A profiling-based algorithm for exams’ scheduling problem

Tawfiq Alrawashdeh, Khaldun G. Al-Moghrabi, Ali M. Al-Ghonmein
Faculty of Information Technology, Al-Hussein Bin Talal University, Ma’an, Jordan

ABSTRACT
Typically, the problem of scheduling exams for universities aims to determine a schedule that satisfies logistics constraints, including the number of available exam rooms and the exam delivery mode (online or paper-based). The objective of this problem varies according to the university’s requirements. For example, some universities may seek to minimize operational costs, while others may work to minimize the schedule’s length. Consequently, the objective imposed by the university affects the complexity of the problem. In this study, we present a grouping-based approach designed to address the problem of scheduling the exam timetable. The approach begins by profiling the courses’ exams based on their requirements, grouping exams with similar requirements to be scheduled at the same time. Then, an insertion strategy is used to obtain the exam schedule while satisfying the imposed constraints of the targeted university. We applied this approach to the problem of exam scheduling at Al-Hussein Bin Talal University in Jordan and achieved a balanced exam schedule that met all the imposed constraints.

1. INTRODUCTION
The problem of scheduling exam timetables is a well-known NP-hard problem [1]. Generally, the objective of this problem is to minimize the cost of running the exam while satisfying logistics constraints. The cost can be considered as the total number of allocated exam rooms. Logistics constraints capture available resources (rooms and invigilators) and university requirements, such as the maximum number of exams a student can perform in a day. Such requirements and resources vary significantly between universities, which impacts the complexity of the problem. Accordingly, scheduling approaches must take into consideration the constraints imposed by the targeted university.

Several proposals have investigated the timetable scheduling problem in the literature [2]–[25]. Some of these proposals have proposed generic approaches [6], [7], [17], [25], while others have proposed solutions that address the requirements of specific universities. The mechanisms employed by these proposals can be generally categorized into nature-inspired solutions [7]–[10] and heuristic-based solutions [1], [8], [13], [16]. Nature-inspired solutions cover several types of techniques, such as genetic algorithms. While nature-inspired solutions have shown advantages in addressing the problem in its generic settings, heuristic solutions’ capabilities to address the problem regardless of the input size have emphasized the use of heuristic approaches as a suitable technique.

In this study, we address the exam scheduling problem at Al-Hussein Bin Talal University, where paper-based and computer-based exams are typically conducted. The number of exam rooms for computer-based exams is limited compared to paper-based exams. Having two types of exams suggests that a
grouping-based approach should be adopted since the requirements of these two types are different. Additionally, in this case study, some courses with different names might use the same material. These courses (synchronized courses) must be handled as one course; therefore, a single exam must be scheduled for each group of synchronized courses. Furthermore, in the presented scenario, several exams can be scheduled in the same room. These distinguishing characteristics emphasize the importance of adopting a profiling strategy, whereby each group of the courses shares the same characteristic (profile) and is scheduled separately.

Accordingly, in this study, we present a grouping-based algorithmic solution to address the problem. The solution consists of three steps: grouping, exam scheduling, and room scheduling. These steps jointly work to determine each day’s exam schedule and the room allocated to each exam. The grouping step aims to divide the exams into groups, where each group consists of exams that must be scheduled during the same time slot. In the exam scheduling step, an insertion strategy is employed to establish a tentative schedule, whereas in the room scheduling step, the established schedule is confirmed by determining the room for each exam. To evaluate the performance of the proposed approach, we conducted several sets of experiments, and the results showed that the approach can obtain a conflict-free schedule that satisfies the imposed constraints. Moreover, we performed experiments to determine the maximum number of exams that can be scheduled during the available exam period.

The rest of the paper is organized as follows. Section 2 presents and discusses the most related proposals from the literature. In section 3, we discuss the proposed grouping-based approach. Section 4 presents the discussion and remarks. In section 5, we discuss the results of the experiments. In section 6, we provide the discussion and remarks. We conclude the study in section 7.

2. RELATED WORK

Several proposals have addressed variations of the exam scheduling problem. Abdullah et al. [15] addressed the scheduling problem using a Tabu search-based mechanism. By utilizing the benefits of Tabu search, the mechanism improved the quality of the obtained schedule. However, increasing the number of exams is expected to challenge the efficiency of the proposed mechanism due to the computational requirements of Tabu search. Other searching-based solutions have investigated the exam scheduling problem [5]–[7], [10], [11], [25], including the use of simulated annealing [5], [11], genetic algorithms [10], [25], and Tabu search [7]. Dener and Calp [6] investigated a version of the problem where people who know each other are not expected to be in the same exam room. They proposed a genetic-based algorithm that considers students’ contact information to determine any relationships between them.

Abela and Abramson [25] presented a two-stage scheduling approach to address the exam scheduling problem at the Arab East College for High Education in Saudi Arabia. The objective of the proposed approach is to minimize the number of exam conflicts and improve the schedule’s quality. The first stage of the proposed approach acts as an input preparation step, while the second stage obtains the actual schedules. Yaldir and Baysal [8] proposed an evolutionary-based method consisting of two steps: data collection in the first step and obtaining the schedule using an evolution-based strategy in the second step.

Mandal and Kahar [9] proposed a two-stage heuristic approach to address the exam scheduling problem. The two stages work recursively to determine the schedule for all exams. In the first stage, a graph-based approach orders and schedules the subset of the exams, and the obtained schedule is optimized using the deluge algorithm. To examine the efficiency of the proposed approach, the authors used the Toronto benchmark datasets, and the results showed significant improvement compared to the traditional deluge approach. Similarly, in [10], the proposed solution starts by constructing an initial schedule using a greedy-based heuristic. Then, the quality of the obtained schedule is improved by using a variable bilinearization and decomposition technique.

Ayob et al. [12] proposed a scheduling approach to address the exam scheduling problem at Universiti Kebangsaan in Malaysia. In the proposed approach, the authors presented scheduling constraints as hard and soft constraints. The approach consists of two phases, each dealing with a different type of constraint. In the first phase, an initial schedule that satisfies the hard constraints is generated. The initial schedule is then used by the second phase, which obtains a schedule that also satisfies the soft constraints. The use of hard and soft constraints was also investigated by [13], whereby Abayomi-Alli [13] addressed the exam scheduling problem at the Federal University of Agriculture in Nigeria. To address this problem, the authors proposed a particular swarm optimization-based algorithm, which uses a local searching technique to improve the performance of the particle swarm method.

As discussed in this section, the scheduling problem’s complexity is impacted by scheduling constraints that capture universities’ resources and requirements. Some proposals have focused on reducing the violations of soft constraints based on available resources, assuming that exam rooms and invigilators are
not enough to satisfy both the hard and soft constraints within a predetermined schedule length. Other proposals have investigated scheduling exams without any assumptions about resource availability. In this work, we address the exam scheduling problem at Al-Hussein bin Talal University in Jordan, which is distinguished by exam types, required schedule length, and resource availability.

3. PROBLEM DEFINITION

The exam scheduling problem can be represented as a directed acyclic graph (DAG), \( G = (V, E) \). Set \( V = \{C \cup S\} \) represents the courses (\( C \)) and the students (\( S \)) registered in these courses. Each course is associated with two variables that represent the number of students registered in this course (\( sn \)) and the exam delivery nature (\( en \)); paper-based (\( p \)) or computerized (\( o \)). Set \( E \) captures the courses registered by each student, where for each direct edge \( e(v'_i, v_j) \), vertex \( v'_i \in S \) and set \( v_j \in C \). Additionally, we are given set \( T = \{t_1, t_2, t_3, .. , t_n\} \) that represents the available exam time slots. Each time slot (\( t_i \in T \)) is associated with an exam room number, day (\( d_i \)). Each time slot has a period \( [ST: ET] \) that represents the start and end time of the time slot. Each exam room has a capacity (\( M \)) that refers to the maximum number of students who can use the exam room at any given time. An exam room could be either a computer laboratory or a standard exam room without any computerized equipment. In presented problem the objective is to determine the courses exams schedule such the operational cost (the number of used rooms) is minimized. In addition to the capacity constraints, the obtained schedule must satisfy the following constraints: i) the exam for each source must be assigned to a single time slot, ii) for any given day, the number of assigned exams to a single student must be at most two, and iii) a computerized exam should be conducted in computer laboratories.

4. ALGORITHMIC SOLUTION

The process of scheduling exams must take into consideration several challenging factors, such as the presence of synchronized courses and the delivery mode of exams (online/offline). Synchronized courses should be scheduled during the same time slot, and the delivery mode of exams has a significant impact on the exam schedule, since online exams can be scheduled in a small subset of the exam rooms. In this work, we present an approach based on profiling to determine the exam schedule. The algorithm presented consists of three main steps: the grouping step, the exam scheduling step, and the exams venue scheduling step.

4.1. Grouping step

The objective of this step is to divide the students into a set of groups, with the aim of treating students registered in a course and all of its synchronized courses as a single group. Figure 1 illustrates the process of the grouping step. It begins with retrieving the most recent course catalog, which displays the updated course names and numbers. Then, in a sequential fashion, starting from the first course, a group is constructed for that course. This group includes all students registered in the course and all of its synchronized courses. For any given course \( c_i \in C \), to identify the list of synchronized courses, we begin by determining the list of all available sections for that course. Then, for each section, we identify the instructor identification number (IID), section time (SI), and location (SL) (lines). This information is used to determine all synchronized courses for course \( c_i \in C \), as any other courses given by the same instructor during the same section time will be considered synchronized courses for course \( c_i \in C \), since any other courses that are given by the same instructor during the same section time will be considered a synchronized course for \( c_i \in C \). The result of this step is a set of groups \( G = \{g_1, g_2, g_3, ..\} \). Therefore, a student \( s_i \in S \) is a member in \( M \) number of groups, where \( M \) is the total number of courses registered by the student in the current active semester.

4.2. Exam scheduling step

Figure 2 illustrates the exam scheduling step process. At most three exam sessions are expected to be scheduled on each exam day. For each session, the algorithm tracks the number of available seats (\( ns \)) in the exam rooms that meet the delivery criteria (paper-based or computer-based). The largest group (\( g_i \in G \)) will be selected and inserted into the schedule on the first available exam session on the first exam day, provided that the available number of seats in this session is greater than or equal to the size of the group. After inserting the group’s exam into the schedule, the algorithm checks whether the current state of the schedule violates the constraint of assigning more than two exams to a student on any given day. If this constraint is violated, the group’s exam is rescheduled on the next available exam day. If there is no available session on the current exam day with a sufficient number of available seats, the group under consideration will be scheduled on the next available exam day. Once the exam schedule for a group (\( g_i \in G \)) is confirmed, the algorithm proceeds to schedule the next group while considering the first available day with a sufficient number of seats as the first possibility for the exam.
4.3. Scheduling rooms step

The objective of this step is to minimize the number of exam rooms allocated to each group \( g_i \in G \). To achieve this objective, several factors, including the number of available rooms and their capacity, must be taken into consideration during the allocation process. Figure 3 illustrates the exam room allocation step. This step works iteratively to determine the exam rooms for each exam session. Starting from the first-day exam session, the allocation process orders the groups allocated to this session in descending order based on their sizes. In addition to the student groups, the available exam rooms are also ordered in a descending style based on their capacity. Then, in a sequential fashion, the student groups are assigned to the exam rooms, with the criteria of this assignment being to allocate the exam room \((s)\) with the highest capacity to the largest.

5. EVALUATION

To evaluate the performance of the presented approach, we conducted an extensive set of experiments using input data representing the students’ registration status during the first semester of the academic year (2021-2022) at Al-Hussein Bin Talal University. The summary of the input data used in the experiments is shown in Table 1. Unless mentioned otherwise, we assume that there are four available exam sessions on each exam day and a total of 50 available exam rooms, of which 10% are prepared for computerized exams and the rest are for paper-based exams. Furthermore, we assume that 15% of exams are for parallel courses, and the percentage of students with more than two exams on the same day is zero in all experiments. We also assume that the total exam period is 14 days in all experiments.

To examine the number of exams allocated per session, we conducted experiments, and the results are shown in Figure 4. Based on Figure 4, the presented algorithm exhibits a greedy behavior, where the number of exams allocated in the morning session is expected to be higher than the number of exams allocated in the evening session. This behavior is desirable when more invigilators are available during morning sessions.
A profiling-based algorithm for exams’ scheduling problem (Tawfiq Alrawashdeh)

Next, we investigated the impact of students on the number of used exam rooms per session. Table 2 shows the number of students scheduled to have an exam in each session, as well as the average number of exam rooms used per session. Based on Table 2, we can see that increasing the number of students does not have a direct linear impact on the number of used exam rooms. The student’s constraint of having only two exams per day, as well as the total number of students per exam, play an important role in determining the number of exams per day, which explains the observed performance.

To study the impact of the number of students on the percentage of students with two exams on the same day, we conducted experiments, and the results are presented in Table 3. Based on the results, we can observe that the percentage of students with two exams on the same day is significantly low throughout the entire exam period. Moreover, we can also notice that there is no direct correlation between the number of students and the reported percentage. To understand this behavior, we considered the mechanism employed by the proposed algorithm. Initially, the expected exams are stored based on the number of students. Thus, exam groups with relatively similar numbers of students are expected to be allocated next to each other in the
Typically, exams for general courses are anticipated to have significantly higher numbers of students compared to specialty (major) course exams. This trend is also expected when comparing the first and second-year courses with the third and fourth-year courses, where exams for later years are expected to have fewer students. Combining this with the fact that each student is expected to be enrolled in major and general courses, the probability of having two exams on the same day is reduced since the groups of the course are distributed across the groups’ list. These factors contribute to the observed behavior.

<table>
<thead>
<tr>
<th>Day</th>
<th>Number of students</th>
<th>Percentage of students with 2 exams on the same day</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2721</td>
<td>0.005145</td>
</tr>
<tr>
<td>2</td>
<td>4788</td>
<td>0.007728</td>
</tr>
<tr>
<td>3</td>
<td>4769</td>
<td>0.077584</td>
</tr>
<tr>
<td>4</td>
<td>3902</td>
<td>0.000256</td>
</tr>
<tr>
<td>5</td>
<td>1876</td>
<td>0.034648</td>
</tr>
<tr>
<td>6</td>
<td>3317</td>
<td>0.000603</td>
</tr>
<tr>
<td>7</td>
<td>1879</td>
<td>0.007451</td>
</tr>
<tr>
<td>8</td>
<td>1879</td>
<td>0.124002</td>
</tr>
<tr>
<td>9</td>
<td>1982</td>
<td>0.005045</td>
</tr>
<tr>
<td>10</td>
<td>3512</td>
<td>0.001708</td>
</tr>
<tr>
<td>11</td>
<td>4161</td>
<td>0.010334</td>
</tr>
<tr>
<td>12</td>
<td>2197</td>
<td>0.019572</td>
</tr>
<tr>
<td>13</td>
<td>2648</td>
<td>0.002266</td>
</tr>
<tr>
<td>14</td>
<td>1654</td>
<td>0.000605</td>
</tr>
<tr>
<td>15</td>
<td>1815</td>
<td>0.002755</td>
</tr>
</tbody>
</table>

6. DISCUSSION AND REMARKS

To analyze the obtained results, the development team has raised several concerns and remarks that we discuss below in detail: Firstly, it is important to ensure the integrity of exams for courses with overlapping material. Some courses are expected to have overlapping material, and if the percentage of such material is relatively high, it can introduce exam integrity issues. Therefore, further analysis should be performed to determine the percentage at which such courses should be treated as a single course. Additionally, as we have discussed, synchronized courses can add pressure to the scheduling process. Hence, eliminating such courses is expected to improve overall efficiency. Secondly, the upper bound of the number of registered students in each division is crucial in determining the quality of the obtained schedule in terms of the number of days and sessions required. Reducing the upper bound of the number of students in each division can lower the number of exams that need to be scheduled simultaneously, prevent overlapping schedules, and minimize the number of required exam days. This can ultimately lead to a more effective exam scheduling process. Lastly, the number of allocated exam halls should also impact the quality of the obtained schedules. Increasing the number of exam halls can ensure that exams are distributed in a timely and efficient manner, thereby reducing the overall duration of the exam period.

7. CONCLUSION

The resources and requirements of universities, including various constraints, have a significant impact on the complexity of scheduling exams. Therefore, in this study, we addressed the challenge of scheduling exams at Al-Hussein Bin Talal University, taking into account several constraints imposed by the university. To tackle this issue, we investigated the grouping approach, which can be considered a profiling mechanism. This approach aims to simplify the problem by scheduling exams belonging to the same group at the same time. The exam strategy employed in the scheduling process minimizes the operational cost while making efficient use of the university’s resources. Our results demonstrated that increasing the number of students does not negatively impact the performance of the proposed approach.

REFERENCES

A profiling-based algorithm for exams’ scheduling problem (Tawfiq Alrawashdeh)

**Tawfiq Alrawashdeh** serves as an assistant professor at the Department of CS, Al-Hussein Bin Talal University, Jordan. He got a bachelor’s in computer science from Al Hussein Bin Talal University (AHU) Jordan in 2005. He then gets a master’s in computer science from Al-Balqa’ Applied Jordan 2011. He received his Ph.D. in 2020 from Universiti Sultan Zainal Abidin, Malaysia. His research interests focus on cloud computing and task scheduling. Tawfiq worked as a Head of programming division at Al Hussein Bin Talal University from 2005 to 2014. Then worked for Al Hussein Bin Talal University from 2014 to present as Lecturer where he teaches a different course in computer science. Tawfiq started working as Director of Computer and Information Technology Center from 2018 to present. He can be contacted at email: tawfiq@ahu.edu.jo.

**BIOGRAPHIES OF AUTHORS**


Khaldun G. Al-Moghrabi serves as an assistant professor at the Department of CIS, Al-Hussein Bin Talal University, Jordan. He received his bachelor’s degree in CIS from Al-Hussein Bin Talal University (2006) and holds a master’s degree in CIS from the Middle East University in Jordan, 2009. In 2018, he earned his doctorate degree in MIS from OIU, Sudan. His research interests include E-learning, decision support systems, database systems, CC, big data, and the IoT. He can be contacted at email: khaldun.g.moghrabi@ahu.edu.jo.

Ali M. Al-Ghonmein serves as an assistant professor at the Department of Computer Information Systems (CIS), Al-Hussein Bin Talal University, Jordan. He received his bachelor’s degree in Computer Science from Al-Hussein Bin Talal University (2004) and obtained his master’s degree in CIS from the Arab Academy for Banking and Financial Sciences in Jordan, 2008. In 2018, he earned his doctorate degree in Management Information Systems (MIS) from Omdurman Islamic University (OIU), Sudan. His research interests include information retrieval, decision support systems, natural language processing, database systems, cloud computing, big data, and the IoT. He can be contacted at email: ali.m.alghonmein@ahu.edu.jo.