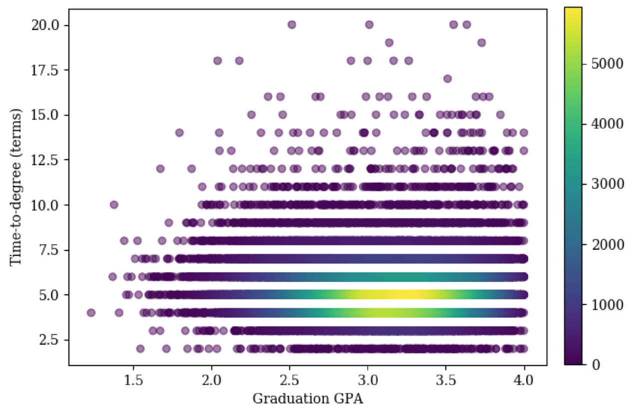

Figure 1: Distribution of graduation GPA vs time to degree across 25 different majors (see Section 2.1).

graduation GPAs in the range [1.8, 4.0], and TTD in the range and [3, 17] terms. This suggests that not all previous degree plans should be treated equally for learning good registration patterns. Despite this large variability in the students' graduation GPA and TTD, analyzing the actual students' degree plans and how/if they relate to their academic performance and TTD has received limited attention. We believe that this analysis should provide good insights about how to best utilize the past degree plans to help promote academic success for current and future students.

In this study, we provide an analysis of the degree plans taken by past students and their relationship with their academic performance in terms of graduation GPA and TTD. We try to answer the two following research questions:

RQ1. How does the timing of taking courses with respect to the student's academic level relate to their GPA and TTD?

RQ2. How does the pairwise degree similarity between pairs of students relate to the similarity in their GPA and TTD?

We use a large-scale dataset that consists of 25 majors that have the highest population of degrees granted from different colleges at the University of Minnesota that spans 16 years. Our analysis shows that: (i) low TTD students tend to take more courses ahead of time, and follow more similar sequencing for the common courses (especially in their later years), than high TTD students; and (ii) low GPA students tend to take more courses ahead of time, and follow more diverse sequencing for the common courses, than high GPA students.

Based on the results of our analysis, it is important for data-driven approaches that utilize student's degree plans, such as course recommendation, course sequence recommendation, and curriculum designing, to: (i) take the graduation GPA and TTD into account when training their models on students' degree plans; (ii) consider the student's academic level when recommending to them a set of courses, and making sure the courses are well-aligned with their academic level; and (iii) account for the student's expected grades and TTD in each course that they recommend. We believe that this can further improve the performance of these methods, especially for marginalized students who struggle with course selection and curriculum planning, and help them towards better academic performance and successful graduation.

Table 1: Dataset statistics.

Major	Students	Courses	Grades
Accounting	848	882	39,996
Art	740	1,461	27,132
Biology	1,311	1,399	53,885
Business & Marketing	738	1,061	33,259
Chemical Engineering	753	742	36,004
Civil Engineering	785	727	33,186
Communication Studies	1,919	2,041	76,504
Computer Science	993	1,011	43,593
Economics	914	1,248	39,059
Electrical Engineering	884	697	37,370
Elementary Education	932	903	37,783
English	1,564	2,144	59,462
Family Social Science	841	976	30,788
Finance	1,194	1,104	56,547
Global Studies	966	1,844	35,942
History	1,055	1,867	40,508
Journalism	2,467	2,256	104,757
Kinesiology	1,117	1,100	57,086
Marketing	1,179	1,157	52,365
Mechanical Engr.	1,266	820	52,786
Nursing	785	794	39,875
Political Science	2,046	2,400	76,296
Psychology	2,688	2,578	104,206
Soc of Law Criminol Devianc	727	1,266	28,253
Spanish Studies	789	1,710	34,365
Total	29,501	34,188	1,231,007

We also perform a case study that tries to predict whether the student at each semester will graduate on-time or over-time, by using features related to course timing and ordering as pursued by that student. TTD prediction has been explored in several previous studies [1, 2, 5, 9, 15], where they used features about student's demographic information, family background, financial aid, on- and off-campus work and experiences, as well as course grades and credit hours. We train several binary classification models using the proposed course timing and ordering features and show that curriculum planning is also a good indicator for TTD prediction.

2 ANALYSIS OF DEGREE PLANNING

2.1 Data Extraction and Pre-processing

The data used in our study was obtained from the University of Minnesota, where it spans a period of 16 years (Fall 2002 to Spring 2017). We extracted the set of students who have already completed their degrees on or before Spring 2017. We selected the degree programs that have at least 1,000 graduated students, which accounted for 25 majors from different colleges. Since our study focuses on the timing of courses and their ordering, we focused our study on full-time students and filtered out students who have been enrolled on a part-time basis for more than two terms. In addition, we removed rare courses that were taken by less than 20 students. The statistics for the final dataset used in our analysis is shown in Table 1.

We define time-to-degree (TTD) as the actual number of Fall and Spring terms taken by the student, divided by two. Since the number of students who transferred credits from other institutions or transferred credits from high school constitutes about two thirds of all students on average over all majors, we included them in our analysis by computing their TTD as the sum of their TTD at the University of Minnesota and the estimated number of terms for taking the transfer credits, which we refer to as transfer terms. We estimated the number of transfer terms as follows. For each student that have transferred credits from another college or from high school, let c and x be the number of transfer credits and the maximum number of credits taken by that student in the Fall or Spring terms, respectively. The number of transfer terms is then estimated by dividing c by x.

2.2 Data Analysis

Our two research questions focus on studying how the student's academic level when they take their courses as well as the pairwise degree similarity between pairs of students relate to their graduation GPA and TTD. To address these two questions, we define two sets of metrics: course timing metrics (Section 2.2.1), and degree similarity metrics (Section 2.2.2).

2.2.1 Course Timing Metrics. In each department, courses can be taken by students of different academic levels, e.g., freshman or sophomore. Previous studies, such as [14], showed that the student academic level plays an important role in accurately predicting their grades in a future course, since students of the same academic level tend to have similar academic maturity, experience, and knowledge. Based on that, we assume that each course needs to be taken in its corresponding level, which is based on the majority population of students of the same major who previously took this course. To address our first research question, which focuses on the timing of courses and how it relates to the student's academic performance, we measure the difference between the student's academic level when they took a course and the course's derived academic level. We also measure the difference between the academic level of pairs of students who took the same course. Let $\text{slevel}(s_i, x)$ be a function that returns the classification code for student s_i (1 for freshman, 2 for sophomore, 3 for juniors and 4 for seniors), when they took course x. And let $\text{clevel}(x)$ be a function that returns the derived level for course x that belongs to a specific major, which we compute as the majority student population's level that belong to that major when they took course x. For instance, for a course CSCI 541, if the overall distribution of the students' academic levels when they took it is: 60% seniors, 30% juniors, and 10% sophomores, then clevel(CSCI541) will return the classification code for seniors, which is 4[2]. Note that we only considered courses whose majority population is at least 60% of their whole population. We define two different metrics for computing course timing as taken by students, as follows:

(1) **Student-to-course Absolute Level Difference:** Given a course x taken by student s_i, we compute the absolute deviation of s's academic level when they took x from x's academic level as:

$$\text{diff}(s_i, x) = |\text{clevel}(x) - \text{slevel}(s_i, x)|, \tag{1}$$

We will refer to this metric as *Student-to-course Absolute Level Difference*. This metric gives a value in the range [0, 3]. Computing the average of $\text{dist}(s_i, x)$ over different student groups tells us how often students in each group take courses at their right academic level, where the smaller the average value is, the more courses that students take at a closer course academic level to their academic level.

(2) **Student-to-course Signed Level Difference:** Eq. 1 measures the absolute deviation of the student's academic level to the course's derived level, but it does not take into consideration the sign of that deviation. Since a student can take a course ahead or behind its derived level, we need another metric that considers this difference. This will show when different students tend to take their courses. We thus define our next metric, which we will refer to as *Student-to-course Signed Level Difference*, and is computed as:

$$\text{diff}(s_i, x) = \text{clevel}(x) - \text{slevel}(s_i, x), \tag{2}$$

This metric gives a value in the range [-3, 3]. Computing the average of this metric over all the courses taken by a student can tell us how often that student tends to take courses at different derived course level from their academic level when taking these courses. The higher the negative direction of this average value, the more lower-level courses the student took , while the higher the positive direction of this average value, the more higher-level courses the student took.

2.2.2 Degree Similarity Metrics. Our second research question focuses on the pairwise degree similarity between pairs of students and it relates to the similarity in their graduation GPA and TTD. To address this research question, we define three different metrics that compute the similarity between a pair of degree plans, as follows.

(1) **Student-to-student Course Time Difference:** For each pair of students, we compute the academic level difference when they took the common courses. We will refer to this metric as the *Student-to-student Course Time Difference*, and we compute it as:

$$\text{diff}(s_1, s_2, x) = |\text{slevel}(s_1, x) - \text{slevel}(s_2, x)|, \tag{3}$$

The average of Student-to-student Course Time Difference over all the common courses taken by a pair of students will be low for pairs of students who take the common courses at similar academic levels, and will be high otherwise.

(2) **Bag Similarity:** The similarity between two degree plans with respect to the set of courses taken in both of them can be measured by using the Jaccard similarity

[2]Though this is a simple way to define the course's academic level, it serves as a good starting metric. We plan to investigate other ways of deriving the course's academic level more efficiently in the future.

coefficient between them, which we will refer to as the bag similarity, and is computed as:

$$\text{sim}(d_1, d_2) = \frac{|C_1 \cap C_2|}{|C_1 \cup C_2|}, \tag{4}$$

where: C_i is the set of courses taken in degree i. This metric gives us an overall idea about the percentage of courses that are taken in common in the two degree plans.

(3) **Sequence Similarity:** The bag similarity metric defined above cannot tell us any information about the ordering of common courses in a pair of degree plans, which can be an important factor for academic performance. Since each course provides specific knowledge that can be useful for performing well in another course, the ordering of courses can affect the student's grades as well as their TTD. Therefore, we define another metric that can tell us how the course sequencing in the two plans aligns with each other. We will refer to this metric as sequence similarity, which we compute as:

$$\text{sim}(d_1, d_2) = \frac{\sum_{(x,y)\in|C_1 \cap C_2|} T(t_{1,x} - t_{1,y}, t_{2,x} - t_{2,y})}{|C_1 \cap C_2|}, \tag{5}$$

where C_i is as defined in Eq. 4, and $t_{i,x}$ is the time, i.e., term number, that course x was taken in d_i, e.g., the first term is numbered 1, the second is numbered 2 and so forth. Note that since students can enroll in summer terms at our university (for one or two courses), we assign the term number for a summer term to half the value of the previous and following spring and fall terms, respectively. This is to ensure that students who enroll in any summer term have the same term numbers (relative to their entry term) as those who do not enroll in it. Function $T(dt_1, dt_2)$ is defined as:

$$T(dt_1, dt_2) = \begin{cases} 1, & \text{if } dt_1 = dt_2 = 0 \\ \exp\Big(-\lambda(|dt_1 - dt_2|)\Big), & \text{if } dt_1 \times dt_2 \geq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

where λ is an exponential decay constant[3]. Function T assigns a value of 1 for pairs of courses taken concurrently, i.e., during the same term, in both plans, and assigns a value of 0 for pairs of courses that are either: (i) taken in reversed order in both plans, or (ii) taken concurrently in one plan and sequentially in the other. For pairs of courses taken in the same order, it assigns a positive value that decays exponentially with $|dt_1 - dt_2|$. Our underlying assumption behind such an approach is that, when courses x and y are taken concurrently or in the same order with similar time difference in both d_1 and d_2, then we assume that this is a more similar ordering of both courses than when there is a larger time difference in both plans, and that a different ordering of x and y in the plans does not contribute to their similarity score.

Note that for all the above three pairwise degree similarity metrics, since the degree requirements and courses change from year to year at our university, we only consider pairs of students of the same cohort, i.e., those who entered college in the same term, when computing these metrics. Moreover, we computed each of the Student-to-student Course Time Difference and sequence similarity metrics for pairs of students who have taken at least 20% of their courses in common. We computed these metrics only for the majors where the number of each group of student pairs is $\geq$ 50 pairs. The statistics for the different pairs of students included in our analysis is shown in Table 1.

3 RESULTS

We present the results of our analysis for different groups of students, based on their graduation GPA and TTD. Since both variables are considered important for academic success, we study the effect of changing one variable while fixing the other to a specific range. For instance, we study the effect of the course timing metrics among students who have low and high TTD of ≤ 9 and ≥ 11 terms, respectively, by assuming that they have achieved a high GPA that is ≥ 3.0.

3.1 How does the timing of taking courses with respect to the student's academic level relate to their graduation GPA and TTD?

Figure 2 shows the box plots of the 25 majors in terms of the course timing metrics (defined in Section 2.2.1) among different GPA and TTD-based student groups. By comparing the different student groups in terms of Student-to-course Absolute Level Difference, we see that there is no significant difference among the low- and high-TTD-based groups (Fig. 2 (a)), while for the high- and low-GPA-based groups (Fig. 2 (c)), we see that high GPA students have lower Student-to-course Absolute Level Difference than low GPA ones.

By comparing the student groups in terms of Student-to-course Signed Level Difference, we see that, in Fig. 2 (b), low TTD students tend to take more courses ahead of time than high TTD students. On the other hand, Fig 2 (d) shows that low GPA students tend to take more courses ahead of time than high GPA students.

To see whether there is statistical significance in these results on a per-major basis, Table 2 shows a summary of the per-major results in terms of the average and standard deviation of the course timing metrics for each student group. It also shows the number of majors that has statistically significant results in one group over the other. These results show that Student-to-course Absolute Level Difference is not a significantly discriminating metric among the different groups of students, as it is statistically significant in 9 and 13 majors only, out of 25 majors, for the TTD- and GPA-based student groups, respectively. In terms of Student-to-course Signed Level Difference, the differences are statistically significant among high and low TTD-based student groups in

[3] In our analysis, we chose a small exponential decay constant $\lambda = \frac{1}{5}$ for a slow decay effect.

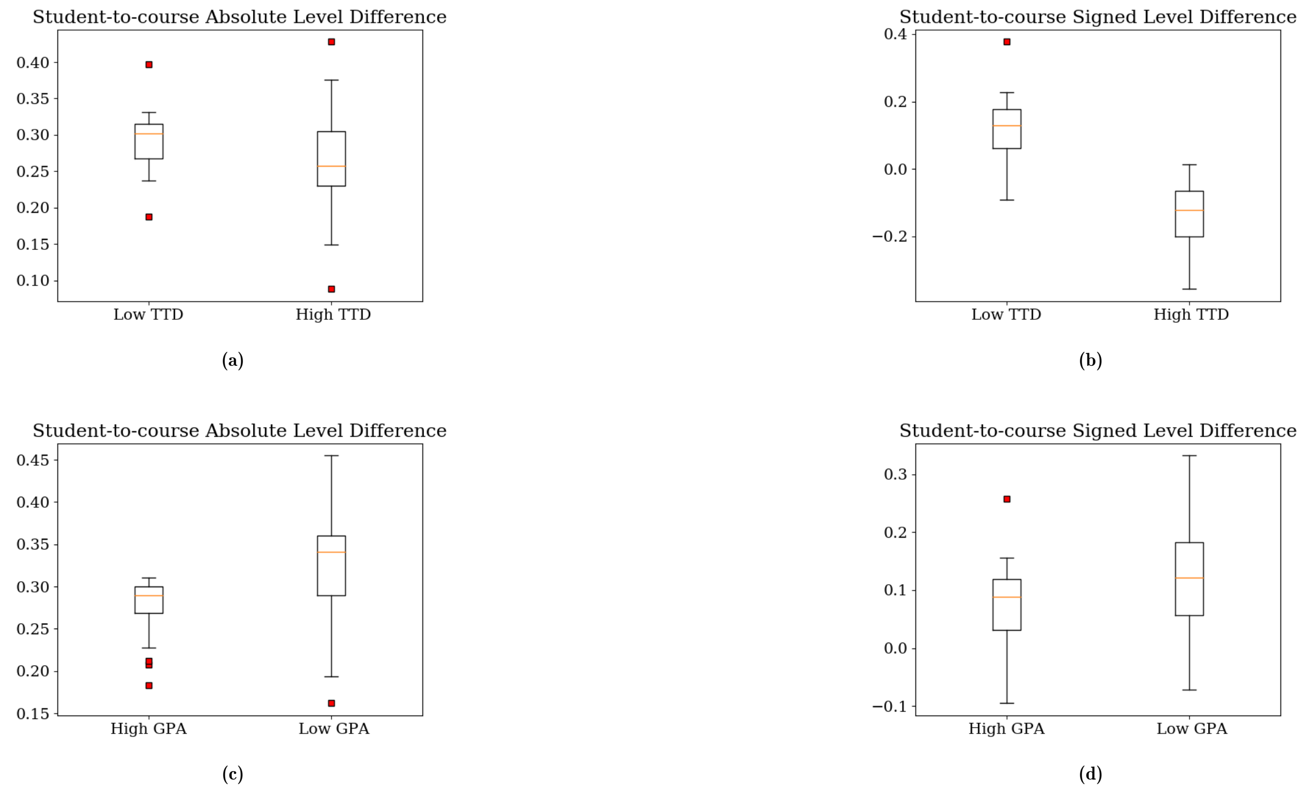

Figure 2: Course timing metrics among different groups of full-time students. TTD is shorthand for time-to-degree. Low and high time-to-degree is one that is ≤ 9 and ≥ 11 terms, respectively, both with GPA ≥ 3.0. High and low GPA is one that is ≥ 3.2 and ≤ 2.8, respectively, both with TTD ≤ 10 terms. The line inside the box denotes the median value. The ends of the whiskers denote the lowest datum still within 1.5 IQR (interquartile range) of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile, while the red squares denote outliers that are outside these ranges.

Table 2: Summary of the course timing metrics results among high and low GPA- and TTD-based student groups across all majors.

Metric	Mean	Std.	Mean	Std	Count(†)
	Low TTD		High TTD		Low vs High
Student-to-course Absolute Level Difference	0.293	0.040	0.266	0.080	9 (25)
Student-to-course Signed Level Difference	0.125	0.094	-0.136	0.101	24 (25)
	High GPA		Low GPA		High vs Low
Student-to-course Absolute Level Difference	0.275	0.035	0.331	0.068	13 (25)
Student-to-course Signed Level Difference	0.080	0.072	0.122	0.095	10 (25)

Low and high TTD denote the set of students with time-to-degree that are ≤ 9 and ≥ 11 terms, respectively, both with GPA ≥ 3.0. High and low GPAs denote the set of students with GPAs that are ≥ 3.2 and ≤ 2.8, respectively, both with TTD ≤ 10 terms. The columns "Mean" and "Std." denote the average and standard deviation of the per-major results of the corresponding student group. Count(†) denotes the number of majors that have statistically significant results between the two compared groups, using Welch's t-test with a significance level of 0.001, and the number between parentheses denote the total number of majors that qualified for the corresponding metric.

24 out of the 25 majors, but only statistically significant in 10 majors among high and low GPA-based groups. This shows that the timing of courses is highly correlated with time to degree, but is not a discriminating factor for the graduation GPA.

3.2 How does the pairwise degree similarity between pairs of students relate to the similarity in their graduation GPA and TTD?

Figure 3 shows the box plots of the 25 majors in terms of the pairwise degree similarity metrics (defined in Section 2.2.2) among different pairs of GPA and TTD-based student groups,

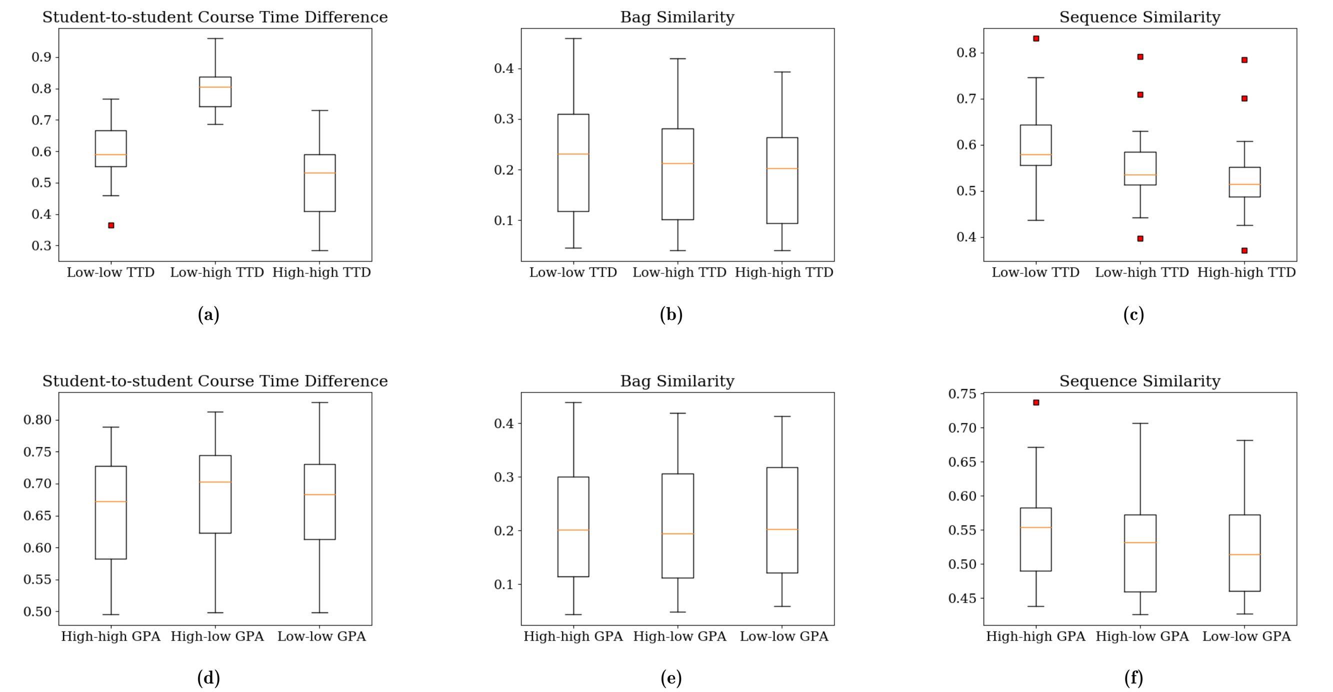

Figure 3: Degree similarity metrics among different groups of full-time students. TTD is shorthand for time-to-degree. Low and high time-to-degree is one that is ≤ 9 and ≥ 11 terms, respectively, both with GPA ≥ 3.0. High and low GPA is one that is ≥ 3.2 and ≤ 2.8, respectively, both with TTD ≤ 10 terms. The line inside the box denotes the median value. The ends of the whiskers denote the lowest datum still within 1.5 IQR (interquartile range) of the lower quartile, and the highest datum still within 1.5 IQR of the upper quartile, while the red squares denote outliers that are outside these ranges.

Table 3: Summary of the degree similarity metrics results among different pairs of GPA- and TTD-based student groups across all majors.

Metric	Mean	Std.	Mean	Std.	Mean	Std.	Count(†)		
	LL TTD		LH TTD		HH TTD		LL vs LH	LL vs HH	HH vs LH
Student-to-student Course Time Difference	0.595	0.102	0.801	0.081	0.506	0.128	18 (18)	13 (18)	18 (18)
Bag Similarity	0.220	0.124	0.198	0.114	0.190	0.110	25 (25)	23 (25)	16 (25)
Sequence Similarity	0.590	0.099	0.547	0.096	0.528	0.097	18 (18)	14 (18)	14 (18)
	HH GPA		LH GPA		LL GPA		HH vs LH	HH vs LL	LL vs LH
Student-to-student Course Time Difference	0.656	0.087	0.677	0.086	0.670	0.088	13 (24)	4 (24)	8 (24)
Bag Similarity	0.209	0.119	0.206	0.115	0.214	0.113	10 (24)	7 (24)	4 (24)
Sequence Similarity	0.550	0.073	0.530	0.070	0.521	0.065	20 (24)	14 (24)	13 (24)

LL, LH and HH denote the pairs of students where each pair belongs to the (low, low), (low, high) and (high, high) corresponding student groups, respectively. Low and high TTD denote the set of students with time-to-degree that are ≤ 9 and ≥ 11 terms, respectively, both with GPA ≥ 3.0. High and low GPAs denote the set of students with GPAs that are ≥ 3.2 and ≤ 2.8, respectively, both with TTD ≤ 10 terms. The columns "Mean" and "Std." denote the average and standard deviation of the per-major results of the corresponding student-pair group. Count(†) denotes the number of majors that have statistical significant results between the two compared groups, using Welch's t-test with a significance level of 0.001, and the number between parentheses denote the total number of majors that qualified for the corresponding metric (see Section 2.2.2).

while Table 3 shows a summary of the per-major results and the statistical significance between different groups of student pairs. By comparing the different student groups in terms of Student-to-student Course Time Difference in the box plots, we see that for the TTD-based groups (Figs. 3 (a)), the high TTD students tend to take their courses at a slightly more similar time together than the low TTD students, with median Student-to-student Course Time Difference value of 0.59 an 0.53, respectively. In addition, pairs of low-high TTD tend to take courses at even a more different timing (with a median Student-to-student Course Time Difference value of 0.8) than pairs of low-low and high-high TTD, aligning with their results of the Student-to-course Signed Level Difference metric in Section 3.1. The statistical significance results in Table 3 also confirm these differences among low and high TTD-based groups, where 13 out of the 18 qualifying majors have statistically significant Student-to-student Course Time

Difference in the pairs of low-low vs pairs of high-high TTD-based groups, while the Student-to-student Course Time Difference is statistically significant in all the 18 majors in each of the pairs of low-low and high-high TTD-based groups vs the pairs of low-high TTD-based groups.

On the other hand, there is not a clear distinction between high and low GPA students in their timing of taking courses among different majors (Fig. 3 (d)). Table 3 also shows that the average Student-to-student Course Time Difference falls in the range $[0.656, 0.677]$ with a standard deviation of ~ 0.087 among the different GPA-based pairs of students, which also aligns with the results of the course timing metrics in Section 3.1 that shows that the timing of courses is not discriminative among different GPA-based student groups.

By comparing the bag similarity among different TTD-based students, we see that low TTD students take more courses in common than high TTD students (average values among pairs of low-low and high-high TTD-based students of 0.22 and 0.19, respectively), with a statistically significant difference in 23 out of the 25 majors.

By looking at the sequence similarity, we see that among the different TTD-based students, low TTD students follow more similar ordering of the courses than high TTD students, with a statistically significant difference (p-value < 0.001) between the two groups in 14 out of the 18 qualifying majors, with an overall sequence similarity that is 0.062 higher in the former group across all majors. An interesting observation is that there is a larger diversity in the sequencing of courses taken by pairs of high TTD students (an average sequence similarity of 0.528) than among pairs of high-low TTD students (an average sequence similarity of 0.547). Along with the course timing results that showed that high TTD students tend to take courses later in time than low TTD students, this could be explained as the former group of students, though they achieve high grades in the courses they take, do not have enough information about their degree requirements. As a result, they end up fulfilling these requirements later in their study than when they should have been fulfilled.

Among different GPA-based students, high GPA students follow more similar sequencing of the courses than low GPA ones, with a difference in their sequence similarities that is statistically significant in 14 out of the 24 qualifying majors (see Section 2.2.2).

To further analyze the differences in the sequence similarity among students, we computed the sequence similarity among different student-pair groups based on their academic levels when they took their courses. At the University of Minnesota, the student is classified into one of four academic levels, based on the total number of credits completed: freshman (< 30 credits), sophomore (< 60 credits), junior (< 90 credits), and senior (≥ 90 credits). Freshmen and sophomores are classified as lower division students, while juniors and seniors are classified as upper division students. Table 4 shows the summary of these results, for different pairs of GPA- and TTD-based students. By comparing lower and upper division TTD-based students, we see that there is much greater difference in the different groups' similarities that belong to the upper division than those that belong to the lower division. This shows that students in their early years tend to take courses in a very similar ordering, regardless of their TTD (average sequence similarities of 0.918, 0.897 and 0.896 for low-low, low-high and high-high TTD-based pairs of students, respectively). In their later years, however, low TTD students continue to follow similar sequencing of their courses (with an average sequence similarity of 0.893), while high TTD students diverge from that sequencing and follow more diverse sequencing of their courses (with an average sequence similarity of 0.806).

Similar trends apply to the lower and upper division GPA-based student groups (Table 4), though the differences between the sequence similarities of the upper division groups are not as high (average sequence similarities of 0.881, 0.874 and 0.869 for high-high, high-low and low-low GPA-based pairs of students, respectively). This again shows that the sequence similarity is more discriminating for TTD than for GPA.

4 CASE STUDY: TTD PREDICTION

So far, we have analyzed the differences between different GPA- and TTD-based students with respect to the course timing and degree similarity metrics that we defined in Section 2.2. Here, we test whether the timing and ordering of courses as taken by the student at each semester can help predict whether he/she will graduate on-time or over-time. There has been a lot of research on TTD prediction and analyzing the possible effects behind over-time graduation [1, 2, 5, 9, 15]. Features like academic features, financial aid, off- and on-campus work and experience, family background, student's demographic information and high school grades have all been investigated and they were found to be good predictors for TTD. In this work, we build a classification model that uses course timing and ordering features to predict students who are at-risk of graduating over-time. We define a student to be at-risk of graduating over-time if he/she graduates in more than four years, i.e., more than nine Fall or Spring terms. We use academic features that have been previously used for TTD prediction as baseline features, to compare their performance against the newly proposed features.

4.1 Features

4.1.1 Academic (Baseline) Features. Similar to previous work [9], we use the following academic features that exist in our dataset:

(1) **General Experience:** We use the following features: initial status (new vs transfer student), number of program major changes, stop-out time since first enrollment, and number of summer enrollment terms.
(2) **Course Grades:** We use percentage of D or F grades, percentage of I (incomplete) or W (withdrawal) grades, individual course grades, and number of repeated courses.
(3) **Credit Hours:** We use the total credits accumulated, total transfer credits, percentage of earned to attempted credits, and average credit load per enrolled term.

Table 4: Summary of the sequence similarity results among different pairs of GPA- and TTD-based student groups, grouped by their academic division, across all majors.

Division Group	Mean	Std.	Mean	Std.	Mean	Std.	Count(†)		
	LL TTD		LH TTD		HH TTD		LL vs LH	LL vs HH	HH vs LH
Lower Division	0.918	0.024	0.897	0.023	0.896	0.023	17 (20)	11 (20)	8 (20)
Upper Division	0.893	0.024	0.837	0.029	0.806	0.035	25 (25)	23 (25)	23 (25)
	HH GPA		LH GPA		LL GPA		HH vs LH	HH vs LL	LL vs LH
Lower Division	0.916	0.019	0.901	0.026	0.893	0.032	24 (25)	22 (25)	18 (25)
Upper Division	0.881	0.024	0.874	0.023	0.869	0.019	17 (25)	18 (25)	15 (25)

Refer to Section 3.2 for the definition of division group. LL, LH and HH denote the pairs of students where each pair belongs to the (low, low), (low, high) and (high, high) corresponding student groups, respectively. Low and high TTD denote the set of students with time-to-degree that are ≤ 9 and ≥ 11 terms, respectively, both with GPA ≥ 3.0. High and low GPAs denote the set of students with GPAs that are ≥ 3.2 and ≤ 2.8, respectively, both with TTD ≤ 10 terms. The columns "Mean" and "Std." denote the average and standard deviation of the per-major results of the corresponding student-pair group. Count(†) denotes the number of majors that have statistical significant results between the two compared groups, using Welch's t-test with a significance level of 0.001, and the number between parentheses denote the total number of majors that qualified for the corresponding metric (see Section 2.2.2).

4.1.2 Course Timing and Ordering (New) Features. Based on the metrics defined in Section 2.2 that consider course timing and pairwise course ordering, we define the following features:

(1) **Course Timing:** For each course, we use the relative term number when the course is taken (starting from 1), and the academic level of the student when he/she took that course.
(2) **Pairwise Course Ordering:** For each pair of courses (c_1, c_2), we use the number of earned credits as well as the number of terms taken between the two terms when the student took c_1 and c_2. Note that a feature "c_1 : earned-credits : c_2" denotes the number of credits that the student earned after taking c_1 and before taking c_2, which is different from the feature "c_2 : earned-credits : c_1", and the same applies for the term difference based features.

4.2 Experimental Setup and Evaluation

We normalized each feature to L2 norm as a pre-processing step. We tested many classifiers (including logistic regression, SVM, Decision Tree, Random Forest, and Multi-layer Perceptrons (MLP)) using scikit-learn library in Python [13], and found MLP to be the best performing classifier. The data for each major was trained separately, with an average percentage of over-time graduating students of 54% with a standard deviation of 17%. We constructed different sets of the data, in order to predict whether the student, at each semester, could be at-risk of graduating over-time. We performed 10-fold cross-validation and we report the average results over the 10 folds averaged over all the 25 majors.

We evaluate the classifier's performance in terms of the following metrics:

- Recall of at-risk: Recall is the ratio of true positives to all actual positives.
- Precision of at-risk: Precision is the ratio of true positives to all predicted positives.
- F_1 of at-risk: F_1 score is the harmonic mean between Precision and Recall, which conveys the balance between the two and computed as:

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{7}$$

- Area under the receiver operating characteristic (ROC) curve (AUC): ROC curve plots the true positive rate against the false positive rate, at various thresholds. AUC corresponds to the probability that the classifier will rank a random positive instance higher than a negative one.

4.3 Experimental Results

Table 5 shows the TTD prediction results when using the academic and course timing and ordering features, by predicting TTD at each semester when the student is enrolled, starting from the second to the seventh semester. The results show that the prediction performance using the proposed course timing and ordering features is similar to that using the standard academic features. Using the course timing and ordering features tends to give more accurate F_1, precision and recall scores in the late years (semesters 5 though 7) than in early years (semesters 2 through 4). In terms of AUC, there are small insignificant differences in the prediction performance using both types of features. This shows that degree planning in terms of the timing of courses and the ordering between them plays an important role in the student's TTD, that is similarly equal to his/her general experience and academic performance in terms of grades and credit hours.

5 CONCLUSIONS AND IMPLICATIONS

In this study, we investigated how the student's academic level when they take different courses as well as the pairwise degree similarity between pairs of students relate to their graduation GPA and time to degree (TTD). Our analyses were conducted on a large-scale dataset that spans 16 years worth of degree plans pursued by students from 25 majors from different colleges at the University of Minnesota. Our findings indicated that:

- Student clusters that are based on their graduation GPA or TTD tend to share more similarities within themselves than with students from different clusters, in the time when they take their courses as well as the set and sequencing of them.
- Low TTD students tend to take courses ahead of time. In addition, they follow more similar sequencing for

Table 5: TTD prediction results using the academic (baseline) and new (course timing and ordering) features.

Metric	Feature Type	Semester Number					
		2	3	4	5	6	7
F1 of at-risk	Academic	0.420	0.424	0.445	0.405	0.435	0.417
	Course Timing and Ordering	0.388	0.419	0.407	0.424	0.463	0.418
Precision of at-risk	Academic	0.365	0.369	0.387	0.352	0.379	0.367
	Course Timing and Ordering	0.363	0.385	0.366	0.389	0.424	0.390
Recall of at-risk	Academic	0.528	0.542	0.557	0.514	0.551	0.525
	Course Timing and Ordering	0.478	0.522	0.519	0.546	0.575	0.520
AUC	Academic	0.550	0.549	0.548	0.548	0.547	0.548
	Course Timing and Ordering	0.549	0.545	0.544	0.546	0.545	0.544

Underlined entries denote the best performance across the two feature types in each semester.

the common courses, especially in their late years than high TTD students.

- Low GPA students tend to take courses ahead of time and follow more diverse sequencing of their courses than high GPA students.

Overall, there is a strong correlation between the timing and ordering of courses and the students' TTD. However, the correlation between them and the student's graduation GPA is not as strong. One potential explanation for this could be that, since each course provides a specific set of knowledge components that can be useful or required for other courses, there is an inherent sequencing among courses through which the students can accumulate their knowledge in a correct way and graduate on time. However, even when students follow the correct sequencing that guarantees on-time graduation, their grades in different courses can be affected by many other factors that can or cannot be measured. For instance, the student's effort in the course and how much time they allocate for learning its material and finishing its assignments and projects is hard to measure in the actual classroom setting. Another factor could be the student's learning style and how it aligns with the instructor's teaching style, the types of evaluation they do, as well as the grading system they follow. A third factor could be the student's network in class and whether they have a good support for understanding the material inside and outside of class. All these factors play an important role in the student's performance in class and hence affect their final grades that together make up their graduation GPA.

From a research perspective, this study contributes to the literature by providing empirical evidence about the timing and ordering of courses as pursued by past students and how these relate to their graduation GPA and TTD. Researchers who develop data-driven approaches that make use of past degree plans, such as course recommendation, course sequencing, and curriculum designing, can use this information to better model the degree plans and develop more robust methods that can better assist students towards academic success, by graduating on-time with high GPA.

From an advisor perspective, this study makes a step forward towards understanding the importance of the timing and ordering of courses and how they are related to the student's graduation GPA and TTD. Advisors can use this information to better guide their students to take courses in the right time that can help them towards their academic success. They can also help them designing their own personalized plans and modify them based on their current performance and end goals, as well as show them the trade-offs they might have to make with respect to their expected graduation GPA and TTD.

From a learner perspective, knowing how the timing and sequencing of courses is related to their academic performance, especially their TTD, students can have better knowledge about how to plan their degree in order to graduate on time and save more money by taking the right set of courses in the optimal sequencing that will help them towards successful graduation in a timely manner.

Since the analysis was conducted on a large-scale dataset that spans 16 years and contains 25 majors from different colleges, we believe that the results of this analysis are generalizable and can apply on data from other universities.

There are some limitations to the current study that readers need to keep in mind for future research. Firstly, this study does not study the effect of the timing of taking courses on the students' grades in these individual courses, i.e., whether taking a course at the same, higher or lower level than the student's academic level will be related to the student's grade in this course. If such a correlation exists, then data-driven approaches need to take this into account while utilizing the degree plans. Secondly, this study does not analyze the causal inference between ordering and timing of courses and academic performance, to test whether one leads to the other. Lastly, we did not study the competition and synergy among courses taken in the same term. This might also affect the student's academic performance, since students have limited amount of time to study for the courses they take simultaneously, which creates competition among these courses. On the other hand, there might be courses in which the knowledge that one course provides during the term helps with the understanding of another course, which creates synergy among them. We plan to address these limitations in the future.

Our study has pointed out some good insights about the timing and sequencing of courses that both students and their advisors could pay attention to. However, further analysis and qualitative research is needed to identify other factors that might affect these results, such as the dynamics of the whole network of students and if closer fellows tend to take more courses together and how this affects their grades. Nonetheless, this study points towards the need for the data-driven approaches that work on course recommendation and sequencing or curriculum designing to consider the differences in the degree plans and know how to best utilize them.

ACKNOWLEDGMENTS

This work was supported in part by NSF (1447788, 1704074, 1757916, 1834251), Army Research Office (W911NF1810344), Intel Corp, and the Digital Technology Center at the University of Minnesota. Access to research and computing facilities was provided by the Digital Technology Center and the Minnesota Supercomputing Institute, http://www.msi.umn.edu.

REFERENCES

[1] Carmen Aina, Eliana Baici, and Giorgia Casalone. 2011. Time to degree: students' abilities, university characteristics or something else? Evidence from Italy. *Education Economics* 19, 3 (2011), 311–325.
[2] Andreas Behr and Katja Theune. 2016. The causal effect of off-campus work on time to degree. *Education Economics* 24, 2 (2016), 189–209.
[3] Narimel Bendakir and Esma Aïmeur. 2006. Using association rules for course recommendation, Vol. 3.
[4] Kiratijuta Bhumichitr, Songsak Channarukul, Nattachai Saejiem, Rachsuda Jiamthapthaksin, and Kwankamol Nongpong. 2017. Recommender Systems for university elective course recommendation. In *Computer Science and Software Engineering (JCSSE), 2017 14th International Joint Conference on*. IEEE, 1–5.
[5] Agar Brugiavini, Carlo Carraro, and Matija Kovacic. 2015. Academic Achievements: Grades Versus Duration. *International Center for Climate Governance* (2015).
[6] Mihai Cucuringu, Charles Z Marshak, Dillon Montag, and Puck Rombach. 2017. Rank aggregation for course sequence discovery. In *International Workshop on Complex Networks and their Applications*. Springer, 139–150.
[7] Asmaa Elbadrawy and George Karypis. 2016. Domain-aware grade prediction and top-n course recommendation. In *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM, 183–190.
[8] Nina Hagemann, Michael P OâĂŹMahony, and Barry Smyth. 2018. Module Advisor: Guiding Students with Recommendations. In *International Conference on Intelligent Tutoring Systems*. Springer, 319–325.
[9] Serge Herzog. 2006. Estimating student retention and degree-completion time: Decision trees and neural networks vis-à-vis regression. *New directions for institutional research* 2006, 131 (2006), 17–33.
[10] Aditya Parameswaran, Petros Venetis, and Hector Garcia-Molina. 2011. Recommendation systems with complex constraints: A course recommendation perspective. *ACM TOIS* 29, 4 (2011), 20.
[11] Aditya G Parameswaran, Hector Garcia-Molina, and Jeffrey D Ullman. 2010. Evaluating, combining and generalizing recommendations with prerequisites. In *Proceedings of the 19th ACM international conference on Information and knowledge management*. ACM, 919–928.
[12] Aditya G Parameswaran, Georgia Koutrika, Benjamin Bercovitz, and Hector Garcia-Molina. 2010. Recsplorer: recommendation algorithms based on precedence mining. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*. ACM, 87–98.
[13] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research* 12, Oct (2011), 2825–2830.
[14] Zhiyun Ren, Xia Ning, and Huzefa Rangwala. 2018. ALE: Additive Latent Effect Models for Grade Prediction. In *Proceedings of the 2018 SIAM International Conference on Data Mining*. SIAM, 477–485.
[15] Katja Theune. 2015. The working status of students and time to degree at German universities. *Higher Education* 70, 4 (2015), 725–752.