Stable Marriage Problem (Hospital Intern Matching)

Jaskaran Singh Bhatti
Department of Computer Science
University of Central Florida
Orlando, United States
Jaskaran.bhatti@knights.ucf.edu

Basundhara Dey
Department of Computer Science
University of Central Florida
Orlando, United States
dey.basundhara@knights.ucf.edu

***Abstract*— In our paper and project, we look at the current headway and more real life applications of the Gale Shapley stable marriage problem. In precise, we concentrate on the case of Hospital Intern matching problem. We study the correlation between the stable marriage problem and matching and try to infer graph theoretic principles that inspire the approaches and methodologies used to solve these real life scenarios. We implement Gale Shapley based greedy method for optimal stable marriage solution in the initial preference list. We also generate the results and special cases after the simulation and study. We also document a few difficulties and hurdles encountered throughout the project.**

***Index Terms*— Gale Shapley Algorithm, Nobel on economics, Optimality Issue, Graph Theoretic Property, NRMP,**

I. Introduction

There are several principle to assign elements of disjoint sets to one another. But mostly problem arises, when each element is associated with different cost and we need to maximize or minimize the total cost of all the assignments. To solve this issue, we find Stable marriage problem is the most suitable one to apply.

In the stable marriage problem, we wish to pair off, or in short marry, n men and n women, to construct a set of n matching pairs in such a way that everyone is satisfied with his/her partner. When setting up the pairs, we need to happen such a situation where no man and no woman prefer each other to their current partner. If this comes true, then we have successfully established n stable marriages. But if not, then we have a 'blocking pair', where 2 people in pair, but both prefer each other to their current matching. To start this pairing, each man and woman have too rank the members of the opposite sex in order of their the same, ordering the women according to their preference. If we represent graphically, we can say that two disjoint sets A and B have the same number of elements in each. We will find a set S of n pairs (a, b) such that a is an element of A and b is an element of B.

This stable marriage problem has a wide variety of practical, precisely real life applications, which range from matching resident doctors to hospitals [1], to matching students with colleges, or more generally between any two-sided market. In our paper, we are mostly concentrating on hospital intern matching problem. That is how hospitals and interns have been assigned to each other, on depending on their preference list of both. For concerning the special case, we also go through the intern- problem of married couple. There may be the cases where interns can also be married and in such a case every couple has preferences over pairs of joint assignments. [2] Such assignments must satisfy certain stability requirements. The problem can be defined such that when stating their preferences, these can give same (tied) preference to agents with which they can be matched.

II. Previous Works on Stable Marriage

In 1962, David Gale and Lloyd Shapley [3] had introduced the stable marriage algorithm. They were working on college admission problem of students and there they implement the SMP algorithm initially. The word "marriage" used to differentiate two sets- one as 'male', another as 'female'.

They had two different groups, representing men (M) and women (W). Each containing n entities and a preference list of size n, where the initial entry represents the most desirable choice of opposite sex and so on. The

the same preference level and where any particular women w∈W doesn't feature in the preference list of any man, m∈M. They represented a diagram (Fig.1) that depicts the two sets of preferences for men and women. This two sets are called "Preference Matrix" of Men and Women.

1:	3 1 5 7 4 2 8 6	1:	4 3 8 1 2 5 7 6
2:	6 1 3 4 8 7 5 2	2:	3 7 5 8 6 4 1 2
3:	7 4 3 6 5 1 2 8	3:	7 5 8 3 6 2 1 4
4:	5 3 8 2 6 1 4 7	4:	6 4 2 7 3 1 5 8
5:	4 1 2 8 7 3 6 5	5:	8 7 1 5 6 4 3 2
6:	6 2 5 7 8 4 3 1	6:	5 4 7 6 2 8 3 1
7:	7 8 1 6 2 3 4 5	7:	1 4 5 6 2 8 3 7
8:	2 6 7 1 8 3 4 5	8:	2 5 4 3 7 8 1 6
male preference lists		**female preference lists**	

Fig.1: Preference Matrix

Now, mapping is simply connection of two elements, resides in two different sets. Marriage S is stable only if for any (m, w) in S, both m and w, prefer each other than to any other women w' and m' respectively, where m, m^'∈M and w, w^'∈W. [3] The objective of the Gale and Shapley paper was to come with a marriage set where everyone is married and the marriage agrees with the stability criterion. Gale and Shapley proved that there exists at least one or more stable marriage stable marriage set. Thus they converted the college admissions problem in to a marriage assignment problem and showed that such an O (n^2) solution always exists. It may be noted that there may be many such possible combinations of marriages which are stable.

[4] On 1971, Mcvitie, and Wilson et. al extended the original Gale and Shapley algorithm, with devising an algorithm which deduces all the possible sets of stable marriages in a recursive manner and also throws light on the nature of optimal solutions. The recursive algorithm ensures each stable solution occur exactly once. If males are allowed to choose proactively and women are reactive to proposals, then it gives us a stable solution which is male optimal as it gives a marriage for no woman can have a better partner than he does in this matching and no woman can have a worse one. On the other hand, if females are allowed to choose and men are reactively accepting or discarding in suspending the as yet best partner, then no women can have a better partner that she does in this matching and indeed in both cases satisfaction levels are skewed towards either male or female. This gave motivation to the next work which was on Efficiency of stable marriage.

Later on 1987, Rob Irving, P. Leather and D. Gusfield [5], discusses the concept where the best marriage that partner, then no women can have a better partner that she will guarantee maximum satisfaction of both groups together and provide a stable solution fair to both the groups. In this paper we have a scenario where optimality is defined as finding the marriage that gives maximum satisfaction (to both M and W) out of all the possible sets of stable marriage. Satisfaction is measured by position of his/her chosen marriage partner in his/her preference list. Let the position be denoted by j. Minimum values of the position will ensure maximum satisfaction. The value of objective function is nothing but the sum of j's in the potential candidate stable marriage. The optimal marriage is found by minimizing the value of an evaluation objective function. Generating all stable marriage and comparing their values will exponential in growth in problem size so it is not a feasible approach. So devise a polynomial time solution for the problem by using concept of exposed rotations and elimination of rotations which follow certain properties. It's a multi-stage iterative process which yields a final rotation. We have two sets Men and Women, where any man m∈M and w ∈W and the number of men and women in each set are equal to n. The mathematical formulation of the problem is as follows:

$$mr(i,k) = j, \quad if\ woman\ k\ is\ the\ j\ th\ choice\ of\ man\ i.$$

$$wr(i,k) = j, \quad if\ man\ k\ is\ the\ j\ th\ choice\ of\ woman\ i.$$

Where, mr: is the preference matrix of men and wr: is the preference matrix of women.

$$S = \{(m_1, w_1), (m_2, w_2), \ldots\ldots\ldots, (m_n, w_n)\}$$

The optimal marriage is the combination of pairs that minimizes total or average satisfaction of 2n individuals, and is optimal with minimum possible value of c(S).

Then the problem raised in stable marriage where there is a tie in the preference or there is incomplete preference. For e.g. the incompleteness means, say, Man m_i feels that a certain women w_i is just not his type, then w_i will not feature in his preference list. Also, for m_i two or more women might have a same preference level. This when combined is known as SMTI (Stable Marriage with Ties and Incomplete Lists- 2002).[6] The first empirical study and experimental set up of an SMTI problem and complete algorithm using a constraint programming encoding with a time complexity of the order of O (n^4). (I.P. Gent & P. Prosser et. al. 2002). The SMTI represents a more real world situation than its earlier counterparts. Although the classical SM problem has a polynomial time solution, solving an SMTI problem to find the maximal cardinality stable marriage is an NP-hard problem. The SMTI problem has solutions where world situation than its earlier counterparts. Although the stable marriages may have singles as they have an incomplete preference list. The marriage which is considered optimal is the marriage set that has maximum marriages and minimum singles (called maximum cardinality marriage). The problem was introduced in [7], and subsequently there has been a body of research dealing with this subclass of problems. We study a particular work mentioned in [6], which exploits a heuristic local search based greedy approach to provide the maximum cardinality marriage for SMTI problem, and claims that the number of steps grows only as (nlogn). We explore this application extensively in the later section.

In 2012,[8] Lloyd S. Shapley and Alvin E. Roth were Awarded with Noble Prize in Economics "for the theory of stable allocations and the practice of market design. The two researchers worked independently of one another, combined with empirical investigations, experiments from Shapley's basic theory and Roth's empirical investigations which resulted in boom in the field of research and performance in many other areas. The economic Engineering was one which was awarded with prize the current year.

III. STRUCTURE OF STABLE MATCHING

To understand stable marriage problem, we consider the basic Gale Shapley algorithm. Below the algorithm is:

```
Initialize all b ∈ B and g ∈ G to free ;
while (some boy b is free & has a girl g to propose)
{
  g := first ranked girl on b's list;
          if    g is free
              (b, g) become engaged;
      else
              some pair (b', g) already exists
      if    g prefers b to b'
              (b, g) become engaged;
               b' becomes free;
      else
              (b', g) remain engaged;
}
```

In that algorithm, b is man and g is woman. B and G are the two different sets of Men and Women. Initially both of them are free. Now we introduce a loop to count all the man and woman in the group one by one. Let us presume that proposal initiate from the men side. So a b will propose the first prioritized g on his list. Now algorithm will first check if g is free. If she is, (b, g) will engage. Else it will be certain that some (b', g) already exists. Now again a check will happen. G will check her preference list and if she finds b's rank greater than b', then she will break the commitment and will be engage with b. b' will be free in that case. What if b' is higher prioritized than b? Then (b', g) will continue. This way the loop will gone till all the woman and man are making pair of (b, g).

For better understanding, we also have constructed pseudo code, better say formalization.

- Set n boy $S_B = \{B_1, B_2, ..., B_n\}$
- Set n women $S_G = \{G_1, G_2, ..., G_n\}$
- Each man ranks the women in S_G in strict order of preference.
- Each woman ranks the men in S_B in strict order of preference.
- A matching S_M is a bijection between the men and women.
- A (boy, girl) pair (B,G) **blocks** S_M if:
  - **b** prefers **g** to his partner in S_M, and
  - g prefers **b** to her partner in S_M.
- In a S_M which has no blocking pair is **Stable.**

For now, let us take a good example to demonstrate the problem.

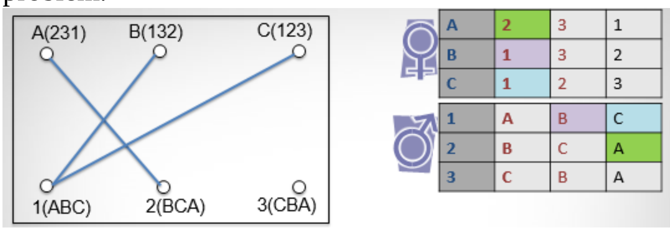

Fig: 2

As stable matching is related to graph theory. So here we also construct the corresponding graph as well, while demonstrating problem.

From fig 2, we can see, {A, B, C} is a set of women and {1, 2, 3} is a set of men. From the graph we can see, it's a $K_{n,n}$

Step 1: A proposes to 2, B proposes to 1 and C proposes to 1. From men's table, 1 get two proposes from B and C. 2 get from A. 3 got no one. In the graph itself, we connect the proposals as edges.

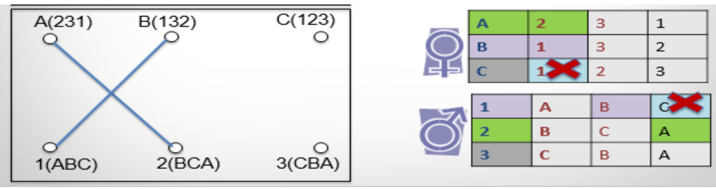

Fig: 3

Step 2: There is a choice option to 1. As B is more preferable than C, so 1 get engaged to B and 2 get engaged to A.

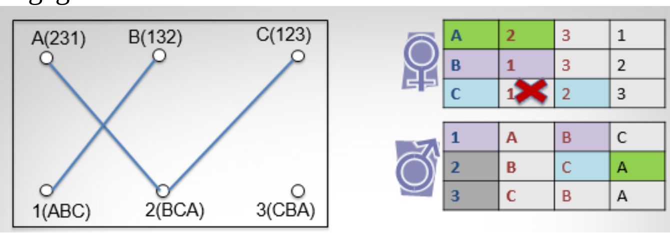

Fig: 4

Step 3: Now, in the next turn, C propose to his second choice 2, as she is already rejected by 1. In the boys table, 1 get proposal to B, with whom he already engaged. By 2 get two options to choose. A and C.

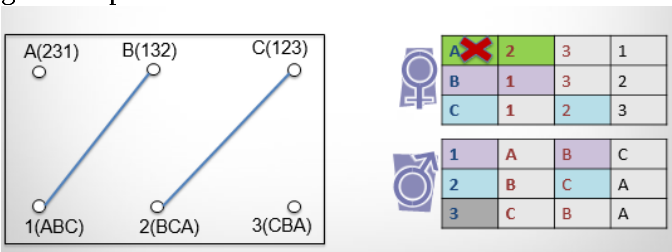

Fig: 5

Step 4: As C is on 2nd choice and A on 3rd, 2 rejects A and break his engagement. He newly get engaged to C. End of this stage, again 2 couples, B-1 and C-2.

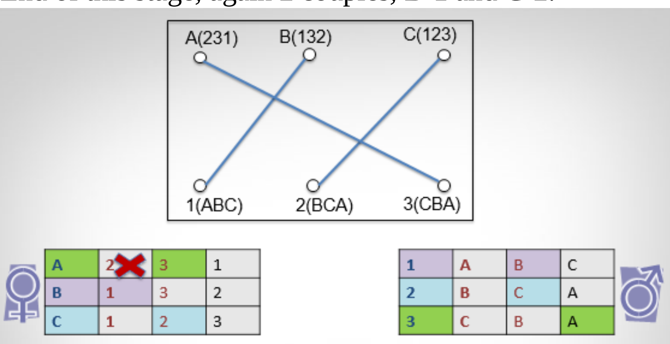

Fig: 6

Step 5: Now single A proposes to her second choice 3. As 3 is still single and has no other proposal, so it gets engaged to A.

So end of this, we get 3 tentative couple, A-3, B-1 and C-2. The system is stable as in each pair, both are preferred by each other.

IV. Basic Observations

While Gale- Shapley has introduced the algorithm and used it for solving the college admission problem with students, they have concluded some of the observations noted below.

Definition. An assignment of applicants to college will be called unstable if there are two applicants a and b who are assigned to colleges A and B, respectively, although b prefers A to B and A prefers b to a. [3]

From the above definition, we get a perfect idea of what an unstable condition could be. But still question remains that, what will be optimal solution for this matching? Galey also have discussed the same.

Definition: A stable assignment is called optimal, if every applicant is at least as well of under any other stable assignment. [3]

One thing is clear, whatever the optimal solution will be, and that will be unique. From the algorithm Gale also introduced some theorem, which helps to clarify stable marriage more clearly.

Theorem1: There always exists a stable set of marriages. [3]

Proof. We can prove it through iterative procedure to find a actual stable set of marriages. If we go through the algorithm (let's assume men propose female) ,in the second stage, whoever got rejection again will propose to his next best choice , and eventually women again reject all but her best proposal till so far. So in this way, at most n^2-2n+2. Because as long as any woman stays single, the iteration will continue. But gents can't propose to the same girl more than once. So every woman will surely get a proposal in due time. As soon as the last female get engaged, the state is called Stable. In that state, every girl is engaged with his favorite (irrespective of the no. of preference) gent.

Theorem2: This algorithm always terminates.

[9]Remember that a woman will accept her first proposal, and will always remain engaged from that point on although she might "trade up" for a better husband. Also, every man will propose to a woman, or will currently be engaged to a woman in each round. Since we have the same number of men and women, we know that we cannot have a single male without having a single female. If these two single individuals did exist, then the man would have to propose to the woman at some point as she is listed on his preference list. Thus, they would become engaged. This ensures the algorithm will end.

Now a crucial question is: male optimal or female optimal? Considering the example we can see, whoever did the proposing ended up with their higher choice for a spouse. [9] A person's optimal mate is that person's favorite from the list of possibilities. Same wise, a person's pessimal mate is that person's least favorite from the list of their possibilities.

In Irving's paper, considering this all fact, he concludes some properties of stable marriage.

Property 1: If every man is paired with the first woman on his shortlist, then the resulting matching is stable; it is called the male optimal solution, for no man can have a better partner than he does in this matching, and indeed no woman can have a worse one.[5]

Property 2: If the roles of males and females are interchanged, and if every woman is paired with the first man on her (female-oriented) shortlist, then the resulting matching is stable; it is called the female optimal solution, for no woman can have a better partner than she does in this matching, and indeed no man can have a worse one. [5]

We find this a shortcoming of this algorithm. Though it calls stable, but any set's element have to compromise with its choice in matter of ranking.

He also conclude that (a) if a certain woman doesn't appear in any certain man's preference list, then there is no possibility of matching them together. (b) Any certain woman will be first on any man's sort list if and only if man is last on woman's.

V. Graph Theoretic Application

A. Independent edge set

We can draw an analogy of the independent edge set problem with the stable marriage problem. For example each element in the independent edge set represents a marriage, then since one man or woman can be married to only one partner, so the nodes are represented by men and women, and independent edge set will ensure that only one node is associated with an element in independent edge set . The following diagram shows such a situation where weights are preferences.

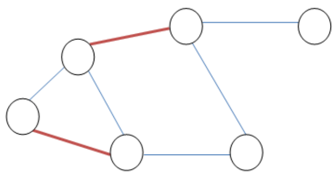

Fig: 7

We can increase the no. of marriages; if we alter the marriage pairings in the following way in which we have independent edge set which preserves the monogamous nature of the marriages and at the same time we increase the number of marriage and get the maximum cardinality marriage.

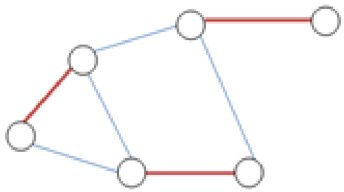

Fig: 8

the above example is an example of perfect matching. Matching number of a graph is the size of the maximum matching. In case of perfect matching, matching number is |V|/2. The maximum -bipartite matching is an analogous problem to find the maximum matching.

B. Bipartite Matching

Now if we recall we discussed about the concept of singles associated with SMTI problem model in the previous section. The concept of singles remaining unmatched in a stable marriage can be realized through the following bipartite Graph example. If we notice the following figure with Men on left and women on right we see that 1,2,3 are married while 4 and 5 remain single because Men 4 have only women 8 in its preference list but women 8 had a higher preference for Man 3 and hence 3 – 8 are matched leaving Man 4 unmatched. The concept of singles in a stable marriage occurs due to the incomplete preference list property, where in this case Man 4 has only women 8 in its preference list. Similar is the case with Man 5 however the next step Man 5 can marry Women 10, or women 9 but for that to happen Women 9, 10 also have to have Man 5 in their preference list. The diagram below is shown with respect to a Male preference list Bipartite Graph, and bold lines are marriages.

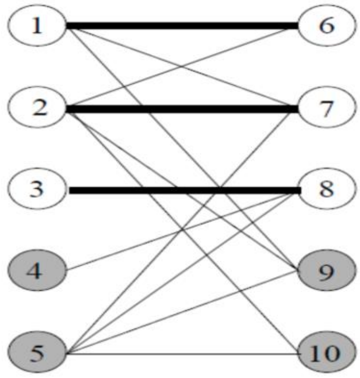

Fig: 9

Now let us analyse how the stable marriage problem is related to Bipartite graphs and matching. The nodes can be divided into two groups of Men group and Women group and an edge represents a possible pairing and there is no edge between the nodes belonging to the same group. Hence it satisfies the bipartite graph property. This being the case we have preferences denoted by edges colored in black (see Figure 10 (a), 10 (b)) and out of the preferences chosen pairs for marriage might be shown in some other color (red for e.g. in Fig 5(a)). An incomplete bipartite graph with black edges will denote a preference list that is incomplete. We observe that the Fig. 10 is an incomplete bipartite graph. Now we have discussed that optimality in stable marriage often deals with increasing cardinality of the stable marriage by trying to find a stable marriage which marries more people.

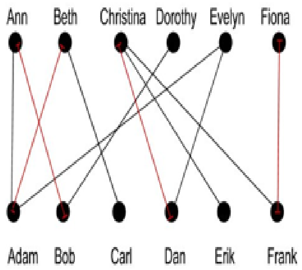

Fig: 10(a)

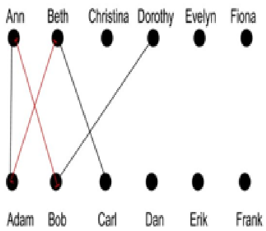

Fig: 10(b)

C. Maximal Flow Min cut

There is a way to improve on a matching by finding a path from an unmatched woman to an unmatched man (dominating) in which every second edge is in the current matching (dominated). Such a path is called an alternating path, relative to the matching M, or a-path for short, because it alternates between edges not in the current matching and edges in the current matching.

If that is the case, we still may be interested to nonetheless extend the benefits of marriage to the largest number of people in the group. So we may want to find a maximum matching that is one with maximum cardinality. This latter generalization of the Marriage Problem is called the Maximum Matching Problem. It has the advantage that it could be applied even if the number of women and men in the group were not the same.

The question is can we always improve on a matching if we find an alternating path? The answer is yes, because an alternating path is a path from a woman to a man in a bipartite graph. Consequently it has odd length. Therefore we have always one more odd-numbered edge on an alternating path than the number of even-numbered edges. Since all odd-numbered edges are not in the matching, while all even-numbered ones are, we increase the number of matching by one if we replace the even-numbered edges by the odd-numbered ones in the matching. The result is a new matching with one more edge than the one we started with.

This concept of finding the alternating path from the existing marriage pair is the underlying graph theoretic motivation that has given rise to local search based greedy approach which we implemented (later in the paper), where there is a similar approach of choosing a random marriage and then try and generate candidate solutions which are possible alternatives that form a local neighborhood and search for the best alternative out of those (through some evaluation/objective function) and then move to that alternative and then again try to get another alternative from the same. This continues until we get a stable marriage with maximum size or some termination criteria is reached.

Hence the lemma,

Lemma: Suppose that G = (V, E) is a bipartite graph with vertex classes X and Y, and M is a matching in G. If there exists an alternating path from an unmatched vertex x in X to an unmatched vertex y in Y, then there exists a matching M' with cardinality |M| + 1. This forms the basis of motivation for later papers.

The ties can be modeled graph theoretically by the fact that there are weights which represent the preference order of a particular member with those whom are in the preference list. Say for Adam there is a tie in the preference for Ann and Evelyn in the first choice and Fiona is his second choice. See figure 11.

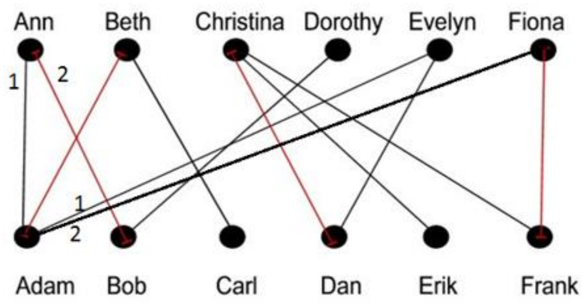

Fig: 11

Suppose there are two sets M and W, you can model such problems as flow problem, by adding a super source s1 with edges to all vertices in M, and a super sink d1, with edges from all vertices in W, and finding a maximal flow from s1 to d1.i.e a transformation of a maximum bipartite matching problem into a maximum flow problem. Figure 7 illustrates such a scenario.

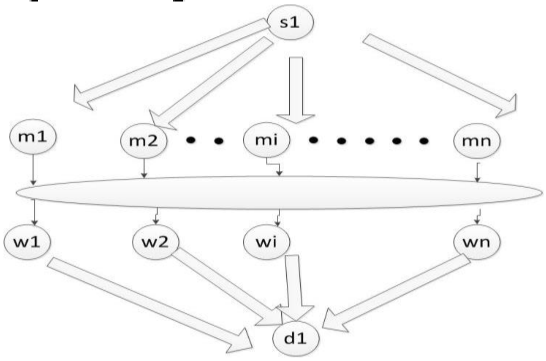

Fig: 12

VI. Hospital Intern matching

[10]In the 1940s, hospitals competed for residents by making offers early in medical school and demanding immediate responses. General instability encouraged private negotiations between students and hospitals. In 1952, after a few false starts, the National Intern Matching Program (NRMP) found a stable solution: the deferred-acceptance algorithm!

"The question of course then arises as to whether these results can be applied 'in practice'. [3] had expressed some reservations on this point,—and then came another surprise. Not only could the method be applied, it had been more than ten years earlier!" Prior to the mid-1990s the hospitals acted as the proposers. After a review by Roth et all the students propose.

[11]Formally stated, the intern assignment problem is a stable matching problem where a set of interns are assigned to a set of hospitals. Each hospital has a preset number of positions it wants to fill and a priority list that ranks the interns in accordance with its preferences over whom to hire to fill those positions. Note that since a hospital's preferences are one-dimensional, one should regard these multiple positions as identical. Different sections in the same hospital should be considered as different hospitals in this scheme. A hospital can rank in its preference list the possibility of leaving a position unassigned over being assigned to some of the interns. On the other hand, each of the interns has a preference list over the hospitals that specifies which hospitals the intern would like to be assigned.

The nature of the intern assignment problem changes profoundly if some of the interns are allowed to act as couples having a joint preference list over pairs of positions. This coupling phenomenon corresponds to the possibility that some of the interns are married to each other and want to be assigned to positions that are geographically close to each other. Existence of a solution for every instance of the problem can no longer be guaranteed. This is NP hard problem. We consider it as a special case.

VII. Simulation Results

We have done the coding part using C compiler.

Below we present the simulation result. We take different hospitals name and different intern and apply the stable marriage algorithm to solve it.

```
Cardiology Pairs with Michael
Endocrinology Pairs with Adams
Surgical_surgery Pairs with Cynthia
Pathology Pairs with Simons
Radiology Pairs with Gillchrist
Neurology Pairs with  Fred
Pediatric_surgery Pairs with Britney
 Unstable Pairing
Surgical_surgery removes Cynthia from pair
 Pairs with next member of their choice
Surgical_surgery Pairs with Taylor
 Unstable Pairing
Cardiology removes Michael from pair
 Pairs with next member of their choice
Cardiology Pairs with Smith
Anesthesiology Pairs with Michael
 Unstable Pairing
Anesthesiology removes Michael from pair
 Pairs with next member of their choice
Anesthesiology Pairs with Mathew
 Unstable Pairing
Pathology removes Simons from pair
 Pairs with next member of their choice
Pathology Pairs with Michael
Oncology Pairs with Cynthia
Cardiothoractic_surgery Pairs with Simons
 Resultant Pairing of the Partners is as follows        :
     Michael - Pathology
       Adams - Endocrinology
     Cynthia - Oncology
      Simons - Cardiothoractic_surgery
  Gillchrist - Radiology
        Fred - Neurology
     Britney - Pediatric_surgery
      Mathew - Anesthesiology
      Taylor - Surgical_surgery
       Smith - Cardiology
```

VIII. Difficulties

The relationship of the problem with graph theory is not discussed in any of the papers. Perhaps using the graph theory techniques to solve did not yield efficient time optimal solutions. We had read additional articles and put in some time to understand how graph theory motivated the modeling of such problems and how solution approaches where also inspired through graph theory.

Paper was lacking with proper visualization of dominating and un-dominated pairs made it difficult to generate the immediate neighborhood of a random marriage.

REFERENCES

[1] A. E. Roth. The evolution of the labor market for medical interns and residents: a case study in game theory. In J. of Political Economy 92:991-1016, 1984.

[2] E Ronn. ‘NP-complete stable matching problems’ - Journal of Algorithms, 1990

[3] Gale. D., and Shapley. L. S. College admissions and the stability of marriage. Am. Math. Monthly 69 (1962), 9- 15.

[4] McVitie and Wilson, The stable marriage problem. Communications of the ACM,14(1971).pp.486-492

[5] Rob Irving, P. Leather and D. Gusfield .An efficient algorithm for the "optimal" stable marriage. Journal of the ACM, 34 (3) (Jul. 1987), pp. 532-543

[6] M. Gelain, M.S. Pini, F. Rossi, K.B.Venable and T.Walsh, “Local Search for Stable Marriage Problems with Ties and Incomplete Lists”, In Proc. of PRICAI 2010, pp. 64-75, Springer-Verlag

[7] I.P. Gent and P. Prosser, “An empirical study of the stable marriage problem with ties and incomplete lists”, Proceedings of the 15th European Conference on Artificial Intelligence, ECAI'2002, Lyon, France

[8] The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2012 .Alvin E. Roth, Lloyd S. Shapley.Nobelprize.org.

[9] Jennie Premer, Fort Calhoun, NE. Stable Marriage Problem. in partial fulfillment of the MAT Degree. Department of Mathematics. University of Nebraska-Lincoln. Summer 2011

[10] Emily Riehl ‘’A solution to the stable marriage problem’.HarvardUniversity

[11] Eytan Ronn “NP complete stable matching Problem”. Computer Science Depurtment, Technion-Isruel Institute of Technology, Haifa 3X00, Ismel Received October 10,1986; revised October 16,1988