Article 5

Matching Final Year Project Topics with Students using Stable Marriage Model

Naimah Mohd Hussin (PhD), Ammar Azlan,
Faculty of Computer & Mathematical Sciences,
Universiti Teknologi MARA Perlis Branch, Malaysia

Abstract
Every semester, a new batch of final year students needs to find a topic and a supervisor to complete their final year project requirement. The problem with the current approach is that it is based on first come first serve. So, the pairing between student and supervisor is not the optimal ones, i.e. some students may not get their preferred topic or supervisor. Plus, it is also time consuming for both students and supervisors. The researcher is motivated to solve this long overdue problem by applying a stable marriage model that is introduced by Gale and Shapley hence the name Gale-Shapley Algorithm. To determine the functionality of this approach, a system prototype has been constructed and a random dataset is used. The result, 60% of the students get their first choice topics while the remaining students get their second or third choice. This is a remarkable outcome considering the time and effort saved compared to the current process. Therefore, stable marriage model is applicable in solving student-topic pairing.

Keywords *Gale-Shapley Algorithm, Stable Marriage Model, matching, optimization, pairing, final year project*

Introduction

Final year students are required to undergo a course, Information System Project Formulation, where they need to do a proposal of a final year project in any area related to their program. This project will be supervised by a lecturer from their department. Some lecturers may supervise more than one student, depending on their schedule and expertise. The lecturers involved in supervising final year projects and the total students they can supervise are determined at the departmental level. The project title however will be provided by the lecturers or proposed by the students.

Therefore, the first step in this final year project development is for the students to find a supervisor with a project title that suits the students' interest. As simple as it may sound, there are many students who had made the wrong choice in finding either project title or supervisor that is suitable for them. This mistake is propagated throughout the semester. As a result, they might not perform well in their final presentation. Some are lucky to get away with a low grade. Unfortunately, for those who are not so lucky, they got a failed grade and had to repeat the course all over again.

This matching problem might occur from the students' lack of knowledge about the project title being proposed by their supervisor. On the other hand, the supervisors do not know the students well enough to determine whether the title is suitable for them or not. Therefore, it would be practical to have a platform which acts as the 'middle-man' to find the best match between students and project titles.

The algorithm that is used in performing this matching is Gale-Shapley's algorithm. This algorithm has been implemented in many areas such as to assign medical students to hospitals based on the preferences of students over hospitals and vice versa. It is implemented in many countries with different names. To name a few, it is known as NRMP in the US (Gusfeld and Irving, 1989), CaRMS in Canada and SPA in Scotland (Irving, 1998). Another application is reported by Teo et al (2001), which assigns students to secondary schools in Singapore. Since all of the problems are solved by using this algorithm, then it is a good hypothesis that the problem of matching students and project titles will be solved too.

Finding a suitable supervisor and project title is essential for final year students in order to perform their best in their final year project. It would also be best if this matching could be done automatically with high resemblance to manual selection.

Unfortunately, until today, the matching between students and supervisors is still done manually. It is done where students would meet and ask supervisors about the project title. Students can choose the supervisor or the topic that they find interesting. The flaw in this system is that, the choosing is only done by students. Supervisors basically just accept the students based on first come first serve basis.

As a result, the quality of students' final year project will be fixed at the same level if no improvement is being done on this matching problem. The current matching approach is also time-consuming for both students and supervisors. These problems are believed can be solved by implementing Gale-Shapley algorithm.

This research studies the Gale-Shapley algorithm and determines if the algorithm can be implemented in matching problem with respect to project titles and students' interest in their final year project. This research focuses on the matching process and the quality of the solution produced.

Literature Review

Final Year Project in UiTM Perlis

Final year project is a project which every final year students have to carry out for two semesters. Currently no application has been implemented to manage the project. This includes from the beginning phase of the final year project, which is finding the suitable supervisor, to the last phase which is preserving the finished project. This project will concentrate on the first phase which is finding the best match between final year students and the project titles.

The perfect match between students and project titles is essential in order to optimize the performance of both parties. Students will produce better project if the project they are working on is suited with their interest and ability. On the other hand, supervisors would give a better guide to the students if the area of project is within their expertise.

Stable Marriage Model

Stable matching problems were first studied by Gale and Shapley (1962). In a stable marriage problem there are two finite sets of participants: the set of men (M) and the set of women (W). In the Gale-Shapley model, the preference lists of the participants are required to be completed, and no one is to be declared as unacceptable. A matching is just a one-to-one mapping between the two sexes such that a man *m* is mapped to a woman *w* if and only if *w* is mapped to *m*, and *m* and *w* are acceptable to each other" (Teo et al, 2001).

The stable marriage model consists of two sides of conditions to be fulfilled which are the male side and the female side. The matching process will go through a number of recurrences until the most optimal match is found. There are also conditional and nested conditional statements exist within the recurrences. The first conditional statement matches a man with a free woman. The second one matches a man with an engaged woman if she prefers the man over his fiancé. This decision is done with a nested conditional statement within the second conditional statement. This whole process is shown in Figure 1.

The structure of this model has proven that there is close similarity between stable marriage problem and final year project problem. With a few modifications on the algorithm, it is possible to implement stable marriage model in finding the best match between students and topics in the final year project. In final year project case, the male and female part is represented by students and topics respectively.

```
functionstableMatching
{
   Initialize all m∈M and w∈W to free
while∃free man m who still has a woman w to propose to
   {
      w = m’s highest ranked such woman
if w is free
         (m, w) become engaged
else some pair (m’, w) already exists
if w prefers m to m’
            (m, w) become engaged
            m’ becomes free
else
            (m’, w) remain engaged
   }
}
```

Figure 1: Gale-Shapley Stable Marriage algorithm

Application of Stable Marriage Model

Gale-Shapley Stable Marriage Problem algorithm is a well-known approach in solving matching problem. Many researchers have done studies where the algorithm is applied to other areas. Dean et al (2006) said that before becoming a doctor, a medical school graduate in the United States is required to complete a residency program at a hospital. The medical field has turned to a centralized marketplace, called the National Residency Matching Program (NRMP) since 1950s to help this marketplace. The program basically works by final-year medical students and hospitals each submit preferences over possible matches. Then, an algorithm determines which match is the most desirable. The system is considered as successful if the generated matches are stable. Computing a stable matching is a classic problem in economics and computer science. It can be solved in polynomial time by the deferred acceptance algorithm of Gale and Shapley.

Furthermore, there is also a case where the stable marriage problem is put into practice in hospital-residents problem. According to Unsworth and Prosser (2005), both the stable marriage problem and hospital-residents problem are stable marriage problems. They consist of two sets of objects that need to be matched to each other. In stable marriage problem, men need to be paired to women meanwhile in hospital-residents problem, residents need to be assigned to hospitals. Each set of objects expresses a weighted preference in the form of a preference list for the objects in the opposite set. The problem is then to find a matching of one set to the other set which is stable.

This shows that stable marriage problem is a very flexible algorithm and has been used in solving problems in many areas. All the researches and implementation of stable marriage problem in other fields have proven to come up with satisfying results. It is believed that it can also be implemented in final year project and more importantly, it will produce good results. This could be further induced by a statement in an article entitled Geometric Stable Roommates as follows: "*The Stable Marriage problem is a true multidisciplinary problem: it is well-studied in economics, computer science, and combinatory. The problem and its numerous extensions continue to receive considerable attention, both from the theoretical point of view and for real-world applications*" (Arkin et al, 2009).

Construction

The system is implemented as a web-based application since it is necessary for the system to be accessible by all final year project students and supervisors. The process flow for the system is as shown in Figure 3.1 and is self-explanatory.

1. Login (as administrator, student or supervisor)
2. Registration (student and supervisor)
3. If (supervisor)
    a. Register topic(s)
    b. For each topic, rank students who have chosen the topic
    c. If (algorithm executed)
        i. Retrieve matching results
4. If (student)
    a. Rank topic(s) Note: 1 as the most preferable topic
    b. If (algorithm executed)
        i. Retrieve matching results
5. If (administrator)
    a. Register all supervisors and students
    b. If (supervisors and students have completed ranking)
        i. Run algorithm to find matching
        ii. Display matching results
6. End

Figure 2: Process Flow

Results and Findings

Table 1 shows the dataset used in the testing. Students are represented by their last two digits of their student number. Each column indicates how each student rank the topics. For example, the entry in the first column and the first row (shown in bold) indicates student whose student number ends with 26 ranks topic with ID number 14 as the first rank.

Table 1: Dataset – Students' ranking

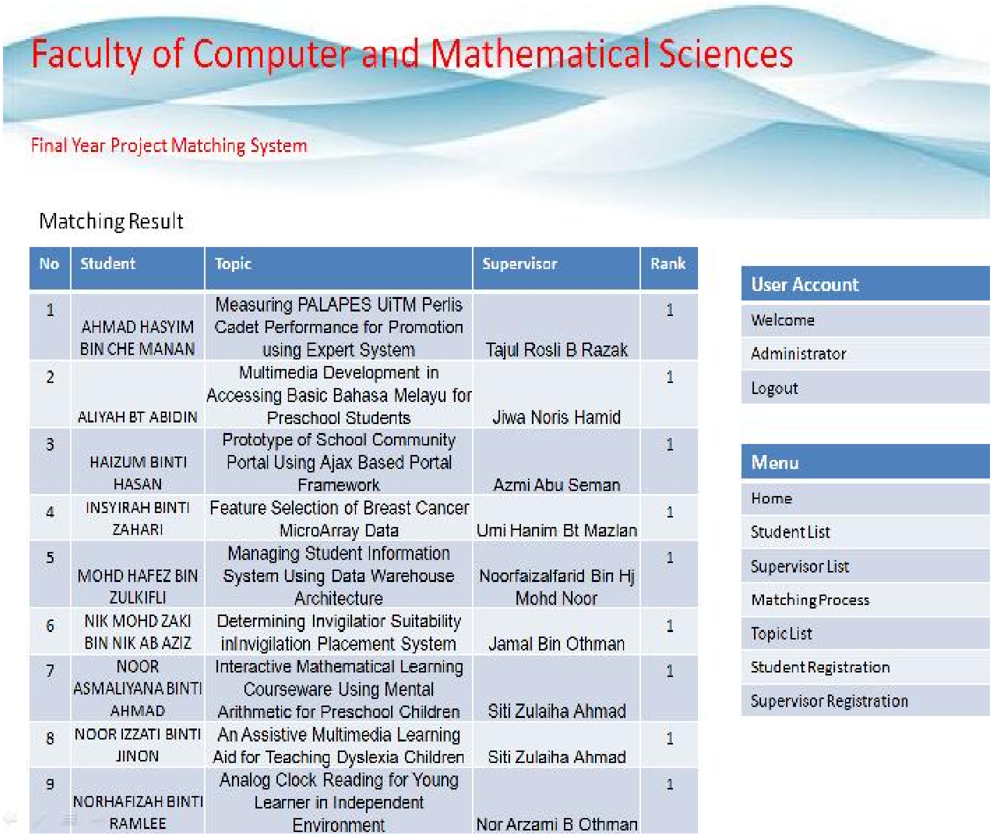
Rank	Student 26	32	36	38	64	76	84	86	94	68
1	**14**	10	3	14	3	11	4	3	5	3
2	2	13	9	9	9	6	6	6	9	2
3	12	12	10	12	4	9	7	8	12	13
4	7	6	12	6	11	4	10	11	13	10
5	15	7	6	13	13	10	13	14	6	7
6	10	9	13	8	6	2	8	1	7	5
7	6	8	5	2	15	1	15	5	10	15
8	9	14	15	4	5	7	2	9	14	11
9	11	15	4	3	8	12	3	12	15	12
10	13	1	8	10	12	13	9	13	2	9

The dataset for supervisors' preferences is also done in similar way. For example, in column 1, supervisor for topic 1 has ranked student with student number ends with 86 as the first rank. Referring to Table 1, it is clear that not all topics have the same number of students to be ranked. The more number of students for that topic shows that the topic is preferable among students. In this case, topic 9 and 13 are the most preferable topics because all students rank them. Topic 1 is the least preferable topic because only three students rank it.

Table 2: Dataset – Supervisors' ranking

Rank	Topic 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	86	38	38	38	64	76	68	38	32	84	76	26	84	38	64
2	76	94	36	84	86	94	32	84	84	32	86	86	32	86	68
3	32	84	64	36	68	86	84	36	86	38	64	76	26	32	32
4		26	86	64	94	64	26	86	64	76	26	36	86	26	84
5		76	84	76	36	26	76	64	26	94	68	32	94	94	26
6		68	68			32	94	32	94	68		38	38		36
7						38			36	36		64	36		94
8						84			38	26		94	76		32
9						36			68			68	68		
10									76				64		

The matching result of these two preferences is shown by listing all the topics and the matching students. Besides generating the matched list between students and topics, it also shows the rank of the topics that students get (See Figure 3)



No	Student	Topic	Supervisor	Rank
1	AHMAD HASYIM BIN CHE MANAN	Measuring PALAPES UiTM Perlis Cadet Performance for Promotion using Expert System	Tajul Rosli B Razak	1
2	ALIYAH BT ABIDIN	Multimedia Development in Accessing Basic Bahasa Melayu for Preschool Students	Jiwa Noris Hamid	1
3	HAIZUM BINTI HASAN	Prototype of School Community Portal Using Ajax Based Portal Framework	Azmi Abu Seman	1
4	INSYIRAH BINTI ZAHARI	Feature Selection of Breast Cancer MicroArray Data	Umi Hanim Bt Mazlan	1
5	MOHD HAFEZ BIN ZULKIFLI	Managing Student Information System Using Data Warehouse Architecture	Noorfaizalfarid Bin Hj Mohd Noor	1
6	NIK MOHD ZAKI BIN NIK AB AZIZ	Determining Invigilatior Suitability inInvigilation Placement System	Jamal Bin Othman	1
7	NOOR ASMALIYANA BINTI AHMAD	Interactive Mathematical Learning Courseware Using Mental Arithmetic for Preschool Children	Siti Zulaiha Ahmad	1
8	NOOR IZZATI BINTI JINON	An Assistive Multimedia Learning Aid for Teaching Dyslexia Children	Siti Zulaiha Ahmad	1
9	NORHAFIZAH BINTI RAMLEE	Analog Clock Reading for Young Learner in Independent Environment	Nor Arzami B Othman	1

Figure 3: Matching result

Figure 4shows the pie chart that compares the number of students with the ranks of topic they get. The darkest section in the pie chart represents the number of students who get their first ranked topics. It can be concluded that the system satisfies more than half the students (60%).

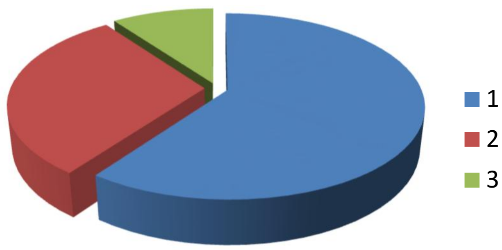


Figure 4: Percentage of students for each rank

Conclusion and Recommendation

In this research, it is proven that stable marriage model can solve the problem in pairing students with final year topics. The system is able to find a match that satisfies 60% of first choice preference.

The actual system is recommended to be used by the Computer Science Department to manage their final year projects. However, the system needs to be improved in terms of its interface and user-friendliness. Secondly, user of the system will be more pleased if the system could be extended to enable the user to either accept or reject the title generated by the system. It will extend the timeline of the system but as with the current course registration used by UiTM, there is add and drop period where users can still alter the courses they have preregistered.

References

Dean, B.C., Goemans, M.X., Immorlica, N., (2006). *"The Unsplittable Stable Marriage Problem"*. In International Federation for Information Processing, Volume 209, Fourth IFIP International Conference on Theoretical Computer Science-TCS 2006, eds. Navarro, G., Bertossi, L., Kohayakwa, Y., (Boston: Springer), pp. 65–75.

Arkin, E. M., Bae, S.W., Efrat, A., Okamoto, K., Mitchell, J. S. B., &Polishchuk, V. (2009).*Geometric Stable Roomates*. In Information Processing Letters, Volume 109 Issue 4, January, 2009. pp 219-224.

Gale, D., & Shapley, L. S., (1962).*College admissions and the stability of marriage*. American Mathematical Monthly, Vol. 69, p. 9–15.

Gusfeld, D., & Irving, R. W. (1989).*The stable marriage problem: structure and algorithms*. MIT Press, Cambridge, MA, USA.

Irving, R. W. (1998). *Matching medical students to pairs of hospitals: a new variation on a well-known theme*. In Proceedings of ESA'98: the Sixth Annual European Symposium on Algorithms, Vol. 1461, p. 381-392.

Teo, C. P., Sethuraman, J., & Tan, W. P. (2001). *Gale-Shapley Stable Marriage Problem Revisited: Strategic Issues and Applications*. In Management Science © 2001 INFORMS Vol. 47, No. 9, September 2001 pp. 1252–1267

Unsworth, C., Prosser, P. (2005) *An n-ary Constraint for the Stable Marriage Problem*, The Fifth Workshop on Modelling and Solving Problems with Constraints, held at the 19th International Joint Conference on Artificial Intelligence (IJCAI 2005)