Optimal Selection of Training Courses for Unemployed People based on Stable Marriage Model

Jorge Martinez-Gil
Software Competence Center Hagenberg GmbH
Hagenberg, Austria
jorge.martinez-gil@scch.at

Bernhard Freudenthaler
Software Competence Center Hagenberg GmbH
Hagenberg, Austria
bernhard.freudenthaler@scch.at

ABSTRACT

The problem that we address here is given *n* job seekers and *n* job offers, where each job seeker has ranked all job offers in order of preference given by a suitability function, and vice versa; the goal is to compute the minimum set of skills to be offered to the job seekers, so that a) a global stable marriage between job seekers and potential employers can be reached, and b) the degree of satisfaction for that stable marriage might be maximum. To achieve this goal, we have designed an iterative algorithmic solution that can be solved in polynomial time. Additionally, we illustrate our solution with an use case based on a numerical example.

CCS CONCEPTS

• **Information systems** → **Information systems applications**; • **Information Systems** → *Data Mining*; • **Retrieval tasks and goals** → Information extraction;

KEYWORDS

Information Retrieval, Knowledge Engineering, Stable Marriage, e-Recruitment

ACM Reference Format:

Jorge Martinez-Gil and Bernhard Freudenthaler. 2019. Optimal Selection of Training Courses for Unemployed People based on Stable Marriage Model. In *Proceedings of The 21st International Conference on Information Integration and Web-based Applications & Services, Munich, Germany, December 2–4, 2019 (iiWAS2019),* 7 pages.
https://doi.org/10.1145/3366030.3366063

1 INTRODUCTION

In this context, a stable marriage model consists of a number of job applicants and potential employers with strict preferences over a possible labor relation. A stable marriage model is a matching of the labor graph so that no job applicant and employer exists who mutually prefer other than their current election to start an eventual labor relation. Gale and Shapley proved that their algorithm always finds such a matching model for any preference model of the parties [8].

iiWAS2019, December 2–4, 2019, Munich, Germany

ACM ISBN 978-1-4503-7179-7/19/12…$15.00
https://doi.org/10.1145/3366030.3366063

In this work, we work over a variant of the Gale and Shapley's stable marriage problem. In fact, we focus on a job market whereby *n* employers request a person holding a number of skills and competences, and *n* unemployed people offer their skills and competences in order to get a job. We assume that it is more plausible for job applicants to get new skills and competences than for employers to change their viewpoint on what their business might need. Therefore, the problem consists of finding the optimal selection of training courses for unemployed people, so that the job market can be optimally satisfied. This means that our proposed algorithm should be able of not only finding a stable marriage between a set of unemployed people and a set of potential employers but to additionally find the optimal degree of global satisfaction by spending the minimum amount of resources in the form of training courses.

Therefore, the contribution of this work can be summarized as follows:

- We design an algorithmic solution being able to perform an optimal selection of training courses for unemployed people that presents two major advantages: a) the solution is stable (no job applicant and potential employer would mutually prefer other than their identified counterpart), and b) an optimal degree of systemic satisfaction is guaranteed.
- We demonstrate the feasibility of such as algorithmic solution by means of a procedural implementation, and we show a complete running case study based on that implementation that solve a prototypical scenario in the IT sector.

The rest of this paper is structured in the following way: Section 2 presents related work on different approaches concerning the stable marriage problem and some interesting variants that have been already studied. Section 3 formally presents the problem of optimal selecting those training courses that can optimally meet the expectations of both companies and unemployed workers. In Section 4, we explain the algorithmic solution that we have designed to solve the problem and compute its associated computational complexity. Section 5 shows how our solution works by illustrating a case study concerning a particular recruitment scenario. Section 6 discusses our contribution. Finally, we provide our concluding remarks and hints for future work.

2 RELATED WORKS

There is a large corpus of literature on the problem of stable marriage and its variants, mainly due to the number of practical applications that can be addressed using a model of this kind [5]. Gale and Shapley presented the classic problem between single men and women looking to form a marriage [8]. Same authors also introduced the problem of the university admissions that differs from the classic problem in which one side could accept proposals of more than one counterpart [8]. Another analog problem considers the fact that one side of the market is made up of colleges, the other side is made up of students, but there is a particularity in relation to college fees to be paid [24]. Additionally, it also worth mentioning the problem of the two bodies, whereby a number of couples should be allocated together, either to the same resource or to a specific pair of resources mutually chosen by them [11].

Other relevant works in this context involve matching job seekers and companies with contracts in which one side can be matched under a variety of contractual conditions [12]. A special case of the contracts could be the possibility to offer flexible remuneration [4]. If we focus on the context of computational solutions in the field of human resources management, our previous works has consisted of considering an underlying lattice as a knowledge base that overcome the problem of semantic heterogeneity and facilitates the accurate and efficient learning [20], matching [21], and querying [22] information on human resources in a number of different scenarios [1]. At the same time, Guedj has also proposed a model to deduce initial preferences on the basis of the profiles viewed as nodes from a taxonomy [9], [10]. The aim here is to offer a list of job positions to a potential job applicant based on preferences, or alternatively, to generate a list of job candidates to a recruiter based on the job requirements. In addition, Cabrera-Diego et al. [2], [3] propose different methods for ranking profiles without using job offers nor any kind of semantic resource. Espenakk et al. [6] propose a method to rank candidates based on insights from previous screening phases. Khobreh et al. [15] proposes an approach for consolidationg abilities and skills resulting from vocational education. In addition, Schönböck et al. [23] propose a framework to determine ranked lists of volunteers from marketplaces. Finally, it is also worth mentioning the works of Tinelli et al. for dealing with semantics in the context of e-recruitment [25].

On the other hand, in this research work, we aim to go a step further by proposing how to calculate the optimal selection of training courses for obtaining a maximum degree of satisfaction in overall terms. We think this method could have an impact on such organizations as public and private employment organizations or even large companies or project teams wishing to plan in advance the training needs that best fit to their strategy, as set out in [13].

3 PROBLEM STATEMENT

The problem that we address here can be formally defined as follows:

- There are two finite and disjoint sets $A = \{a_1, a_2, ..., a_n\}$ of job applicants and $E = \{e_1, e_2, ..., e_m\}$ of employers.
- Each job applicant has ordered, transitive, and complete preferences over the employers, and each employer has ordered, transitive, and complete preferences over the job applicants. The ordered list for a job applicant is represented as $P(a_i) = (e_{j_1}, e_{j_2}, ..., e_{j_n})$ and for a employer $P(e_i) = (a_{j_1}, a_{j_2}, ..., a_{j_n})$, and it can be calculated by means of a suitability function f that we assume to be stable.
- A *job market* is a tuple $(A, E, \mathbf{P}, \mathbf{M_p})$, where $\mathbf{P}$ is the set of all the preferences of all the parties, and $\mathbf{M_p}$ is the set of established relations taking into account the preferences P.
- An job market $(A, E, \mathbf{P}, \mathbf{M_p})$ is stable when there are no $a \in A$ nor $e \in E$ that would both rather have each other than their current counterpart in $\mathbf{M_p}$.
- An job market $(A, E, \mathbf{P}, \mathbf{M_p})$ is perfect if and only if each $a \in A$ is paired in M with the $e \in E$ that it prefers most of all possible employers AND each $e \in E$ is paired in M with the $a \in A$ that it prefers most of all possible job applicants.
- The distance between two job markets $(A, E, \mathbf{P}, \mathbf{M_p})$ and $(A, E, \mathbf{Q}, \mathbf{N_q})$ is the difference between $\mathbf{M_p}$ and $\mathbf{N_q}$.
- An *optimal job market* is a tuple $(A, E, \mathbf{P}, \mathbf{M_p})$, so that the distance to a perfect job market $(A, E, \mathbf{P}', \mathbf{M}'_{\mathbf{p}'})$ is minimal (i.e. mean, it is not possible to find a shorter distance.

Assuming that:

(1) Employers E will not change the requirements what they are looking for during the whole process
(2) The suitability function f will remain the same for all job seekers and during the whole process
(3) Job applicants A taking training courses acquire the new skills and competences at the end of those courses

Our aim in this work is to build an algorithmic solution for automatically finding the minimum set of skills and competences to be offered to job applicants A in the form of training courses, so that f can generate a job market $(A, E, \mathbf{P}, \mathbf{M_p})$ meeting the following requirements:

- $(A, E, \mathbf{P}, \mathbf{M_p})$ is stable
- $(A, E, \mathbf{P}, \mathbf{M_p})$ is optimal

Our hypothesis is that a solution of this kind can have potential to make an impact in, at least, such organizations as public and private employment agencies wishing to plan in advance the training courses that best fit to their strategic needs.

4 CONTRIBUTION

We present here our contribution to face the problem. First of all, we focus on the design of the algorithmic approach, and secondly, on the computation of the associated complexity.

4.1 Design of the algorithmic solution

Our algorithmic solution works as follows:

- In the first round, we use a suitability function considering two facts: (a) each job applicant calculates its suitability function f_a, i.e. how well its profile fits in relation to each job offer, (b) each employer calculates its suitability function f_e, i.e. how well its offer fits in relation to each job applicant profile. To do that, we can use a measure such as profile matching [19].

$$f_a(a, e) = \# \frac{Prof(a) \cap Prof(e)}{Prof(a)} \ (a)$$

$$f_e(a, e) = \# \frac{Prof(a) \cap Prof(e)}{Prof(e)} \ (b)$$

- In the second round, we proceed to order the preference list of each applicant and each employer by considering the results of the first round. The order is descending, what means that the highest value ranks first, the second highest value ranks second, and so on. In case of tie, we order by index (by convention).

- In the third round, we proceed in the following way: (i) each job applicant proposes to its most preferred employer, (ii) each employer then "conditionally accepts" its most preferred job applicant among those who proposed to it, (iii) all the job applicants whose proposal was not accepted remain unemployed.
- In each following round, (i) each unemployed job applicant proposes to its most preferred employer whom it has not yet proposed to, (ii) each employer retains its most preferred job applicant among those who proposed to and its current corresponding applicant, (iii) all the job applicants whose proposal was not conditionally accepted or who were "dumped" are unemployed. This loop has to run until there are no unemployed job applicants[1]. At this moment, the marriage is stable as proved by Gale and Shapley [8].
- In the fourth round, we compute the distance of the obtained job market to the perfect one (i.e. that market where every party got a labor relation with its first preference). This distance is the difference between $\mathbf{M_p}$ and $\mathbf{N_q}$, and we define it as follows:

$$(A, E, \mathbf{P}, \mathbf{M_p}) - (A, E, \mathbf{P'}, \mathbf{M'_{p'}}) =$$

$$\mathbf{M_p} - \mathbf{M'_{p'}} = \sum_{i=0}^{n} (M_p a_i + M_p e_i) - 2 \cdot i$$

- In the fifth (and final) round, we compute the distance between a perfect job market and every job market generated by the permutations of the set of all skills demanded by the employers from the job market[2]. The number of permutations of n elements (where n is the count of all skills and competences requested by the employers) is given by the formula:

$$\binom{n}{1} + \cdots + \binom{n}{n-1} + \binom{n}{n} = 2^n - 1$$

The permutation leading to a minimum distance to a perfect market is the solution that we are looking for, i.e. the set of skills and competences that achieves a stable marriage and an optimal market at the same time. In case two or more solutions might be found, we always prefer that solution associated to the smaller set of skills and competences (what it will make us saving resources in the form of time, money or effort). In case we found a distance of 0 between both job markets, we can stop the algorithm since we are sure that we cannot achieve a better solution, but please note that this only applies in the case whereby the solution might start computing by the smaller sets of permutations. Otherwise, there is always the possibility that we might find an equally good but cheaper solution.

[1]It is important to note here that this iterative process is not symmetric, but it could be easily modified in order to the employers propose a labor relationship to the unemployed people

[2]The reason is that we assume that it makes no sense considering training courses for skills and competences not demanded by the employers

4.2 Computational complexity

We consider here the Big O approach since this notation can give us some hints on the growth rate that our algorithmic solution presents, and thus, its capability to work in scenarios requiring very large inputs. Therefore, we have that the first, second, and fourth rounds can be solved in linear time, and therefore, they present $O(n)$ where n is number of job seekers or employers. Iwama and Mayazki shows that the computational complexity for our third round is $n(n-1)$ what in practice means $O(n^2)$ for the worst case [14]. Finally, the fifth round can be represented by just a loop where the rest of the functionality is nested. Therefore, we have a complexity of $O(n)$. Since the first, second, third and fourth rounds are nested on this last round, we have that the overall complexity of the proposed solution is $O(n^3)$, what means that the solution is suitable to be operated in real settings.

5 CASE STUDY

In order to illustrate how our method works in practice, we show here a case study concerning a recruitment scenario in the field of software development[3]. Let us suppose we have a pool of unemployed people who have been working as programmers in the past. At the same time, we have a pool of potential employers looking for filling some programming positions that they need for accomplishing some tasks related to their business. We want to know what is the minimum set of skills and competences that could be offered in the form of training courses so that we can reach an optimal job market as defined in Section 3.

For practical reasons, we show here just two iterations of our method; firstly a single iteration of the proposed algorithm, and secondly the iteration leading to the solution of the problem. However, the evolution of the job market through several iterations is trivial, and it can be studied in depth by following the steps already indicated.

5.1 First and second round

First of all, we are considering the requirements of the job applicant profiles and job offers as we can see in Table 1 and Table 2 respectively. For each of these parties, we have to calculate each suitability function so that we can obtain every preference list.

Job Seeker	Competences
Applicant 1	C, C++, C#
Applicant 2	Javascript, CSS, HTML, Java
Applicant 3	Python, R, C, SQL
Applicant 4	C, C++, SQL, Java, Scala

Table 1: Initial set of skills and competences of each unemployed person

5.2 Third and subsequent rounds

The third round starts with the information that it is shown in Table 3 and Table 4. Initially, each job applicant proposes to its most preferred employer, each employer then "conditionally accepts"

[3]Please note that this method would be exactly the same in any of the other fields

Employer	Requirements
Company 1	C, C++
Company 2	Python, R, Julia, Ruby
Company 3	Java, Javascript, SQL
Company 4	Scala, Perl

Table 2: Initial set of skills and competences required by employers

its most preferred job applicant among those who proposed to it, and all the job applicants whose proposal was not accepted remain unemployed, etc. As explained before, this stage is a well-known generalization of the stable marriage problem from Gale and Shapley [8]. This loop has to run until there are no employers left that can offer a job as mentioned in Section 3.1.

Job Seeker	Employer
Applicant 1	Company 1 (2/3), Company 2 (0/3), Company 3 (0/3), Company 4 (0/3)
Applicant 2	Company 3 (2/4), Company 1 (0/4), Company 2 (0/4), Company 4 (0/4)
Applicant 3	Company 2 (2/4), Company 1 (1/4), Company 3 (1/4), Company 4 (0/4)
Applicant 4	Company 1 (2/5), Company 3 (2/5), Company 4 (1/5), Company 2 (0/5)

Table 3: Initial preference list of each unemployed person

Employer	Job Seeker
Company 1	Applicant 1 (2/2), Applicant 4 (2/2), Applicant 3 (1/2), Applicant 2 (0/2)
Company 2	Applicant 3 (2/4), Applicant 1 (0/4), Applicant 2 (0/4), Applicant 4 (0/4)
Company 3	Applicant 2 (2/3), Applicant 4 (2/3), Applicant 3 (1/3), Applicant 1 (0/3)
Company 4	Applicant 4 (1/2), Applicant 1 (0/2), Applicant 2 (0/2), Applicant 3 (0/2)

Table 4: Initial preference list of each employer

5.3 Fourth round

As a result of the prior computation rounds, we have achieved a stable marriage (see Table 5). Therefore, this means that we can now proceed to compute the distance from the generated job market to a perfect job market, i.e. whereby each party involved might get its first choice:

$$\mathbf{M_p} - \mathbf{M'_{p'}} = \sum_{i=0}^{n}(M_p a_i + M_p e_i) - 2 \cdot i$$
$$= 15 - 8 = 7$$

The overall degree of satisfaction for the given job market is 7 units far from a perfect job market. In next rounds, we will see if this value can be improved by offering to the job applicants the possibility to take training courses with the aim to get new competences.

Figure 1 shows us the initial assignment for applicants and potential companies whereby no training course has yet been offered. This assignment can be obtained through the traditional Gale and Shapley approach.

Job Seeker	Employer	Job seeker's Pref.	Employer's Pref.
Applicant 1	Company 1	1	1
Applicant 2	Company 4	4	3
Applicant 3	Company 2	1	1
Applicant 4	Company 3	2	2

Table 5: Final stable matching. The distance of the job market generated by this matching to a perfect job market is 15 - 8 = 7

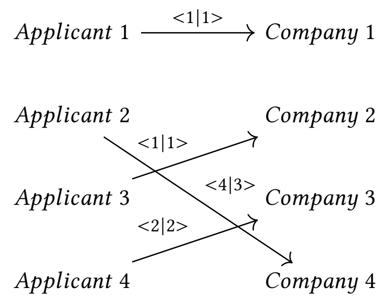

Figure 1: Initial job assignment for job applicants and potential companies whereby no training course has yet been offered

5.4 Fifth round

Now, we want to know what it happens when we offer the possibility of learning the language SQL[4] to those job applicants who do not master it yet. Please note that this is just one iteration.

Job Seeker	Competences
Applicant 1	C, C++, C#, **+SQL**
Applicant 2	Javascript, CSS, HTML, Java, **+SQL**
Applicant 3	Python, R, C, *SQL*
Applicant 4	C, C++, *SQL*, Java, Scala

Table 6: Skills and competences of each unemployed person after learning SQL

Let us see the behavior of our proposed algorithmic solution when evaluation this iteration:

(1) Job applicants not mastering SQL yet, acquire this new skill, as we can see in Table 6.
(2) Every job applicant has to re-calculate its preference list, as we can see in Table 7.
(3) Each employer has to re-calculate its preference list too, as we can see in Table 8.
(4) We compute the third and subsequent rounds by means of Gale and Shapley as discussed, and we achieve a stable matching.

[4]There is no particular reason to choose this skill, we just pick a random skill in order to show how our method works

Job Seeker	Employer
Applicant 1	Company 1 (2/4), Company 3 (1/4), Company 2 (0/4), Company 4 (0/4)
Applicant 2	Company 3 (3/5), Company 1 (0/5), Company 2 (0/5), Company 4 (0/5)
Applicant 3	Company 2 (2/4), Company 1 (1/4), Company 3 (1/4), Company 4 (0/4)
Applicant 4	Company 1 (2/5), Company 3 (2/5), Company 4 (1/5), Company 2 (0/5)

Table 7: Updated preference list of each unemployed person after learning SQL

Employer	Job Seeker
Company 1	Applicant 1 (2/2), Applicant 4 (2/2), Applicant 3 (1/2), Applicant 2 (0/2)
Company 2	Applicant 3 (2/4), Applicant 1 (0/4), Applicant 2 (0/4), Applicant 4 (0/4)
Company 3	Applicant 2 (3/3), Applicant 4 (2/3), Applicant 1 (1/3), Applicant 3 (1/3)
Company 4	Applicant 4 (1/2), Applicant 1 (0/2), Applicant 2 (0/2), Applicant 3 (0/2)

Table 8: Updated preference list of each employer after training courses on SQL were performed

Job Seeker	Employer	Job seeker's Pref.	Employer's Pref.
Applicant 1	Company 1	1	1
Applicant 2	Company 3	1	1
Applicant 3	Company 2	1	1
Applicant 4	Company 4	3	1

Table 9: Final stable matching. The distance of the job market generated by this matching to a perfect job market is 10 - 8 = 2

(5) Once that we have achieved the stable matching, we calculate the distance to a perfect market (see Table 9).

$$\mathbf{M_p} - \mathbf{M'_{p'}} = \sum_{i=0}^{n} (M_p a_i + M_p e_i) - 2 \cdot i$$
$$= 10 - 8 = 2$$

The distance has been reduced to 2 units. This means that by offering a training course on the language SQL, we have been able to increase the global satisfaction of the parties with respect to the initial stable marriage, which is already a step forward. However, we want to keep looking for the optimum solution.
Figure 2 shows us the temporary global assignment of applicants to potential companies after the acquisition of the skill SQL by all the applicants that did not have that skill before.

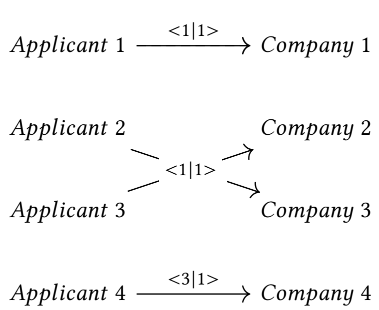

Figure 2: Temporary assignment for applicants and companies after the acquisition of the skill SQL by all the applicants

(6) Now, it is time to evaluate each permutation of the skills and competences requested by the employers. Our set of skills and competences is the set of all those that have been required at least once by at least one employer[5]. Therefore, it is necessary to consider the set {*C, C++, C#, Javascript, CSS, HTML, Java, Python, R, SQL, Scala*}. This means that we have to produce all the possible permutations for 11 different skills and competences. Therefore,

$$\binom{11}{1} + \cdots + \binom{11}{10} + \binom{11}{11} = 2^{11} - 1$$
$$= 2047 \ permutations$$

Since we have already evaluated the permutation *SQL*, we need to still evaluate each of the pending 2046 different permutations. The permutation that is able to get a minimum distance to the perfect job market is our solution. Please note, that it is perfectly possible that we might find more than one solution. In that case, and for economic reasons, we prefer the solution associated to the smaller set of skills and competences as mentioned in Section 3.1.

5.5 Optimal solution

Now, we are going to explore how to reach the optimal solution for our case study. To do that, we have obtained that optimal solution by using an implementation of our method over the 2047 possible permutations[6]. As a result, we know that the optimal (and cheapest) solution is achieved when offering the set {Perl, SQL} in the form of training courses. Please note, that it is perfectly possible that we might find more than one solution, but we can assure you these solutions will not be cheaper. Let us see the behavior of the proposed method when achieving that solution.

(1) Firstly, we have to offer the possibility of learning Perl and SQL to those candidates that do not master these skills and competences yet, as we can see in Table 10.
(2) Every job applicant has to re-calculate its preference list, as we can see in Table 11.
(3) Each employer has also to re-calculate its preference list too, as we can see in Table 12.
(4) We compute the third subsequent rounds by means of Gale and Shapley until there are no employer left that can offer a job, as recurrently discussed before.

Job Seeker	Competences
Applicant 1	C, C++, C#, **+SQL**, **+Perl**
Applicant 2	Javascript, CSS, HTML, Java, **+SQL**, **+Perl**
Applicant 3	Python, R, C, *SQL*, **+Perl**
Applicant 4	C, C++, *SQL*, Java, Scala, **+Perl**

Table 10: Skills and competences of each unemployed person after learning Perl and SQL

And finally, in the fourth round of this particular iteration, we have achieved the stable matching between unemployed people

[5]Because considering other elements would have no positive effect.
[6]In fact, there is no alternative way to shorten the process of getting the solution

Job Seeker	Employer
Applicant 1	Company 1 (2/5), Company 3 (1/5), Company 4 (1/5), Company 2 (0/5)
Applicant 2	Company 3 (3/6), Company 4 (2/6), Company 1 (0/6), Company 2 (0/6)
Applicant 3	Company 2 (2/5), Company 3 (1/5), Company 4 (1/5), Company 1 (1/5)
Applicant 4	Company 3 (2/6), Company 4 (2/6), Company 1 (1/6), Company 2 (1/6)

Table 11: Updated preference list of each unemployed person after learning Perl and SQL

Employer	Job Seeker
Company 1	Applicant 1 (2/2), Applicant 4 (2/2), Applicant 3 (1/2), Applicant 2 (0/2)
Company 2	Applicant 3 (2/4), Applicant 1 (0/4), Applicant 2 (0/4), Applicant 3 (0/4)
Company 3	Applicant 2 (3/3), Applicant 4 (2/3), Applicant 1 (1/3), Applicant 3 (1/3)
Company 4	Applicant 4 (2/2), Applicant 1 (1/2), Applicant 2 (1/2), Applicant 3 (1/2)

Table 12: Updated preference list of each employer after training courses on Perl and SQL were performed

looking for an opportunity and potential employers, as we can see in Table 13.

Job Seeker	Employer	Job seeker's Pref.	Employer's Pref.
Applicant 1	Company 1	1	1
Applicant 2	Company 3	1	1
Applicant 3	Company 2	1	1
Applicant 4	Company 4	2	1

Table 13: Final solution for our case study. The distance of the job market generated by offering {Perl, SQL} as training courses to a perfect job market is 9 - 8 = 1

Therefore, it is now time to calculate the distance from the obtained job market to a perfect job market, i.e. whereby we have the ideal situation in which every party involved in the matching process gets its first choice. This distance between job markets is calculated as follows:

$$\begin{aligned}\mathbf{M_p} - \mathbf{M'_{p'}} &= \sum_{i=0}^{n} (M_p a_i + M_p e_i) - 2 \cdot i \\ &= 9 - 8 = 1\end{aligned}$$

What means that, after this iteration, we can see that the overall distance has been reduced to 1 unit. This means that by offering training courses on such programming languages as Perl and SQL, we have been able to decrease the distance 6 units from the initial matching that we calculated. We already know that by evaluating each of the pending permutations, we will not be able to reach a lower distance. Moreover, we have started computing the smaller sets of permutations, so we can assure you that there is no less expensive solution. Therefore, the set of training courses **{Perl, SQL}** is the solution that we were looking for satisfactorily solving this use case.

Figure 3 shows us the final global assignment of applicants to potential companies after the acquisition of the skills Perl and SQL by all the applicants that did not have those skill before. In this way, the given job market is optimal (although not perfect), since we cannot find a configuration with a smaller distance to the ideal job market.

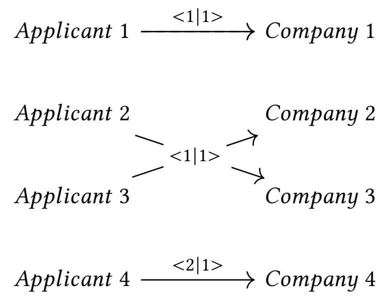

Figure 3: Final assignment for applicants and potential employers after the acquisition of the skills Perl and SQL by all the applicants

6 DISCUSSION

As we have seen, our proposed method is able to find in polynomial time an optimal stable marriage between a set of unemployed people and potential employers. Due to the initial conditions of the problem, it will not be always possible to find a perfect job market. In fact, from the results of our case study, we can see that Applicant 4 was not assigned to its first choice, but its second. This means that given the initial input scenario, it has not been possible to reach a perfect job market. However, it is important to note that we have been able to reach a systemic optimum. This means that we have reached the maximum degree of global satisfaction for the job market that we have analyzed. In that case, this maximum degree of global satisfaction is just one unit far from a perfect job market, and it has been achieved by offering just two training courses to the unemployed people.

It is also interesting to remark that from this work, some interesting further questions could be derived. For example, one could question itself if the proportion of global satisfaction gained over the resources spent to achieve that goal it is favorable or not, cecause there may be cases in which a small global improvement involves a large economic investment. Another possible question might be who should bear the cost of the training courses that make the job market optimal. For obvious reasons, this requires addressing those questions from a pure subjective perspective, and therefore, it is out of the scope of this research work.

7 CONCLUSIONS AND FUTURE WORK

In this work, we have presented a novel approach for automatically finding the optimal selection of training courses for unemployed people based on a stable marriage model. This approach is able of not only finding a stable marriage between a set of unemployed people and potential employers but to find the optimal degree of global satisfaction for the system by spending the minimum amount of resources in the form of training courses. Moreover, we have shown how such solution can be found in polynomial time, so it can be put into production in real environments.

The proposed approach can have an impact in such organizations as public and private employment agencies or large public and private organizations wishing to plan in advance the training courses that best fit to their strategic needs, or even can be exploited at smaller scale, in order to help project-oriented companies in the process of preparing their employees to work in future projects [25]. In general, this method can work in every scenario meeting our initial assumptions.

As future work, it is necessary to further investigate the suitability function to generate the list of preferences of each party involved. We have used here an intuitive function that computes the degree of overlapping between applicant profiles and job offers by just counting the number of skills and competences in the intersection of both profiles. A further improvement could consist of using suitability functions being able to semantically identify what the employer requires and what the applicant offers such as those proposed by [7] and [17], or by computing semantic similarity [18] between them, or even injecting background knowledge into the problem [16]. For example, if an employer needs a good programmer in Java, candidates with strong expertise in C++ should not be strongly penalized, since both programming languages have a lot of features in common.

ACKNOWLEDGEMENTS

We would like to thank the anonymous reviewers for their time and consideration. The research reported in this paper has been supported by the Austrian Ministry for Transport, Innovation and Technology, the Federal Ministry of Science, Research and Economy, and the Province of Upper Austria in the frame of the COMET center SCCH.

REFERENCES

[1] Buga, A., Freudenthaler, B., Martinez-Gil, J., Nemes, S. T., and Paoletti, A. L. (2017). Management of accurate profile matching using multi-cloud service interaction. In *Proceedings of the 19th International Conference on Information Integration and Web-based Applications & Services, iiWAS 2017, Salzburg, Austria, December 4-6, 2017*, pages 161–165.

[2] Cabrera-Diego, L. A., El-Bèze, M., Torres-Moreno, J., and Durette, B. (2019). Ranking résumés automatically using only résumés: A method free of job offers. *Expert Syst. Appl.*, 123:91–107.

[3] Cabrera-Diego, L. A. and Torres-Moreno, J. (2018). Summtriver: A new trivergent model to evaluate summaries automatically without human references. *Data Knowl. Eng.*, 113:184–197.

[4] Crawford, V. and Knoer, E. M. (1981). Job matching with heterogeneous firms and workers. *Econometrica*, 49(2):437–50.

[5] Dias, V. M. F., da Fonseca, G. D., de Figueiredo, C. M. H., and Szwarcfiter, J. L. (2003). The stable marriage problem with restricted pairs. *Theor. Comput. Sci.*, 306(1-3):391–405.

[6] Espenakk, E., Knalstad, M. J., and Kofod-Petersen, A. (2019). Lazy learned screening for efficient recruitment. In *Case-Based Reasoning Research and Development - 27th International Conference, ICCBR 2019, Otzenhausen, Germany, September 8-12, 2019, Proceedings*, pages 64–78.

[7] Faliagka, E., Iliadis, L. S., Karydis, I., Rigou, M., Sioutas, S., Tsakalidis, A. K., and Tzimas, G. (2014). On-line consistent ranking on e-recruitment: seeking the truth behind a well-formed CV. *Artif. Intell. Rev.*, 42(3):515–528.

[8] Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.

[9] Guedj, M. (2016a). Levelized taxonomy approach for the job seeking/recruitment problem. In *2016 IEEE Intl Conference on Computational Science and Engineering, CSE 2016, and IEEE Intl Conference on Embedded and Ubiquitous Computing, EUC 2016, and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering, DCABES 2016, Paris, France, August 24-26, 2016*, pages 448–451.

[10] Guedj, M. (2016b). Ranking preferences deduction based on semantic similarity for the stable marriage problem. In *15th IEEE International Conference on Cognitive Informatics & Cognitive Computing ,ICCI*CC 2016, Palo Alto, CA, USA, August 22-23, 2016*, pages 530–534.

[11] Gusfield, D. and Irving, R. W. (1989). *The Stable marriage problem - structure and algorithms*. Foundations of computing series. MIT Press.

[12] Hatfield, J. W. and Kojima, F. (2010). Substitutes and stability for matching with contracts. *J. Economic Theory*, 145(5):1704–1723.

[13] Heidemann, J., Muschter, S., and Rauch, C. (2013). How to increase public e-services usage in governments - A case study of the german federal employment agency. In *21st European Conference on Information Systems, ECIS 2013, Utrecht, The Netherlands, June 5-8, 2013*, page 128.

[14] Iwama, K. and Miyazaki, S. (2008). A survey of the stable marriage problem and its variants. *International Conference on Informatics Education and Research for Knowledge-Circulating Society (icks 2008)*, pages 131–136.

[15] Khobreh, M., Ansari, F., Fathi, M., Vas, R., Mol, S. T., Berkers, H. A., and Varga, K. (2016). An ontology-based approach for the semantic representation of job knowledge. *IEEE Trans. Emerging Topics Comput.*, 4(3):462–473.

[16] Martinez-Gil, J. (2014). An overview of knowledge management techniques for e-recruitment. *JIKM*, 13(2).

[17] Martinez-Gil, J., Freudenthaler, B., and Natschläger, T. (2018). Recommendation of job offers using random forests and support vector machines. In Augsten, N., editor, *Proceedings of the Workshops of the EDBT/ICDT 2018 Joint Conference (EDBT/ICDT 2018), Vienna, Austria, March 26, 2018*, volume 2083 of *CEUR Workshop Proceedings*, pages 22–27. CEUR-WS.org.

[18] Martinez-Gil, J., Paoletti, A. L., Rácz, G., Sali, A., and Schewe, K. (2016). Maintenance of profile matchings in knowledge bases. In Bellatreche, L., Pastor, O., Almendros-Jiménez, J. M., and Ameur, Y. A., editors, *Model and Data Engineering - 6th International Conference, MEDI 2016, Almería, Spain, September 21-23, 2016, Proceedings*, volume 9893 of *Lecture Notes in Computer Science*, pages 132–141. Springer.

[19] Martinez-Gil, J., Paoletti, A. L., Rácz, G., Sali, A., and Schewe, K. (2018). Accurate and efficient profile matching in knowledge bases. *Data Knowl. Eng.*, 117:195–215.

[20] Martinez-Gil, J., Paoletti, A. L., and Schewe, K. (2016). A smart approach for matching, learning and querying information from the human resources domain. In *New Trends in Databases and Information Systems - ADBIS 2016 Short Papers and Workshops, BigDap, DCSA, DC, Prague, Czech Republic, August 28-31, 2016, Proceedings*, pages 157–167.

[21] Paoletti, A. L., Martinez-Gil, J., and Schewe, K. (2015). Extending knowledge-based profile matching in the human resources domain. In *Database and Expert Systems Applications - 26th International Conference, DEXA 2015, Valencia, Spain, September 1-4, 2015, Proceedings, Part II*, pages 21–35.

[22] Paoletti, A. L., Martinez-Gil, J., and Schewe, K. (2016). Top-k matching queries for filter-based profile matching in knowledge bases. In *Database and Expert Systems Applications - 27th International Conference, DEXA 2016, Porto, Portugal, September 5-8, 2016, Proceedings, Part II*, pages 295–302.

[23] Schönböck, J., Altmann, J., Kapsammer, E., Kimmerstorfer, E., Pröll, B., Retschitzegger, W., and Schwinger, W. (2018). A semantic matchmaking framework for volunteering marketplaces. In *Trends and Advances in Information Systems and Technologies - Volume 1 [WorldCIST'18, Naples, Italy, March 27-29, 2018]*, pages 701–711.

[24] Shmoys, D. B. and Tardos, É. (1993). An approximation algorithm for the generalized assignment problem. *Math. Program.*, 62:461–474.

[25] Tinelli, E., Colucci, S., Donini, F. M., Sciascio, E. D., and Giannini, S. (2017). Embedding semantics in human resources management automation via SQL. *Appl. Intell.*, 46(4):952–982.