Automatic Timetable generation using Permutation and Decision Making

1. Amit Kumar, computer dept, DIT, Pune
2. Anuradha Singh, computer dept, DIT, Pune
3. Harshveen Khanduja, computer dept, DIT, Pune
4. Apoorva Vibhute, computer dept, DIT, Pune
5. Prof. Rachna Somkunwar ,computer dept, DIT, pune

Abstract

Assigning timetable for all the faculty to finish the syllabus in time is always a tough task. This is mainly because of the problem of assignments of the subjects to the faculties in a given stipulated time. The reason behind this is, every teacher is loaded with a certain amount of workload hours every day, and in some scenarios this workload is being distributed in shifts. So for managing time, according to each of the staff's requirements while creating the Timetable is really a cumbersome task. The major hurdle that is facing in this task is overlapping of the teaching hours. So a very few existing software tools are there in online which are eventually creates automatic time table based on the input of the staff load. But all of them are very costly or they never reveal their characteristics to the before buying. So to introduce a powerful and cost effective method of generating the time table for the input details like number of allocated hours for theory and a number of allocated hours for lab for a given teaching staff is the much needed thing in the research. So this proposed system puts forwards an idea of using the Shannon Information gain theory and Data permutation technique for the synchronous data allocation process. And this idea is catalyzed by decision making process which eventually yields a good Timetable without overlapping of the scheduled time.

Keywords: Timetable. Shannon Information Gain, Decision making, Synchronous data allocation. Decision making.

I. INTRODUCTION

Timetable Generation is one of the oldest problems faced by academicians and managers of a company. It is essential to optimize the assignment of jobs to the machines, jobs to the workers and classes to the teachers so as to minimize and reduce conflicts and to be able to tread the path of least resistance. The problem of timetabling is not new and has influenced the researchers since time immemorial.

To achieve optimum timetable, even nowadays, the predominant technique is doing it manually. The assignment of tasks and other jobs manually introduces a lot of discrepancies in the process and it is more prone to the introduction of human error in the process. there has been a lot of research in this area but nothing concrete has been presented by the authors of those publications to reduce the burden as many universities and colleges employ manual timetabling.

This paper utilizes various techniques for solving the creation of a timetable. The technique proposed in this paper utilizes, Entropy estimation, Shannon Information Gain with the help of Genetic algorithm and Hungarian task allocation to achieve an innovative solution to this problem.

Entropy is a term used in physics as the measure of randomness. Randomness is explained in a very elaborate fashion in the laws of physics. In the beginning of time and our universe itself, all the matter in the universe was collected together in one point just before the big bang happened. At the moment when all the matter was

concentrated in a point, the entropy was 0, that would mean very little information can be used to describe the information it contained.

But right after the big bang, the entropy of the universe increased, that is the information required to describe the system with the plethora of stars and galaxies that are in this universe. The entropy of the universe is ever increasing and will keep increasing continuously.

Entropy estimation in computer science is defined as the measure of messiness or randomness in a given set of data. This is usually accompanied by some amount of information gain as the randomness adds quite a lot of information by increasing its entropy. The entropy used here is Shannon Information Gain. It is a widely used mechanism for the classification of data.

Shannon Information Gain is one of the most commonly used algorithms for the classification of data for the Data Mining procedure. The Shannon Information gain is very similar to the concept of Entropy, which is a compelling measure of disorderliness or the messiness of data. Classification algorithms have a very distinct job to perform, that is splitting the data into smaller subsets. This is done with the idea that the messiness will reduce with the further division. Therefore, the Shannon Entropy has one central goal, that is the goal of any classification algorithm, to reduce the entropy of the whole data.

This research paper dedicates section 2 for analysis of past work as literature survey, section 3 deeply elaborates the proposed technique and whereas section 4 evaluates the performance of the system and finally section 5 concludes the paper with traces of future enhancement.

II. LITERATURE SURVEY

This section of the literature survey eventually reveals some facts based on thoughtful analysis of many authors work as follows.

Y. Rehman [1] explores the realm of education by isolating the most common problems found in laboratory teaching in engineering universities, which impact the overall understanding of a student. The authors performed an experiment for the assessment of the instructors as the academic growth of a student is highly dependent on the effectiveness of the instructor. The authors, therefore, observed the results that prove that it is indeed true. The researchers also proposed a model for a timetable that can segregate the resources efficiently and achieve lab parallelization.

R. Badoni [2] states that the UCTP or University Course Timetabling Problem is widely used for scheduling and allocating various rooms according to the student's needs and is maximized to achieve high efficiency. Sometimes, there have been some inconsistencies that have crept into the scheduling process and that has led to a decrease in the efficiency of the system. Therefore, the authors have proposed the addition of the ACO or Ant Colony Optimisation algorithm to the system which eliminates all the drawbacks in the system.

J. Zhong [3] expresses that the problem of railway timetable scheduling is one of the most basic problems faced by the railway industry. As the railways are needed for maintaining the quality of the service for the transport system, it is imperative to solve the problem of the railway timetable scheduling efficiently. Therefore, the authors have attempted to solve the problem by introducing the enhanced Differential Evolution algorithm or the DE algorithm. this algorithm eliminates all the inconsistencies of the previous PRTS versions and reduces the wait time for the passengers drastically.

S. Ribic [4] introduces Integer Linear programming as avery popular technique for the construction of school and university timetables. As most of the timetabling software is usually designed for use in making timetables for the university, as that is a lot more complicated than thegeneration of timetables for the school. The timetables are varied in nature and differ from school to school and country to country; therefore, it is difficult to achieve a generalized constraint for each and ever implementation. Therefore, the authors have proposed a timetabling algorithm based on integer linear programming that meets all the constraints in different school timetabling systems.

S. Limanto [5] explains that the examination of a thesis is one of the most important aspects of graduation. There are extensive techniques for the examination of a thesis which starts from the organization of the examination timetable, which is still being according to the traditional standard scheduling techniques. Due to techniques being very old and dated, they are not as efficient and fast. Therefore, to ameliorate this effect and incorporate all the constraints, the authors have implemented a web-based system that utilizes the genetic algorithm for increased efficiency in the timetabling process.

X. Yang [6] states that the UTCP or the University Course Timetabling Problem is a complicated problem that relies on the organization of the time-slots of the lecturers and the classrooms and laboratories to get an efficient combination of them to flow accurately. The area of UTCP has been a hot topic for r3esearchers for a long time and a lot of solutions were suggested on this problem. There has been an increased amount of interest in solving this problem as it can be utilized for different applications outside of the university and be applied to solve some practical problems.

M. Almeida [7] expresses that the process of generation of an academic timetable is one of the most difficult of tasks in the area of scholarly planning. This is due to the fact that it requires a lot of time and resources to also allocate and cannot be done if the professors are busy with the workload. Most of the time timetabling technique has largely been done manually which is a very time and resource intensive approach. As most of the techniques that have been proposed rely on optimization techniques, but the authors believe this problem could be solved by using heuristics and achieve a better output.

S. Awang [8] stresses the importance of the generation of a timetable as it is one the most widely used methods for the organization of an event or a day. For a university, there are a lot of constraints and a lot of data that needs to be considered for the formulation of a timetable. The method of reliable generation of a timetable is one of the most demanding tasks at it requires a lot of time and resources. Therefore, the authors have presented an innovative technique for the generation of the timetable by utilizing random selection and heuristic. This combination is one of the most reliable and fast techniques for the generation of the timetable.

Y. Ting-Hong [9] investigates the urgency of the timetabling for the purpose of course planning, as it is very complicated to generate a working model of the timetable. There are a lot of requirements that need to be fulfilled before planning a course timetable. The authors developed a technique that can calculate the urgency of a particular timetable in the course planning. As the urgency of the timetable task is calculated, it can be used to find the urgency function of space and time. All of these predicted values are then adjusted to the specific college through an analytic hierarchy process. This generated a generalized form that can be used on any application or college.

F. Guo [10] introduces an innovative technique for the process of generating timetable with the help of reinforcement learning. The authors have stated that there are various problems that are faced by a person creating the timetable for the first time. To eliminate various difficulties that are faced when altering timetable scheduling action vector and timetable eigenvector, the researchers have utilized a Naïve Bayesian classification algorithm for mining the historical data efficiently. The proposed system can generate and schedule timetables very efficiently in comparison with the conventional approach.

P. Kaewchanid [11] presents a unique technique for the teaching assistant scheduling problem with the introduction of the constraint-based approach. This is one of the classical approaches used when generating a timetable manually. The authors have also implemented an EDF (Earliest Deadline First) and cluster method to group the resources and to schedule the task with the shortest deadline first. This leads to a lot of bottlenecks and problems and the authors have implemented the constraints to eliminate the problems and perform the timetabling efficiently and faster in comparison with other conventional techniques.

III PROPOSED METHODOLOGY

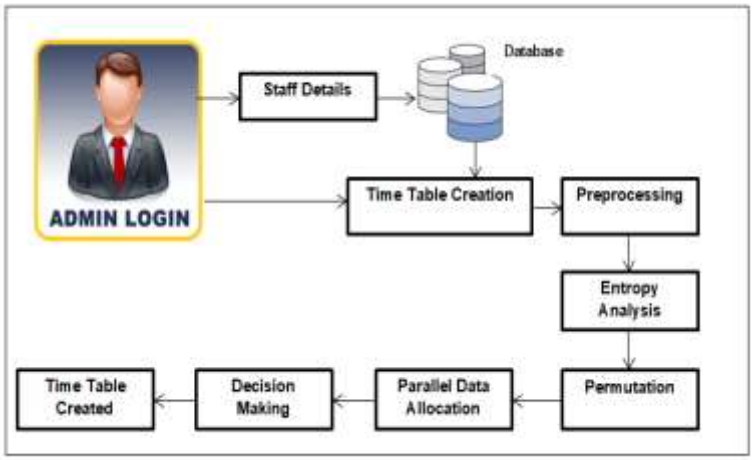

Figure 1: The Overview of the proposed methodology

The proposed methodology for the automatic timetable generation overview is depicted in the figure 1. The steps that involve in making of timetable is broadly narrated in the below mentioned steps.

Step 1: Data Collection and Preprocessing - This is the primitive step of the proposed model where an Administrator of an Engineering college who is in charge of handling the automatic timetable creation software feed all the details of the staff into the system. The details are including the Staff Complete name, Staff name abbreviation, Staff handling theory subject name, Theory subject abbreviation, subject load per week in hours, practical subject name. practical subject abbreviation, practical subject load per week, Semester and the Division.

Here the proposed model is designed to generate the timetable for 3 Even semesters for the Computer science department. That includes the semesters like 4th, 6th and 8th with two Divisions A and B. Once admin enters all the data into the system, then all the input data is properly validated and stored in the database.

Here proposed model is designed for considering a constraint of 7 working hours a day and five days a week. Where 5 hours for theory and 2 hours for practical is assigned for the day. So after feeding all the staff details to the system, the Administrator will ready to generate the Timetable automatically. In this process of automatic generation of the timetable initially all the details of the staff from the database is fetched and stored in a double dimension list. And then this list is subject to create the timetable.

Step 2: Entropy Analysis - Here in this process each and every staff from the database is listed for the separate list for Theory time table generation and practical time table generation to merge both in the end. For this purpose a double list is created which contains the columns like semester, division, Staff abbreviation and theory load. And another list with the practical details.

Once this detail is being collected in the list, then the unique staff name is collected in a list by subjecting all the name to hash set function. Then each of the unique staff is counted for his name in all the semester for both the divisions. So in this process total 6 classes are there like 4A,4B,6A,6B,8A and 8B.

So a count is estimated for a staff for all this 6 classes to evaluate the information gain value which will be in the range of 0 to 1 as shown in the equation 1. Then all the staffs are labeled with this information gain value to sort them in descending order using the Bubble sort technique. So the staff who is having highest gain value holds the top position and this indicates that, the staff is having theory or practical load in almost all the classes. So the proposed model gives first priority to schedule his class timings which generating the timetable.

$$IG = -\frac{P}{T}\log\frac{P}{T} \ -\frac{N}{T}\log\frac{N}{T} \quad \text{_____(1)}$$

Where

P= Frequency of the staff in all classes
T= Total number of classes that is 6.
N= T-P
IG = Information Gain values

Step 3: Permutation, Parallel Data Allocation and Decision making - This is the core step of the proposed model where a timetable is generated automatically. This process is started by creating 6 5 X 5 string data matrices each for one class. Here 5 X 5 indicates five theory hours for a day, for 5 working days from Monday to Friday. And this all the String matrices are stored in the list.

Here in the process a matrix is considered for the size 5 X 5 for each of the class simultaneously. Then a staff abbreviation is considered along with his theory load. Then the staff is trying to allocate in that position of the matrix by simultaneously checking his abbreviation name in the same position in all other classes matrices.

If the staff is present in that position or time slot, then another staff is selected to repeat the same thing to find a staff who is not at all allocated in that position of the matrices for any the class. Then a place will be found for a staff for the given position, then that staff is placed in that time slot and then his/ her count for the work load will be reduced by a unit. This will be repeated for all the staff till the complete matrices are filled with unique staffs that eventfully indicated the time slot.

In the similar way a another set of matrices is created for the practical subjects and then both are merged to form a desired format of the Timetable. This contains the subject name in abbreviated form, Staff name in abbreviated form and class name. Which is eventually free from any of overlapping issues. This process is depicted in the algorithm 1.

Algorithm 1 : Timetable Generation

// Input : Staff Detail set S_D = {Staff name S_N, load L_D}
Class List of Matrices of ORDER 5 X 5 C_L
// Output : Timetable Matrix T_M
Function : timeTableGeneration(S_D)
Step 0: Start
Step 1: T_M = Ø
Step 2: ***for*** i=0 TO 4
Step 3: ***for*** j=0 TO 4
Step 4: ***for*** k=0 to SIZE of S_D
Step 5: T_{SET}=S_D[k]
Step 6: $T_{SET\ [0]}$= S_N , $T_{SET\ [1]}$= L_D
Step 7: ***Flag***=isStaffNotExisted(i,j,S_N))
Step 8: ***if Flag*** = TRUE, **THEN**
Step 9: $T_{M[i,j]}$ = S_N, SUB
Step 10: L_D =L_D - 1
Step 11: update L_D into S_D
Step 12: **ELSE**
Step 13:T_{SET}=SD[k+1]
Step 14: Repeat STEPS 6 TO 13 TILL ***Flag*** =TRUE
Step 15: **End *for***
Step 16: **End *for***
Step 17: **End *for***
Step 18: return T_M

Step 19: Stop

IV RESULT AND DISCUSSIONS

The methodology for the creation of enhanced electronic lecture time-table scheduler has been coded in Java programming language on a NetBeans 8.0 Integrated Development Environment. The database capabilities have been handled by a MySQL database server. The machine on which the proposed technique has been implemented runs on Windows Operating System and is powered by a Core i5 Central Processing Unit and 6GB of Physical Memory. To successfully evaluate the effectiveness and accuracy of the proposed system, various experiments are conducted.

One of the most widely used techniques for the valuation of error is the Mean Absolute Error or (MAE). This technique is capable of evaluating the margin of error in terms of percentage and provides an effective and robust technique for the evaluation of performance for the presented technique. The MAE parameter has been utilized due to the fact that the parameters in this technique are continuous entities which is a similar phenomenon.

The percentage error is evaluated in this technique of Mean Absolute Error for the number of accurate timetables generated. The Mean Absolute Error is evaluated on the absolute difference betweenan actual number of trails and the number of accurate timetables generated through the Equation given below.

$$\text{MAE} = \frac{(\sum_{i=1}^{n} | xi - yi |)}{n}$$

Where,

xi - Number of Actual Trials Conducted.
yi - Number of Accurate Timetables Generated.
n- Number of Experiments Conducted.

Table 1: Measurement Data for MAE.

Experiment Number	Number of Trials Conducted (xi)	Number of Accurate Timetables Generated (yi)	Difference (xi-yi)
1	10	9	1
2	10	10	0
3	10	10	0
4	10	10	0
5	10	9	1
6	10	10	0
7	10	9	1
8	10	10	0
9	10	10	0
10	10	10	0
		MAE	0.03

Table 1 above discloses some important aspects about the calculation done in this process which is eventually utilized to generate the Mean Absolute Error. There were 10 Experiments done in total, with each experiment containing 10 trials. The resultant values have been plotted onto a graph depicting the process given infigure 2 above. The Mean absolute error estimated for this technique is 0.03, which is exceptionally low. Which is a good sign for the proposed technique.

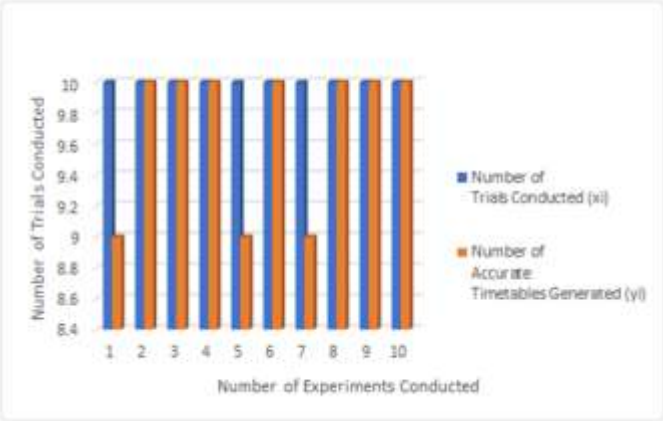

Figure 2: Evaluation of MAE.

V CONCLUSION AND FUTURESCOPE

The person who is in charge of timetable creation through a manual work in educational organization knows the complexity of this task. As mentioned earlier, there are very few costlier softwares are available in the market to generate the timetable automatically. This proposed model generates the time table for the computer department of a Engineering college for the semesters 4th,6th and 8th with two divisions namely A and B.

Here proposed model uses the Shannon information theory to estimate the importance of a staff based on his distribution of the classes in all the divisions. Then by using the permutation and parallel data allocation through decision making efficient timetable is generated for all the divisions of all the semesters without any overlapping issues. The performance of the model is evaluated using the Mean absolute Error which yields a better result of around 0.03, which is eventually represented the error rate, which is in turn is very low and indicates better performance of the system.

In the future this system can be enhanced to work on user defined time slots with more possibilities of timetable generation.

REFERENCES

[1] Yawar Rehman, Ghous Bakhsh Narejo, and Shoaib Zaidi, “An Experiment for TestingEfficiency of EffectiveTeaching Model and Comprehensive Use ofLimited Resources”, IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE) 2012.

[2] R. Badoni and D. Gupta, “A New Algorithm Based On Students Groupings For University Course Timetabling Problem”, RAECS UIET Panjab University Chandigarh, December 2015.

[3] J. Zhong, M. Shen, J. Zhang, H. Shu-Hung, Y. Shi and Y. Li, “A Differential Evolution Algorithm With DualPopulations for Solving Periodic RailwayTimetable Scheduling Problem”, IEEE Transactions On Evolutionary Computation, Vol. 17, No. 4, August 2013.

[4] Samir Ribi´c, RazijaTurˇcinhoˇzi´c, AmelaMuratovi´c-Ribi´c, “Modelling Constraints in School Timetablingusing Integer Linear Programming”, International Conference on Information, Communication and Automation (ICAT), 2015.

[5] S. Limanto, N. Benarkah, and T. Adelia, “Thesis Examination Timetabling using Genetic Algorithm”, International Electronics Symposium on Knowledge Creation and Intelligent Computing, 2016.

[6] Xiao Fei Yang, MasriAyob, MohdZakree Ahmad Nazri, “An Investigation of Timetable Satisfaction Factors for A Practical University Course Timetabling Problem”, International Conference on Electrical Engineering and Informatics (ICEEI), 2017.

[7] Maria Weslane Sousa Almeida, Jo˜ao Paulo Souza Medeiros, and Patrıcia Rufino Oliveira, "Solving the Academic Timetable Problem Thinkingon Student Needs", IEEE 14th International Conference on Machine Learning and Applications, 2015.

[8] Suryanti Awang, Mohamad Fairol, Marzuki Khalid andRubiyahYusof, "Generating University Timetable by using a Combination of Heuristic and Random Selection", IEEE International Conference on Computer Science and Information Technology, 2009.

[9] Yang Ting-Hong, Dan Qi1, Shen Xiao-Na, Zhao Jingand Xie Li, "The Urgency Of Timetable Task In Course Timetable Planning", The 2010 International Conference on Apperceiving Computing and Intelligence Analysis Proceeding, 2010.

[10] F. Guo and H. Song, "Research and Application of Data-Mining Technique in Timetable Scheduling", 2nd International Conference on Computer Engineering and Technology, 2010.

[11] P. Kaewchanid, P. Wangmaeteekul, "Solving a Teaching Assistant Timetabling using Constraint-Based Approach", International Conference on Computational Intelligence and Applications, 2016.

*****