Using Business Intelligence to Mitigate Graduation Delay Issues

Khaled Almgren
PhD Candidate
Department of Computer science and Engineering
University of Bridgeport

Abstract

Graduate master students usually have an expected time to graduate. Graduation delay causes some problems for the universities, such as consuming extra resources. In addition, the number of accepted graduate students is limited in each semester. Consequently, this will reduce the number of new admitted students to graduate programs. The paper solves the graduation delay problem by analyzing and studying the collected data to find out the main reasons behind the problem using some of the Seven Major Statistical Process Control Problem-Solving tools, namely the Histogram, Pareto chart, Cause-and-Effect diagram and control chart. As a result, the optimal solution is reachable using a web-based solution integrated with a Business Intelligence system. The result shows significant changes in the percentage of graduate students graduating in less time.

Keywords: Business Intelligence, Constraint Satisfaction Problem, The Magnificent Seven, Timetabling problem

Introduction

Computer systems were used in organizations and businesses to facilitate their business processes. This can be achieved by automating the business processes. However, information systems sometimes suffer from automating algorithms because of their high complexity. Since an algorithm needs to be efficient, many computer scientists come up with their own efficient algorithms to solve complicated problems [1].

Universities have many business processes, such as class registration, class scheduling, etc. University administrations are a major customer for tech companies because they need to automate their business processes. However, these problems are not as easy as expected; they are actually very challenging. Business Intelligence is employed to solve complicated problems. Business Intelligence is used to create new opportunities for businesses and provide them with efficient strategy techniques that give the business a competitive advantage and stability [2].

One certain problem that has been focused on by computer scientists and experts to reach an optimal solution is the Graduation Delay Problem. In universities, graduate programs usually have pre-defined timelines for completion. The Graduation Delay Problem causes many issues for universities. Two of the most important issues are consuming university resources and reducing the opportunities for new students to join the graduate programs. This is because the number of spaces in graduate programs is limited. One of the problems that affects the Graduation Delay Problem is the class scheduling problem, i.e., when students attempt to register for a course, but the class is already full.

Therefore, the problem of class scheduling using traditional methods is considered to be an NP-Complete problem. Thus, many researchers from the education and technology fields have conducted research for this problem because of its wide use in universities. Many came up with algorithms to solve this problem that was based on local search techniques [1]. Since many algorithms have not reached an optimal solution, the area of research is still wide open [2].

This kind of problem is considered to be a constraint satisfaction problem (CSP). This means that the problem parameter and its domain are restricted by constraints. For example, a student cannot graduate unless he passes the core courses, the core courses are taught by a limited number of professors, etc. For CSP problems, scientists tend to prove a solution or prove that there is no solution [4].

Wasfy and Aloul tried to solve the class-scheduling problem by integrating integer linear programming, SAT- and generic-based solvers to reach a complete solution. They considered three constraints. They are: 1) classroom is only being assigned for one course; 2) courses cannot be fit into one classroom; and 3) number of students per course must be within the classroom capacity. Using the techniques and constraints, a solution was reached [3].
Gunawan, Ng, and Poh approached the problem of class scheduling and timetabling. They proposed a combination of an integer programming approach, greedy heuristic, and modified simulated annealing algorithm under the cover of hybrid algorithm to solve this problem. The proposed algorithm is composed of three main phases. They are pre-processing, construction, and improvement. The first phase involves the data for classes and their teachers. The second phase deals with building a solution. The third phase deals with improving the solution obtained in the second phase. The proposed hybrid algorithm shows the ability to solve the issue [4].

Gonzalez and Esparza built project management tools that aim to assist the students to get the academic degree sooner. The project management tools consist of two main parts. The first tool is designed to provide a private visualization map for each student for course sequences, as well as making advisory adjustments that will optimize the time to get the degree under a restricted set of resources. The second tool aims to collect information from many students in different semesters to determine the problems in the curriculum [5].

Shatnawi, Al-Rababah and Bani-Ismail proposed a clustering approach based on an FP-tree algorithm to solve the problems found in course scheduling times during earlier semesters that did not allow students to register for required classes, which resulted in delayed graduation [6].

In this paper, a new approach has been proposed to solve the graduation delay problem by integrating a web-based solution and Business Intelligence. Some of The Magnificent Seven are used to analyze variation to define potential improvements, reduce variation, and remove waste in the graduation delay problem. The rest of the paper is organized as follows: Section 2 discusses the problem environment, Section 3 describes the problem identification, Section 4 shows the model development, Section 5 presents the proposed solution, Section 6 shows the result, and Section 7 concludes the paper.

Problem Environment

The graduation delay problem occurs in universities and community colleges. Many students from various ages and countries suffer from this problem. In order to approach this problem, we must collect data about the number of graduated students and the number of semesters that it took the graduated students to graduate. This data is collected for a period of fifty years.

The data was collected from 1968 to 2012, but the technologies that were used have evolved. Before the technology was used in universities, the issue with the traditional registration module is that there was no communication between the student and the university; for example, the university had no idea about the student concentration. A university usually offers courses based on the professor and classroom availabilities. As shown in Figure 1 for the traditional registration system, the student has to follow fixed procedures to register for courses for the next semester. This methodology has many flaws that reduce the system performance.

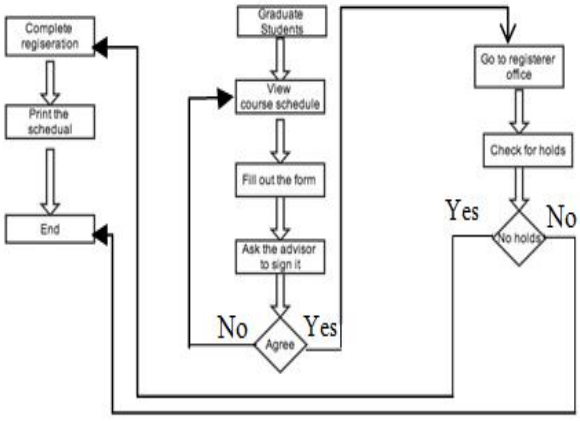

Figure 1: Traditional Registration Module

Problem Identification

Graduation delay problem is the term used to characterize students taking extra time to graduate. For example, undergraduate students usually take four years to graduate from science schools, but some students take up to five years; in other words, there is a one-year delay. The consequences of this delay are time and cost. There are many motivations for solving this problem and its negative effects on universities. The first effect that motivates universities to look for a solution is overconsumption of resources. For example, when a university has many students who have not graduated on time in addition to newly admitted students, the university needs to increase the number of faculty members or cause an overload on current professors that will affect their performance. The issue here is that hiring new faculty members is expensive and the number of students is changing rapidly and not constant. The change in the number of graduated students will cost the department extra money when they have extra professors. The second issue is that the number of spaces is limited in graduate programs. Thus, when a student graduates late, their space is occupied for that extra time and unavailable to a new graduate student. As mentioned earlier, the problem has a huge impact on universities and has motivated many scientists to solve it [7].

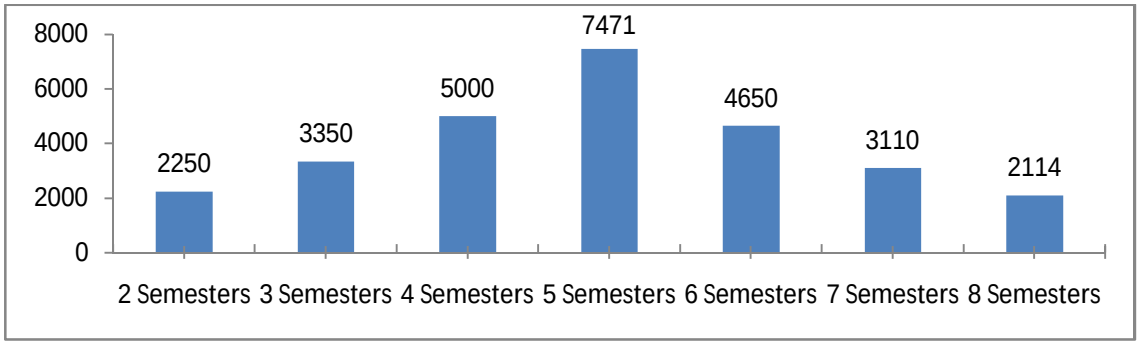

Figure 2: Average Number of Graduating Students for Each Semester

The control chart in Figure 3 illustrates that the process is under control. Therefore, it can be concluded that the process variation is consistent and the output is predictable within the bounds of control limits.

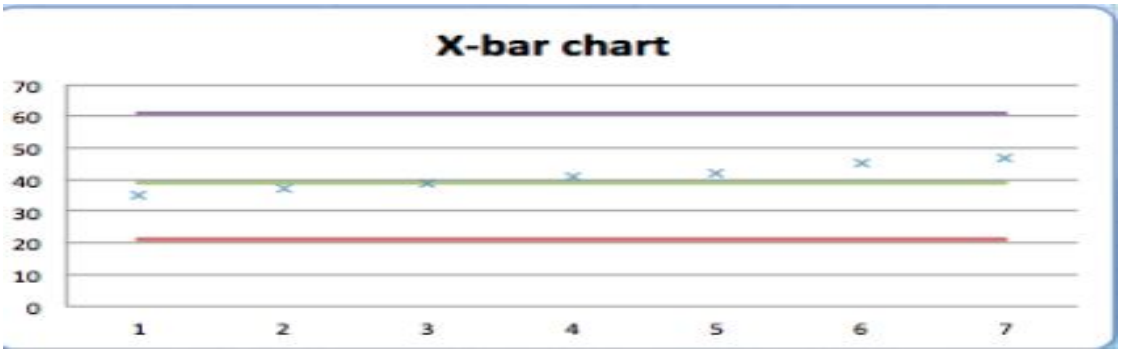

Figure 3: X-Bar Chart for Initial Collected Data

Model Development

In order to determine the reasons behind the fact that students take more semesters than usual to graduate, we conducted interviews with admission officials and students, as well as developed on-line questionnaires for students in order to collect the data.

A Pareto chart and a Cause-and-Effect diagram were designed to troubleshoot problems in statistical quality modules. Therefore, this paper used these modules as tools in identifying the main causes of the student's graduation delay problem. In addition, a control chart has been used to monitor the stability of the process [8, 9].

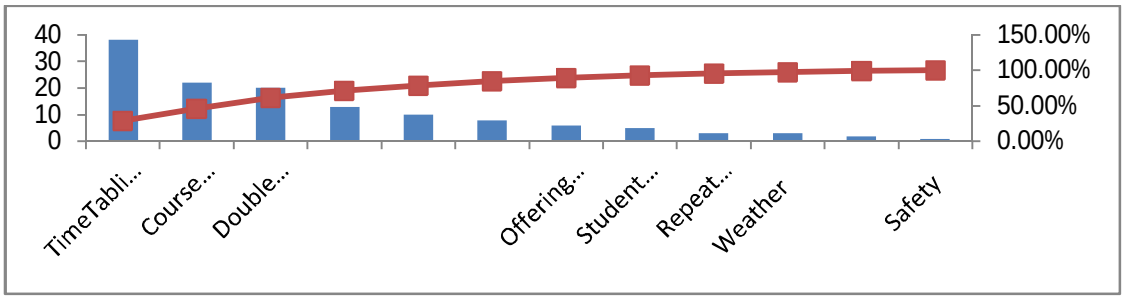

Figure 4: Pareto Chart

From observing the Pareto chart, we see that 80% of the problems identified by the first three major causes are timetabling problem, double major, and project delay. Based on the Pareto chart methodology, solving the first three major problems will increase the number of graduated students. Thus, we consider this the proposed solution to overcome these problems.

All types of problems are solvable using a Fishbone diagram because it makes it easy to understand and organize the problem. In addition, it specifies the causes of the problem. Furthermore, the visual technique that is used in this methodology makes it easier to illustrate the problem and follow the flow of problems. In addition, it gives a general idea about the problem to the reader without having to read the problem statement. The following Fishbone diagram was drawn to understand the source of Students' Graduation Delay.

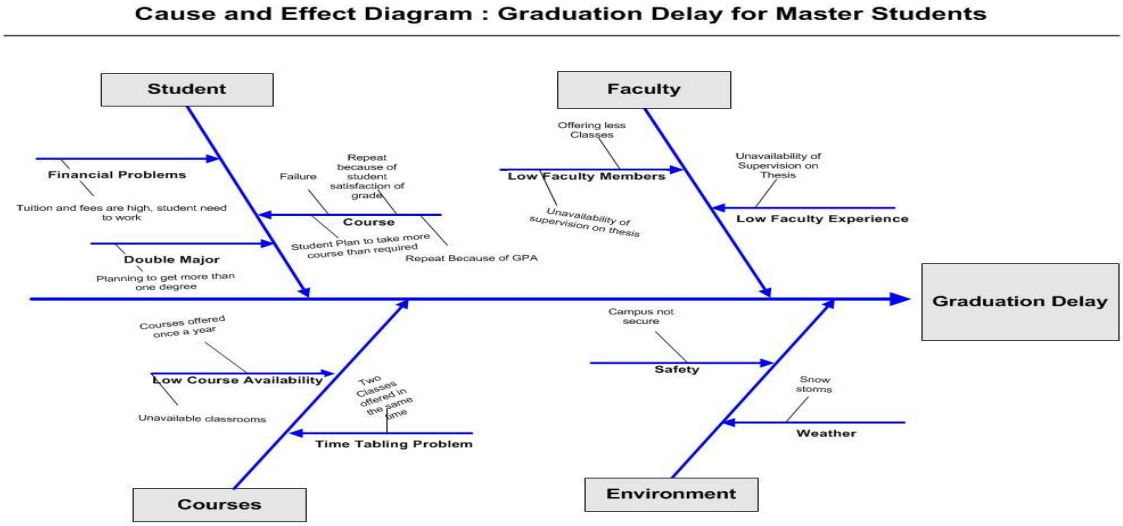

Figure 5: Cause and Effect: Graduation Delay

According to Figure 4, the four main causes that affect the delay of student graduation are Student, Faculty, Courses and Environment. The first cause is the student. A student can have many problems, such as a low GPA, and thus may need to retake the course. The second cause is course. The main problem that faces the student is the conflict between courses, which can be caused by the timetabling problem. For example, the student may want to register for two courses, but these two courses are offered on the same day at the same time.

The third cause is environment. This can delay student graduation. For example, if a natural disaster occurs, the university can be closed for months. The fourth cause is professors. A low number of faculty members leads to offering fewer classes. Thus, students have less opportunity to enroll in classes.

Proposed Solution

The purpose of the solution is to increase the number of graduated students in three to four semesters, which is the average. The proposed solution integrates a web-based solution and Business intelligence (BI).

Business Intelligence (BI) is a set of software tools that manipulate a large volume of data to assist institutes with making better decisions, reduce cost and define new business opportunities. The architecture of the BI is shown in Figure 5 [10].

Business Intelligence has been used widely during the last few decades. The main advantage of using such tools and advanced technology is to provide the consumers with more data that is valuable to achieve the goals of solving the complex problems with optimum solutions. Data is the raw material of the entire Design making. Applying the BI system needs to have the raw material for the design, by providing the raw data that could be used during the process. The first step to provide the data to the intelligence system is gathering the data from the data warehouse; furthermore, the raw data must be normalized. This normalization includes the build of the data and the relation diagram between all the required fields. This standardization is provided to the BI system [11].
The BI build is of four main components that are integrated to achieve a solution to the problem. The integration is mainly integrated with the data warehouse that is mainly focused on combining the data from data sources to one integrated data source. The second phase works on using tools of AI to analyze the data and give reports in the form of charts and data analysis forms as shown in Figure 6.

The proposed solution gives the students and universities the optimal solution to reduce the time for graduated students to graduate. Firstly, before the registration process starts, the students will be able to pick their courses from a set of offered courses based on their individual concentrations. The web-solution will then generate a report. This report will contain data about the most selective courses and concentrations. This data will be used in a business intelligence system to assist and assess institutes with a list of courses that students will choose from to graduate early.

The web-based solution is based on gathering pre-information that gives pre-identification for the student concentration and courses. This measurement tool gives the university the ability to offer the most selective courses [11].

The Business Intelligence system will offer to each student a private page showing a list of courses based on his/her concentration and graduation requirement constraints. For example, a student in the MS in computer science program has five core courses, such as CS550 Advanced database, so the system will employ this fact to generate the list of courses; thus, the system will force students to take the core courses in a limited time, such as three to four semesters. Moreover, the BI system will give the students the optimum study plan that allows students to graduate in a shorter time.

Based on the Pareto chart, the main cause of the problem is the Timetabling problem. The BI system will solve this problem by employing different mathematical techniques in generating a course timetable that satisfies students, professors, and program requirements.

The BI system solves the double major issue. The BI system will consider the students who plan to major in more than one area. The student will notify the BI system with his/her plan of getting a double major. Then the BI system will consider this fact, as well as early graduation, in the timetable generation process.

The project delay is another problem that causes students to graduate late. The BI system solves this problem by allowing the students to register for the project early.

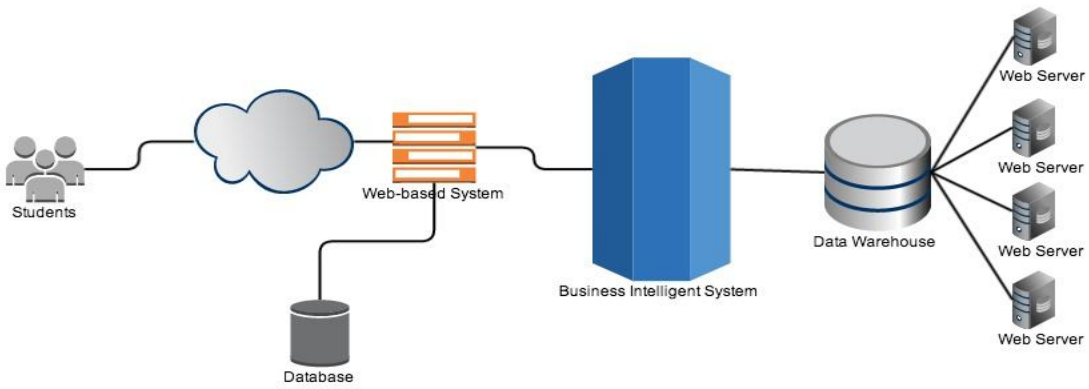

Figure 6: Solution Architecture

Results

The goal of this paper is to offer a solution to students and universities to allow students to graduate sooner. Implementing the web-based solution and using the report from the web-based solution in the Business Intelligence system solved the problem. After applying the statistical techniques, the new figure shows significant improvement.

In Figure 7, the average of graduated students graduating in three to four semesters has increased significantly. By comparing the average in Figure 2 and Figure 7, the average to graduate was five semesters in Figure 2, but in Figure 7, the average became four semesters.

It can be seen clearly from Figure 7 that the number of students who graduated in seven to eight semesters has remarkably decreased. The number of students who graduated in three semesters has doubled from Figure 2 to Figure 7.

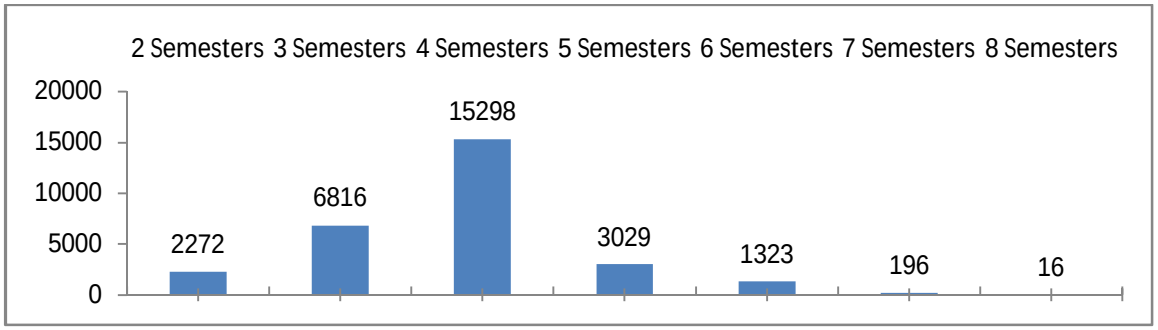

Figure 7: The Histogram Representing the Results

X-bar as shown in Figure 8 shows that the data after implementing the BI system are under control and outcomes are predictable within the bounds of control limits.

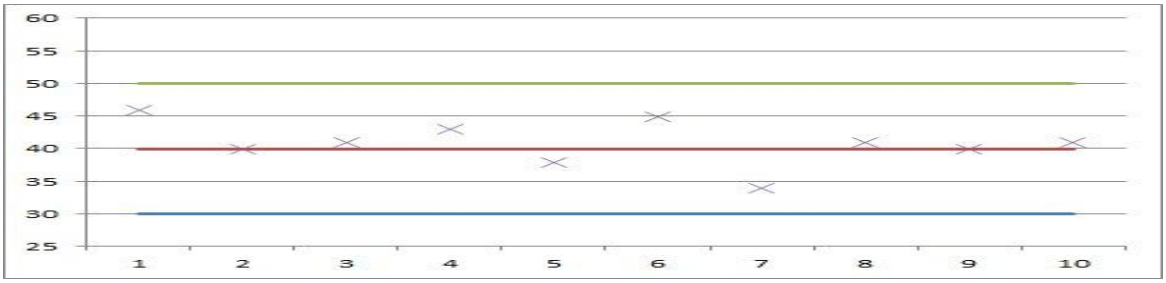

Figure 8: X-Bar Chart after Implementing the Solution

Conclusion

In conclusion, we applied a solution that integrates between a web-based reporting system and Business Intelligence to solve the graduation delay problem. We applied some of the Seven Major Statistical Process Control Problem-Solving tools, such as the Histogram, Pareto chart, and Cause-and-Effect diagram to understand and study the problem. It helped us to find out the causes of the problem. After implementing the solution, we applied the statistical techniques again and we reached great results.

Much can be done to further improve the results. The power of genetic algorithms can be used to solve NP problems, and if applied in our solution, it can get better results. Neural Networking can also be used to enhance the results by predicting the students' behaviors. Profile analysis can also be used to improve the course selection by predicting student performance.

References

D. Callon, *Competitive advantage through information technology*: McGraw-Hill Higher Education, 1995.

L. W. Porter and L. E. McKibbin, *Management Education and Development: Drift or Thrust into the 21st Century?*: ERIC, 1988.

A. Wasfy and F. Aloul, "Solving the university class scheduling problem using advanced ILP techniques," in *IEEE GCC Conference*, 2007.

A. Gunawan, K. M. Ng, and K. L. Poh, "Solving the teacher assignment-course scheduling problem by a hybrid algorithm," *World Academy of Science, Engineering and Technology*, vol. 33, pp. 259-264, 2007.

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 *Frontiers in Education Conference (FIE), 2010 IEEE*, 2010, pp. T3H-1-T3H-2.

S. Shatnawi, K. Al-Rababah, and B. Bani-Ismail, "Applying a novel clustering technique based on FP-tree to university timetabling problem: A case study," in *Computer Engineering and Systems (ICCES), 2010 International Conference on*, 2010, pp. 314-319.

R. Kolisch and A. Sprecher, "PSPLIB-a project scheduling problem library: OR software-ORSEP operations research software exchange program," *European Journal of Operational Research*, vol. 96, pp. 205-216, 1997.

M. J. Chandra, *Statistical quality control*: CRC Press, 2010.

S. Roberts, "A comparison of some control chart procedures," *Technometrics*, vol. 8, pp. 411-430, 1966.

B. Liautaud and M. Hammond, *e-Business intelligence: turning information into knowledge into profit*: McGraw-Hill, Inc., 2000.

W. B. Ashton and R. A. Klavans, *Keeping abreast of science and technology: technical intelligence for business* vol. 11: Battelle Press Columbus, 1997.