EXAMINATION SCHEDULING SYSTEM USING GENERIC ALGORITHM

BY

MENDES OLUWADAMILARE P/HD/20/3210031

TOMORI OLAMILEKAN P/HD/20/3210019

ONYINYECHI MARY NWANKWO P/HD/20/321009

BEING A PROJECT SUBMITTED TO THE DEPARTMENT OF COMPUTER TECHNOLOGY, SCHOOL OF TECHNOLOGY, YABA COLLEGE OF TECHNOLOGY,
YABA, LAGOS.

IN PARTIAL FULFILLMENT OF THE REQUIREMENT REQUIREMENT FOR THE AWARD OF HIGHER NATIONAL DIPLOMA IN COMPUTER SCIENCE.

CHAPTER ONE

1.0 INTRODUCTION

Time-table is a structured schedule of an event with time at which they occur, especially time of arrival and departure according to Adedokun J. B. (2014).

A time-table is an organized set of courses in a tabular form, which makes information available about a series of arranged events such as the time range for the event to take place, venues, invigilators and supervisors that will monitor how the event is carried out.

Examination scheduling system is an important and one of the recurring administrative activities in almost all educational institutions. This system helps in arranging examinations for students which determines when, where and how examination is to be conducted. Creating a good examination timetable system that will satisfy students, lecturers and the institution management is a very difficult task due to the limited resources (Burke and .Newall, 2014).

Examination scheduling is a widely studied type of timetabling problem due to the difficulties of developing an appropriate examination timetable system for high institutions. An examination scheduling system should ensure that both students, lecturers and the school management are satisfied with the examination arrangement, making sure that all examinations are scheduled within the stipulated time frame and both hard and soft constraints are well satisfied.

At its most basic, the exam timetabling problem is concerned with distributing a collection of university exams among a limited number of timeslots (periods). This is, of course, subject to a set of regulations and limitations (often termed constraints), which vary widely from institution to institution. There are certain constraints which must be satisfied under any circumstances such as the

requirement that no student can sit two exams simultaneously, or that exam rooms have a certain physical capacity which must not be exceeded. Such constraints are known as "hard". Solutions, which satisfy all the hard constraints, are often called "feasible" solutions (Qu , 2018).

soft constraints in examination scheduling can reflect the situation where students prefer to spread exams evenly throughout the examination session or at least have some time interval between exams. On the other hand, the institution often wants to schedule large exams earlier (in order to leave more time for marking). Specific preferences may also be expressed by particular members of staff concerning, for example, invigilation duties. Of course, in any real world situation it would be extremely rare if it were possible to satisfy all the soft constraints. Therefore, a useful measure of the quality of a timetable can be taken to be the number of violations of these constraints (Burke, Petrovic 2012).

The difficulties of developing appropriate examination time tables for institutions and tertiary schools is increasing. Institutions are enrolling more students into a wider variety of courses in many different fields.

One of the major approaches in exam timetabling over the years has been constraint programming approach (Simulated Annealing). This method of constraint programming logic language also to solve exam timetabling problems. Constraint programme phase provides an initial solution and a simulated annealing phase to improve the quality of solutions (Duong and Lam, 2004).

1.1 BACKGROUND OF THE STUDY

The need for an automated information system for educational development cannot be over emphasized. The assertion by many professionals that information is power is continuously gaining affirmation as the countries of the world continue to witness dynamic trends in information technology. The manual system of

accessing and scheduling of an examination time-table in higher institutions is always very tedious. The scheduling is done by the school's exams and record department. The number of exam halls in the school and their various capacities must be put into consideration, if not there will be a clash of exam venues. For each of the sections, levels and departments, there will be a time slot for each of the examination days. Manually scheduling is confusing and takes a long period of time to produce an excellent and perfect result that is free from clashes.

1.2 STATEMENT OF PROBLEM

School examination time table has been the only way of informing students of higher institutions about their examination days and time.

The following are some of the challenges encountered in the Manual Process of Exam Time Table Allocation;

- The current manually prepared exam time table is always error prone and takes a lot of time to reschedule.
- Manual system of planning and accessing this timetable is stressful.
- Actual class sizes are not captured and as such there is error in Venue Allocation.
- Clashes take place in cases of carryover students.

1.3 AIM AND OBJECTIVES OF STUDY

The principal aim of this project work is to develop examination timetabling software that will be useful to our education institutions, using Generic Algorithms.

The following are the set objective

- Exploratory study of Generic Algorithm and simulated annealing in resolving conflicts associated with exam timetabling.

- Develop a computer software to automatically generate exam timetable using the appropriate algorithm
- To evaluate and compare the performance of the algorithms in terms of their computational complexity.

1.4 SIGNIFICANCE OF STUDY

This research work is greatly hoped to eliminate the manual way of scheduling examination time tables in the polytechnics and colleges. It will also eliminate stress in planning the time table. It will eliminate examination clashes.

1.5 SCOPE OF STUDY

In this project work, attention is focused on formulating mathematical models for the examinations timetable for Computer Science, Yaba College of Technology. This will act as a benchmark for testing heuristic algorithms and help future reformations of the problem models.

1.8 DEFINITION OF TERMS

Scheduling : Scheduling is the process of arranging, controlling and optimizing work and workloads in a production process

Software: Software is a set of instructions, data or programs used to operate computers and execute specific tasks.

Algorithm: a process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.

Slots: A space where some information will display such as Time, Courses, Invigilators.

Optimization: the design and operation of a system or process to make it as good as possible.

CHAPTER TWO
LITERATURE REVIEW

2.0 INTRODUCTION

Examination timetabling is one of the most important administrative activities that takes place in all academic institutions. In this project paper work we present a critical discussion of the research on exam timetabling in the last decade. This last ten years has seen a significantly increased level of research attention for this important area. There has been a range of insightful contributions to the scientific literature both in terms of theoretical issues and practical aspects. The main aim of this survey is to highlight the new trends and key research achievements that have been carried out in the last decade. We also aim to outline a range of relevant important research issues and challenges that have been generated by this body of work.

According to Kong, S. C. & Kwok, L. F. (1999), timetabling system involves a heuristic function to increase the scheduling performance, as well as producing a best outcome.Currently, the well known solutions for the timetabling system are Genetic Algorithms and Memetic Algorithms (Mohd. Dain, A. A., Shaari, N. S. Gom, Y. S. & Bacheck, Z. A., 2001). However, Causmaecker, P. D. and his friends introduced the Semantic Web as a solution in the domain of timetabling. Berger, J. & Barkaouia, M. (2002) also introduced a Parallel Hybrid Genetic Algorithm for the vehicle routing problem, which they argue to be faster, more cost-effective and highly competitive than the best-known heuristic routing procedures and solutions. Obviously, researchers are still looking forward to heuristics that are suitable for their particular problems (Causmaecker, P. D., Demeester, P. & Vanden B. G., 2002). On the other hand, there are many solutions for the timetabling system, while each of them has their strengths and weaknesses.

2.1 DEFINITION OF TIMETABLE

Souza (2000) defines a timetable as the total schedule of a specific teaching, period of work or time to do a specific task, which will be attended by a group of students and the lecturer, at a specified time. He further stated that it requires specific resources such as venue, examination aids such as papers-answer script and question papers, probably writing material. Scheduling is parallel to the resources made available besides fulfilling other needs. In other words, timetables can also be referred to as a schedule of events that guides a school activities throughout the day, week, term or year. CommonWealth of Learning(2000). Furthermore it stated that for each activity, a timetable generally specifies a starting time and an ending time and in some cases indicates who is involved and how the activity will be conducted.

The resolution of the examination timetable problems can be claimed by different areas, such as the School Administration, Artificial Intelligence, Mathematics or Operational Research (Burke, et al., 2016).

Probably, we must appeal to the techniques of simulation imported from fields as diverse as physics or biology, to solve the problem (Zhu and Tha, 2012). An examination takes place in a given time slot, using a set of rooms and a set of invigilators (José, 20016). Each invigilator is assigned for a given room, although several invigilators can supervise the same room. Each exam is scheduled into one slot variable. For each exam, one or several exam rooms are allocated and for each room, a set of invigilators is defined. This suggests a three phased approach, and as a result we have different types of variables: Time Slot Variable, Room Variable, and Invigilator Variables. Timetabling problem in educational institutions is categorized in three groups: institution examination, institution course, and school timetabling. Institution examination timetabling defines the exact day, time slot,

and room that each exam is held such that either no conflicts or a minimum number of conflicts occur. The main objective in the examination timetabling problem is no student takes more than one examination at any time period. This conflict can be regarded as a hard constraint and must be eliminated.

2.2 BRIEF HISTORY OF TIMETABLE

Timetabling problems arise in various forms including educational timetabling (e.g.), nurse scheduling, sports timetabling and transportation timetabling. They have represented a challenging and important problem area for researchers across both Operational Research and Artificial Intelligence since the 1960s. Recent years have seen an increased level of research activity in this area. This is evidenced (among other things) by the emergence of a series of international conferences on the Practice and theory on automated Timetabling, and the establishment of a EURO (European Association of Operational Research Societies) working group on automated timetabling. Burke, Kingston and de Werra (2004) gave a definition of general timetabling, which covers many cases:

A timetabling problem is a problem with four parameters: T, a finite set of times; R, a finite set of resources; M, a finite set of meetings; and C, a finite set of constraints. The problem is to assign times and resources to the meetings so as to satisfy the constraints as far as possible. Among the wide variety of timetabling problems, educational timetabling is one of the most widely studied, from a practical viewpoint. It is one of the most important and time-consuming tasks which occur periodically (i.e. annually, quarterly, etc.) in all academic institutions. The quality of the timetable has a great impact on a broad range of different stakeholders including lecturers, students and administrators. Variants of educational timetabling include school timetabling (class-teacher scheduling), university course timetabling, exam timetabling, and faculty timetabling and

classroom assignment. It has been observed that course and exam timetabling are relatively close problems but very significant differences do exist. This survey will concentrate on examination timetabling.

An excellent survey of examination timetabling was published in 1986 and an insightful follow up paper appeared in 1996. However, a significant number of research papers in the area have been published since 1996. This paper will concentrate upon the research that has appeared since the publication of. The last decade has seen the establishment of a collection of benchmark exam timetabling problems which have been used by many of the examination timetabling research papers that have appeared since 1996. Moreover, there has been some confusion in the literature caused by the existence of different benchmark problem datasets with the same names. This paper aims to eradicate such confusion by presenting a definitive re-naming of the sets and by clarifying the situation over which papers dealt with which problems.

2.3 MACHINE LEARNING

Machine Learning is one of the branches in artificial intelligence that handles the automatic learning of systems in a way to improve the accuracy and quality of prediction, improvising the experienced learning process. In Machine Learning, learning indicates the methods to understand and recognize the data input to the system as well as in making right decisions at the right time as a brain to human beings. Even though the machine can learn inputs and replicate them, it is very hard for machines to understand all inputs based on human emotions and indications. Hence, algorithms pave the way for machines to understand and learn knowledge with particular data input from dataset using a past experience with statistical principles and theory of probabilities.

Machine language being a wide and vast area of study is segregated into two major categories of information;

i.Unsupervised Machine Learning

Unsupervised Learning methods involve no predefined dataset called training dataset and they build patterns for futuristic predictions. It is an extremely powerful technique used in artificial intelligence and robotics where machines learn by themselves. They are capable of predicting the future, a forecast unknown and unlabeled data without any solid inputs. It normally splits the data into logical groups of information and gives results based on imminent perspectives. Some of the unsupervised algorithms include simple K-Means, Fuzzy C-Means and hierarchical clustering, etc.

ii. Supervised Machine Learning

The Supervised Machine Learning Algorithms enable the learning process of machines using a function that is predefined and made available by the developer using an algorithm. The algorithms used in supervised model gains from the training data used for prediction and after learning, produces inference solutions that can further be mapped to new dimensions of data within the boundaries. When supervised, it can produce results within the limit of inputs and will not predict knowledge that is unknown. Some of the supervised algorithms are used in searching for contents in the internet, in classification of emails as inbox and spams files automatically, in content labeling, in voice and speech recognition from different input sources

2.4 WHAT IS AUTOMATION?

Automation can be defined to be the use of computers to control a particular process in order to increase reliability and efficiency, often through the replacement of employees. For a manufacturer, this could entail using robotic assembly lines to manufacture a product. In this project research work, the automation aims at eliminating the manual way of preparing examination time tables by making use of an automated Time Table Allocation System. In automating the manual way of preparing time tables, there are many approaches that one can follow in order to automate timetabling.

2.5 APPROACHES USED IN AUTOMATING TIME TABLE

2.5.1 Heuristic approach

The heuristic approach is one of the first automation approaches used in scheduling problems. This approach attempts to obtain the best solution selection despite the fact that there is no guarantee to that effect. Looking for the best solution will take a considerable time as problems pertaining to scheduling have many probability which must be evaluated.

Through this approach, the number of probability may be limited so that a solution may be made within an appropriate time. It attempts to arrange venues suitable to the students that no department will have exams in two venues at the same time slots no level- HND1 and HND2 will have examinations at the same time slot.

Bresina 2006, was among the early researchers who used this approach. He made several modifications in the manual approach conducted at universities. It commences by allotting the combination of department, level and courses without the existence of any conflict between the resources available. In this approach, the search process will allot day and date first. In the event that there are courses that are not allotted, a change process between the courses will be done. This is done by

shifting the courses that are already allotted to a time slot that may be suitable to enable the said new courses to be allotted a time slot.

2.5.1.0 Heuristic approach Methods

i. Simulated Annealing

Simulated annealing is a variant of the well known modern heuristic strategy for finding good solutions to a global stochastic optimization technique that has been widely used in several types of combinatorial optimization problems. It is a variant of local search which allows up till move to be accepted in a controlled manner.

The Simulated Annealing theory has been investigated over the last few years for exam timetabling with some level of success. In 1991 Abramson applied simulated annealing to construct high school timetables. In 1996 Thompson and Dows land considered an adaptive cooling technique depending upon the success of the move the two authors employed Kempe chain as neighborhood structure.

In 2001, Burke et al proposed a time Predefined variant of simulated annealing.

The basic significant amount of heuristic techniques such as genetic algorithm tabu search and simulated annealing. Some of the earliest practical application of simulated annealing is on very large scale integration. Design packages of thunderbird and thunder wolf. Recently (Kelly etal, 1999,) applied simulated Annealing (SA) to the controlled rounding problem in the three dimensions.

ii. Genetic Algorithm

Genetic algorithms were invented by John Holland in the 1960s and were developed by Holland, his students and colleagues at the University of Michigan (Mitchell 1991). GA is an optimization technique based on nature evolution. It maintains a population of strings called chromosomes that encode candidate solutions to a problem. The algorithm selects some parent chromosomes from the

population set according to their fitness value, fitness values are calculated using fitness function, which performs a major role in Genetic Algorithm.

2.5.2 Integer Programming Approach

Linear programming is a mathematical technique to aid in obtaining the most effective decision by utilizing all available resources. In general definition, it is also known as mathematical programming. Integer programming yields solutions for linear programming hitches. Integer programming faces the same setback as linear programming because nearly every variable must be optimized with positive integers.

For example, in the complication of linear programming, the variable can use fraction numbers (not round numbers). However most of the difficulties in the scheduling issue uses variables with integer value. Another example is a venue must comprise a single department and not a fraction of department. The definition of linear programming:

Maximizing (or minimizing) one linear function to a definite variable. The said functionality either maximizing or minimizing is known as objective function.

The definite variable value must adhere to existing constraints. Each constraint must be either linear equation or linear irregularity.

All variables must not have negative values. Negative value for physical quantities such as department, venue and courses is not permitted.

Integer programming undertook the attempt to solve scheduling issues.

Dahal et, 1997, defined that problem must be divided into two portions which is courses collection against time followed by unification of courses and venue.

Difficulties faced when using this approach is the event of too many or all variables in the formula is integer in value and normally the size of the problem is big. It involves extensive calculation time and probability to solve the situation

without utilizing a mathematical approach. It further stressed that each problem must be solved mathematically, the number of variables and constraints will be too complicated to manage when facing a large problem.

2.6.3 Logical Constraints Arithmetic Approach

Logical constraints arithmetic is a technique that inherits some similarities with Al-Khwarizmi's Heuristic. Waterloo PROLOG (WPROLOG) is used by some universities' scheduling systems without having any conflict. Logical programming was used to represent the constraints while pattern matching and backtracking was used for the search engine.

The timetable was divided into two sub problems. First, the scheduling for the courses, day, and venue- is given a time slot and secondly the venue scheduling will be joining the previous state. The use of logical constraints arithmetic could ensure that all high constriction conditions were fulfilled to ensure that the timetable is feasible.

Concentrating on the venue scheduling, the venue that suits the needs for any configuration, faces some possibilities of not being available. However, bigger venues may have the probability of vacancy. Venues are reunified by size and are given a running number. Those courses that are unsuccessful with any allocated venue can be reconsidered to loosen the size constraints by allocating bigger room size with bigger running number. A mechanism known as size matching will determine the maximum venue size for any course that may be requested.

On the other hand, De Werra (1995) employed a two-execution phase in the attempt to minimize searching difficulties. Problems are broken into subproblems of ND1 and HND1 and ND2 and HND2. In the first phase, courses will be predetermined for its days and the second phase will allocate time for the selected days. If the situation cannot be resolved, the first phase will be repeated and a

different day will be selected. Priority is given to a time and selection of time slotting activity for allocating a reasonable time slot. Knowledge base defines that every activity and also time slots are represented separately. Every time slot has a knowledge base. It will be revoked during the joining or allocation process or a process to resolve any issues. This knowledge base will determine whether the time slot can be used or not. The knowledge base is divided into several resolution constraints mechanisms.

Knowledge base manipulation is mentioned through the usage of courses shifting and courses exclusion. Exclusion is a mechanism to broaden the searching scope utilized in the paradigm.

Resolution constraints mechanism is used to show how the time slot can be implemented to activities that are yet to be combined. If the time slot cannot be used for any activities, others that share the same resources are extracted to enable the first activity to be joined.

This is for the purpose of shortening the search extent due to the previous joining containing limited relationships between the activities. The previously extracted activity will be allocated to another time slot.

2.6.5 The greedy algorithm

The greedy algorithm, with a simple priority function, dictates that all subjects are sorted according to the number of students in the subject from large class size to small class size. The algorithm will assign rooms for each day by separating the morning section and afternoon section. At each section of each day, all subjects of that time are sorted according to their class size from large to small size. The available rooms of that time are also sorted according to their capacity from large rooms to small rooms. The algorithm then maps the large subjects to the large room. However, the whole subject might not fit in one room because some seats

must be assigned to other subjects. The algorithm will search and assign the next subject to the room until the room is filled. Then, the algorithm moves to the next room and performs the same steps. The algorithm stops when there is no subject left to be assigned or there is no seat available. Either case, the algorithm will report to the user. Once the room assignment is done the seating layout of each room must be produced or printed. All results are stored in the database so that they can be viewed, printed, edited or retrieved later as the case may be by the authorized users (Prabnarong and Vasupongayya, 2011).

2.7 REVIEW OF RELATED WORKS

Abramson 1991, implemented Al-Khwarizmi Genetic for solving problems in timetabling at university. Cost is accounted for from the total collision in the timetable and if cost is empty (no collision), therefore the timetable is accepted. The cost for time includes courses, level and venue are given different weight age suitable for its importance. For instance, the optimization problem is to minimize the cost by lessening the amount of collision occurring in the timetable. Timetable is represented with a set of diagraph (a combination of level, venue, courses and department) in which will be joined with time to create a complete diagraph.

Sagir and Ozturk, 2010 tried to address the issue of assigning invigilators to examinations who may be faculty lecturers or graduate students in educational institutions with a computer based system. They adopted the Analytic Network Process model for estimating exam and objective weights The model assigned the invigilators using arbitrarily assigned ratio scale numbers for the exam weights. The model was presented in two directional novelty.The first one is that a course can be offered more than once in a semester. If a course is requested by a few students, then it is enough to be offered once. If the number of students requesting

a course is more than the maximum number of students who are allowed to attend a single class, then the course is multi-offered. The second novelty is that sharing a room for two simultaneous exams is allowed. Also, the model considers some hard and soft constraints, and the objective function is set in such a way that soft constraints are satisfied as much as possible.

Nasser R. Sabar and Masri Ayob Rong 2011 investigated a new graph coloring constructive hyper-heuristic for solving examination timetabling problems using hierarchical hybridizations of four low level graph coloring heuristics. The solution lists all examinations by the level of difficulty of assignment and ensures the most difficult exam to be scheduled is scheduled first. This method was considered a simple and yet efficient approach but suffers setbacks in handling higher level conflicts in examination scheduling. Graph coloring has the limitation-which neglects hard constraints such as students cannot take two examinations consecutively by formulating the basic examination timetabling problem as a Quadratic Assignment Problem. The Aldy model assumes that the number of examinations is equal to the number of time periods and a hybrid algorithm based on a combination of Greedy Randomized Adaptive Search Procedure and Simulated Annealing and Tabu Search was introduced to get an optimal solution. This solution appears to be an efficient heuristic algorithm for solving examination scheduling problems for its ability to obtain the optimal or the best known solutions within reasonable computation time. One of the major limitations of this solution is its inability to handle a very large size problem and has not really been applied to real world problems.

2.8 RESEARCH GAPS

Authors	Article Title	Methodology	Results
Dimopoulou and Miliotis (2011)	Implementation of a university of a university course and examination time-tabling system	Integer programming	Restricted availability of classrooms Increase flexibility of the student's choices of courses
Alireza Rashidi and Mehradad Nouri Koupaei (2012)	A new binary model for university examination timetabling	Binary Model	Multi-offered courses in a single semester Room sharing for two simultaneous exams
Sagi and Ozturk 2010	Exam scheduling: Mathematical modelling and parameter estimation with analytic network	Process approach	Analytic network process model Invigilators assignment
Salem M. Al-Yakoob, Hanif D. Sherali, Mona Al-Jazzaf 2012	A mixed-integer mathematical modelling approach to exam timetabling	Mixed integer	Assignment to designated exam periods and classrooms Gender base Proctor assignment problem