Curricular Complexity of Student Schedules Compared to a Canonical Degree Roadmap

Anthony M. DeRocchis, Laura E. Boucheron, Mark Garcia, and Steven J. Stochaj
Klipsch School of Electrical & Computer Engineering
New Mexico State University
Las Cruces, NM 88003
Email: {tonydero, lboucher, mgarcia3, sstochaj}@nmsu.edu

***Abstract*—This Research-to-Practice Work In Progress paper presents a combination of curricular analytics and student performance analysis in a Bachelor of Science in Electrical Engineering (BSEE) degree program. The study of curricular complexity has gained attention in recent years as a means to quantify the ease of progression through a degree program and to highlight bottleneck courses. While there is much attention to canonical curricular complexity of a recommended degree roadmap, there is little analysis of what paths students actually take through a curriculum. Here, we apply the concept of curricular complexity to students' progression through a BSEE degree and compare those results to the curricular complexity of the canonical degree roadmap. We study the relationship between time-to-graduation, student schedule complexity, and number of extraneous courses. We find that many students take breaks during their degree, extending time-to-graduation, even while many enter the degree with transfer credits. We find a direct relationship between schedule complexity and time-to-graduation, with a one-term increase in time-to-graduation similar to but slightly lower than the average term-weighted complexity. Finally, we find that students take extraneous courses consistent with three to four extra terms, indicating an additional factor in time-to-graduation.**

I. Introduction

It is increasingly recognized that there are significant retention issues in STEM fields [1], [2]. Positive effects of student-centered interventions such as supplemental instruction, mentoring, and incorporation of pedagogical variety in the curriculum have been studied [1]–[3]. At the same time, there are concerns about increases in or inaccurate estimates of time-to-graduation [4], [5]. Time-to-graduation is significantly impacted by curricular characteristics such as total number of credits and the prerequisite structure of the curriculum [6] and can have negative financial impacts to students [7]. Here, we study students' progression through a Bachelor of Science in Electrical Engineering (BSEE) degree as measured with curricular complexity and compare those results to the curricular complexity of the canonical degree roadmap.

This work will provide insights not only into the curricular design of a degree, but to the typical use cases of students which may indicate a need for reconsideration of curricular structure. For example, roadmaps often assume a first semester enrollment and successful completion of Calculus I. Anecdotally, many students may not be calculus ready in the first semester, but the actual degree paths taken by students and curricular complexity of those paths has not been considered.

There are many implications of this work. First, this work will allow for a more quantitative approach to curricular design to improve student success and decrease time-to-graduation. For example, new models for engineering math education have been successful in addressing the calculus bottleneck [3]. Second, this will allow for strategic investment of resources (e.g., teaching assistants or supplemental instruction) to courses most critical in the degree program. Third, this will provide tools for academic advisors who help students plan class schedules to promote success in the program. Fourth, this will allow students to be more proactive in their learning by understanding broader implications of their current performance.

There are three main contributions we discuss in this paper. First, we establish the distribution of time-to-graduation for our student population. Second, we compute students' schedule complexity and relate it to time-to-graduation. Third, we analyze the number of extraneous courses taken by students as a potential factor in increased time-to-graduation.

II. Background and Theoretical Framework

The theoretical framework of this work is based in curricular analytics [6] which defines quantitative measures of the importance of courses to curricula and the complexity of curricula based on prerequisite structure. Curricular design, even when considering measures such as curricular complexity, may provide a misleading metric if students cannot (e.g., due to lack of preparation) or do not (e.g., due to poor advising or planning) follow the roadmap. We use concepts of curricular analytics to measure the canonical degree roadmap for the BSEE and compare to the complexity of student paths through the BSEE.

We represent a curriculum as a directed acyclic graph (DAG) G using the NetworkX network analysis library in python [8]. The vertices V_i of G are courses and the directed edges $E_{ij} = (V_i, V_j)$ represent that course V_i is a prerequisite to course V_j. We use the concepts of course cruciality and curricular complexity [9]. A course that is more crucial to a degree is, intuitively, one that is a prerequisite for many courses. The course cruciality C_i is defined as the sum of the blocking factor B_i and delay factor D_i:

$$C_i = B_i + D_i = desc(V_i) + lp(V_i), \tag{1}$$

where $B_i = desc(V_i)$ is the number of descendants of V_i (courses with V_i as a direct prerequisite or through a chain of

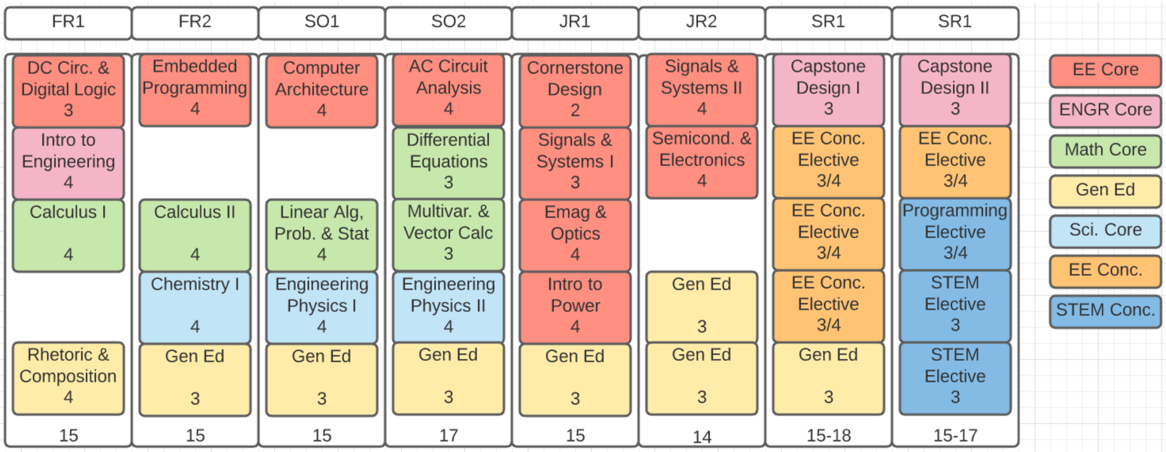


Fig. 1. BSEE curriculum roadmap for the 2019 catalog. The number specified for each course is the credit hours and the number specified for each semester is the semester credit hours.

prerequisites) and $D_i = lp(V_i)$ is the length of the longest path containing V_i. We use the standard definition of path length for D_i, but note that this this results in values 1 less than [9]. In turn, a curriculum that is more complex is, intuitively, one that involves many and long paths of prerequisites. The complexity of a curriculum is defined as the sum of the n constituent course crucialities,

$$S = \sum_{i=1}^{n} C_i. \tag{2}$$

In this work, we define two further quantities, the term-weighted course cruciality (TWC) and the term-weighted curricular complexity (TWCC). The TWC is the product of C_i and the term T_i in which the course is taken,

$$TWC_i = C_i \cdot T_i, \tag{3}$$

where, e.g., $T_i = 2$ for the second semester. A course with large TWC is one that is crucial to the curriculum and/or taken later in the degree, indicating the potential for less flexibility to retake the course in the event of a non-passing grade. As such TWC can be used to strategically focus resources (e.g., supplemental instruction) to enhance student success. The TWCC is defined as the sum of the n TWCs,

$$TWCC = \sum_{i=1}^{n} TWC_i. \tag{4}$$

TWCC favors curricula in which the most crucial courses are taken earlier. This measure is identical to the objective function used in the cruciality-based curriculum balancing of [10]. Whereas TWCC is used in [10] to optimize a curriculum, we use it to compare a canonical roadmap to student schedules.

III. Data

A. The BSEE Curriculum Roadmap

We illustrate a roadmap for our BSEE in Fig. 1. The EE Concentration Electives, the STEM Electives, and the Programming Elective in the Senior year are constrained by each BSEE concentration area. Here we focus on the core of the curriculum, including the EE, engineering, math, and science core as these are common to all BSEE students. We do not consider the effect of General Education courses as those significantly vary between students and have fewer prerequisites and thus a small impact on curricular complexity. We outline pre- and co-requisites for the BSEE core in Table I.

TABLE I
BSEE CORE ALONG WITH COURSE CRUCIALITIES (C), TERM-WEIGHTED CRUCIALITIES (TWC), AND PRE- AND CO-REQUISITES.

Course	C	TWC	**Pre-/Co-requisites**[a]
(1) DC Circ. & Dig. Log.	18	18	Trig. & Precalc.[b]
(2) Intro to Engineering	0	0	College Algebra[b]
(3) Calc. I	21	21	Trig. & Precalc.[b]
(4) Embed. Prog.	14	28	(1)
(5) Calc. II	19	38	(3)
(6) Chem. I	6	12	Int. Alg.[b]
(7) Comp. Arch.	7	21	(1), (4-C), Trig. & Precalc.[b]
(8) Lin. Alg., Prob., Stat.	10	30	(4), (5)
(9) Eng. Phys. I	16	48	(3)
(10) AC Circ.	14	56	(1), (5), (15-C)
(11) Diff. Eq.	9	36	(5)
(12) Multivar. & Vec. Calc.	8	32	(4), (5)
(13) Eng. Phys. II	15	60	(9), (5)
(14) Cornerstone Des.	7	35	(4), (7), (10)
(15) Sig. & Syst. I	9	45	(8), (10-C), (11-C)
(16) Electromag. & Optics	7	35	(10), (12), (15)
(17) Intro to Power	7	35	(10)
(18) Sig. & Sys. II	8	48	(15), (11)
(19) Semicond. & Electr.	7	42	(10), (6)
(20) Capstone Des. I	7	49	(14), (10), (18), (17), (16)
(21) Capstone Des. II	6	48	(20)
TOTAL (TWCC)		737	

[a] Corequisites are denoted with -C [b] Competencies assumed prior to the start of the degree.

B. Student Data

The use of student data was approved by our university's Institutional Review Board (IRB) and all data was de-identified

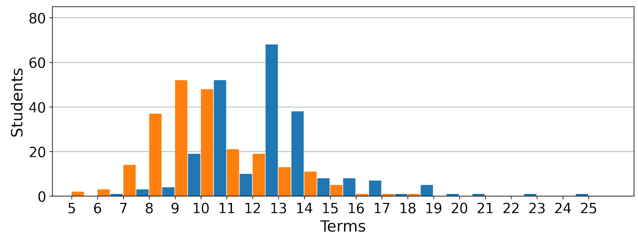


Fig. 2. Number of terms to graduate. Blue: total number of terms, including terms with no enrollment. Orange: total number of enrolled terms.

prior to analysis as described below. We downloaded data for all declared EE majors from Spring 2008 to Spring 2020 from our university's records management system. These data provide all classes enrolled in and grades earned by EE students. The data were combined into a single comma separated variable (CSV) file with an additional column indicating the term of origin, extracted from the filenames. All personally identifiable information (PII) was removed and a unique alphanumeric string was assigned to each student using the SHA512 hashing algorithm [11] applied to the students' ID numbers. The deidentified dataset is used for all subsequent analysis which is conducted using the pandas data processing library in python [12].

The base dataset includes data for $n = 2000$ students who took courses during 36 terms. Removing students with a first term prior to 2008 (or other invalid term values) and focusing on students who were awarded the BSEE on or before Spring 2020, we have a sample of $n = 333$ students. Finally, we consider the number of core courses completed and use the 25^{th} percentile to exclude those students that took the 25% fewest core courses. Since TWCC analyzes the order in which courses are taken, we focus on those students who completed the majority of the BSEE core. Finally, the dataset has $n = 228$ students who graduated between 2012 and 2020.

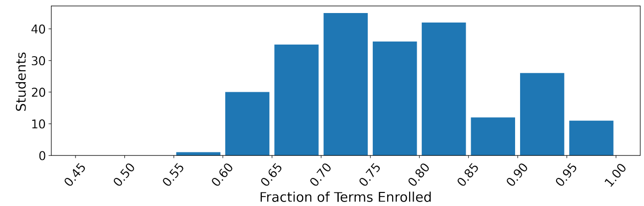
We add columns to the dataset for B_i, D_i, and C_i for each course as in Eq. (1). A column for T_i is used in combination with C_i to calculate TWC_i for each course as in Eq. (3). Finally, as in Eq. (4), these are summed to provide TWCC for student schedules.

IV. Analysis of Student Degree Progression

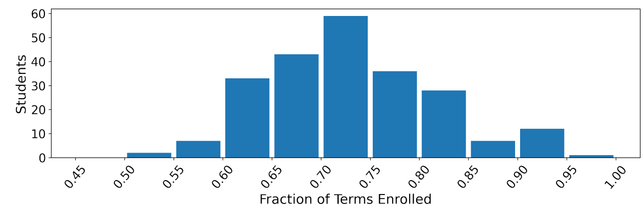
A. Time to Graduate

We show a histogram of number of terms to graduate in Fig. 2 where we see that a significant portion of students took breaks during their degree, visible by the shift to the left from the blue to orange histogram bars. This same effect is seen in Fig. 3 where we show the fraction of total terms enrolled in any courses and BSEE core courses.

In Fig. 2 the number of terms to graduate (total time earning the degree) was 12.87 ± 2.46 and the number of enrolled terms to graduate was 9.99 ± 2.19. Ideally, number of terms to graduate is 12, corresponding to the four year degree plan. In Fig. 3(a), we further note that there are a significant number of students ($n = 205$) who enroll in summer courses,



(a) Any courses



(b) Core courses

Fig. 3. Histogram of the fraction of the terms enrolled at our university.

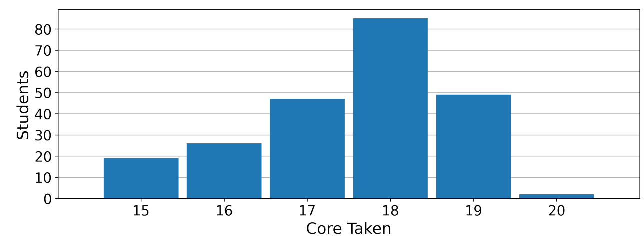


Fig. 4. Number of BSEE core courses taken.

corresponding to a fraction of terms enrolled > 0.67. The distribution for BSEE core courses in Fig. 3(b) has a less significant tail for fraction of terms > 0.67 ($n = 180$), likely due to the fact that many upper division BSEE core courses are not commonly offered in the summer. While these conclusions may appear simple, this is a quantitative indication of students taking breaks in their degree and the utilization of summer courses by those students which is otherwise difficult to quantify without gathering additional data from students.

B. Student Schedule Complexities

We use TWCC to compare student paths through the curriculum. We neglect course retakes in this analysis and use the first term the student took the course as term value T_i for computation of TWCC. This corresponds to initial student plans and allows us to more readily compare to the canonical roadmap. All courses are converted to 2019 catalog course equivalents for comparison. Note that Cornerstone Design did not exist in any form prior to the 2016 catalog, and thus the majority of students in our dataset have not taken that course.

Fig. 4 shows that no student took all 21 core courses and Fig. 5 shows that many ($n = 103$) completed the degree with a TWCC less than the canonical value of 737. These results indicate that students are entering the program with credit for many courses. This insight is particularly relevant at our university since it is otherwise difficult to quantify the totality of reasons for existing credit, including transfer credit,

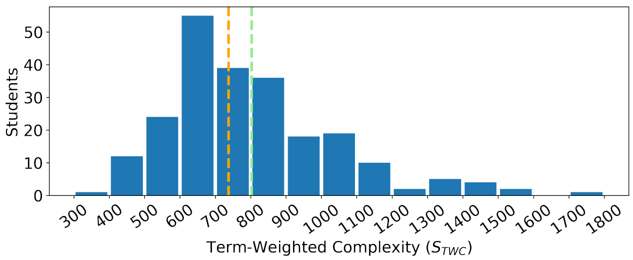


Fig. 5. Histogram of the calculated term-weighted schedule complexities for students. The mean and canonical values are shown by the horizontal green and orange dashed lines, respectively.

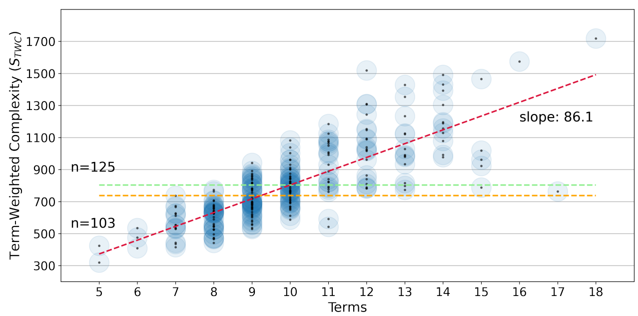


Fig. 6. Scatterplot of number of enrolled terms versus TWCC. The mean and canonical values are shown by the horizontal green and orange dashed lines, respectively. The crimson dashed line shows the linear fit line.

advanced placement credit, and equivalencies and exceptions. Even with this indication of significant entering credits, we see a heavy tail toward larger TWCC and an average TWCC of 803 ± 238, which is larger than the canonical value of 737. To the best of our knowledge, this is the first analysis of schedule complexities using actual student data; the studies in [13] only discuss the complexity of degree roadmaps. Our definition for D_i results in curricular complexity values off by a constant equal to the number of courses in the curriculum; adjusting for this, the complexity of our curriculum is similar to those studied in [13]. Future work will investigate more specific scenarios that result in increased student schedule complexity.

In Fig. 6 we examine the linear relation between TWCC and number of terms enrolled using NumPy's polyfit module [14] which indicates a slope of 86.1. This gives an average TWCC value per term for the students in our dataset. Interestingly, this value is close to, but less than, the canonical average value of $737/8 = 92.1$ which implies students are taking a lighter crucial course load than expected each term. This value could be used by academic advisors, in conjunction with other information, to determine whether students are progressing through the program at an appropriate rate.

C. Analysis of Extraneous Coursework

Given the dichotomy between number of terms to graduation (longer than the nominal 12 terms) versus number of core courses completed (fewer than the nominal 21), we study the number of extraneous courses completed by students. We define an extraneous course as one that does not satisfy any of the degree requirements of the BSEE. We find that students take 11.3 ± 6.8 extraneous courses while pursuing their BSEE; see also Fig. 7. Given that full-time enrollment is 3-4 courses per term, these extraneous courses contribute to 3-4 extra terms of enrollment and are thus one source of increased time-to-graduation. Since these extraneous courses do not factor into the TWCC, this could be one explanation for the average TWCC value per term being smaller than the canonical value.

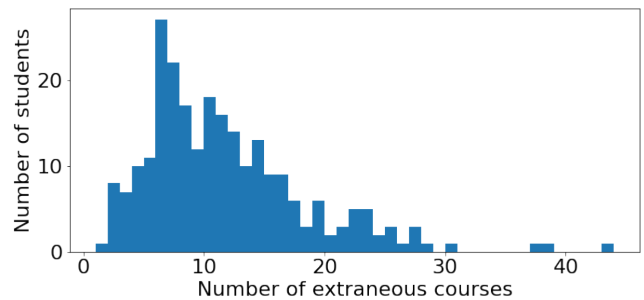


Fig. 7. Histogram of number of extraneous courses taken by students.

There are multiple reasons for extraneous courses including pursuit of multiple degrees or minors, completion of deficiency coursework, "padding" of semester schedules to remain full time and retain financial aid, and inadvertent choice of courses that do not satisfy degree requirements. These results warrant future study into reasons for extraneous courses. In particular, curricular design with attention to curricular complexity may mitigate the "padding" of schedules and more proactive advising may mitigate inadvertent choice of courses that do not satisfy degree requirements.

V. Conclusions and Future Work

We have shown that meaningful insights can be extracted by calculating the TWCC to evaluate student schedules. We see that more than half our students are completing the BSEE with a TWCC less than the canonical value, but that the mean is larger and the distribution is skewed toward larger values. Relating the TWCC to time-to-graduation, we find that students are taking lighter crucial course loads each term than the canonical roadmap. Finally, we find that students take the equivalent of several extra terms worth of extraneous courses which extends their time to graduation.

In addition to the future work discussed throughout Section IV, there are many aspects of this dataset yet to be examined. We wish to incorporate the BSEE concentrations, and their respective courses, into our calculations as a more complete picture of students paths through our curriculum. Creating a more robust framework for analyzing student's schedules will also allow us to provide real-time feedback for students while planning their schedule for future terms.

Acknowledgment

The authors gratefully acknowledge US Department of Education Grant P120A170064 which supported this work.

References

[1] B. N. Geisinger and D. R. Raman, "Why they leave: Understanding student attrition from engineering majors," *International Journal of Engineering Education*, vol. 29, no. 4, p. 914, 2013.

[2] X. Chen, "STEM attrition: College students' paths into and out of STEM fields. Statistical analysis report. NCES 2014-001." *National Center for Education Statistics*, 2013.

[3] N. W. Klingbeil and A. Bourne, "The Wright State model for engineering mathematics education: A longitudinal study of student perception data," in *121st ASEE Annual Conference & Exposition*, 2014, pp. 1–11.

[4] V. Gonzalez and D. Esparza, "Work in progress—advising tool to improve the time for graduation and the transfer of students from a community college to engineering school," in *2010 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2010, pp. T3H–1.

[5] L. DeAngelo, R. Franke, S. Hurtado, J. H. Pryor, and S. Tran, "Completing college: Assessing graduation rates at four-year institutions," Los Angeles: Higher Education Research Institute, UCLA, Tech. Rep., 2011.

[6] G. L. Heileman, C. T. Abdallah, A. Slim, and M. Hickman, "Curricular analytics: A framework for quantifying the impact of curricular reforms and pedagogical innovations," *arXiv preprint arXiv:1811.09676*, 2018.

[7] G. D. Kuh, J. L. Kinzie, J. A. Buckley, B. K. Bridges, and J. C. Hayek, *What matters to student success: A review of the literature*. National Postsecondary Education Cooperative Washington, DC, 2006, vol. 8.

[8] A. Hagberg, D. Schult, and P. Swart, "NetworkX reference, release 2.6rc1.dev0," https://networkx.org/documentation/latest/_downloads/networkx_reference.pdf, Tech. Rep.

[9] A. Slim, J. Kozlick, G. L. Heileman, and C. T. Abdallah, "The complexity of university curricula according to course cruciality," in *2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems*. IEEE, 2014, pp. 242–248.

[10] A. Slim, G. L. Heileman, E. Lopez, H. Al Yusuf, and C. T. Abdallah, "Crucial based curriculum balancing: A new model for curriculum balancing," in *2015 10th International Conference on Computer Science & Education (ICCSE)*. IEEE, 2015, pp. 243–248.

[11] Q. H. Dang, "Secure hash standard," Federal Inf. Process. Stds. (NIST FIPS), National Institute of Standards and Technology, Tech. Rep., 2015-08-04 2015.

[12] The pandas development team, "pandas-dev/pandas: Pandas," Feb. 2020. [Online]. Available: https://doi.org/10.5281/zenodo.4524629

[13] G. L. Heileman, W. Thompson-Arjona, O. Abar, and H. Free, "Does curricular complexity imply program quality," in *Proceedings of the 2019 American Society for Engineering Education (ASEE) Annual Conference, Tampa, June*, 2019, pp. 10–16.

[14] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. [Online]. Available: https://doi.org/10.1038/s41586-020-2649-2