Team #11394 **Problem Chosen: A**

2020 HiMCM Summary Sheet

Solution Abstract

In a world where humans have been leaning even more towards industrialization, employment has been more important due to the ever-increasing demands in empty slots in large companies. We rely on young people to fill in demands; however, they do not have any incentives due to society's current trends.

With the decrease in incentives for employment demands for younger people, we require a specific method to incentivize people to take the jobs to fill the increasing demands of the people in the future.

This paper seeks to provide a ranking algorithm implemented through an application that would provide an easier time for people to select a summer job, which addresses components such as economic and psychological demands on a specific individual from numerous sources, providing incentives from components such as interests and salaries provided to the person.

An application implementation consisting of user-end designs, employer-end designs, and back-end (also known as server-side) designs has also been created to provide a real-life implementation of this ranking algorithm.

It is concluded that this ranking algorithm can provide the best decision for individuals, regardless of circumstance, for the jobs they are supposed to take. Even if it does not break contentions due to the small size of the data in this problem, we believe that this would break contentions in larger data sets. This algorithm also accounts for any non-mathematical demands of the person, concluding that we can provide incentives to people, solving the low young person employment situation.

Table of Contents

Introduction

This paper seeks to investigate a model that would determine the factors that teenagers, who are the people targeted by this problem, apply for summer jobs. This model would then be tested on ten different fictional people with valid data. This paper also facilitates the discussion on the incentives of teenagers in their applications for a summer job.

A Higher Demand in Employment

Humanity has been leaning towards industrialization, which includes the creation of large groups that are tasked to operate in individual facilities (e.g., electrical and power plants) and groups that would generate income for large corporations (e.g., vendors in malls). Thus, there is a higher demand for employment, especially among countries developing under a modernization system (Khayitov & Giyosova, 2020). As more companies (regardless of size) have to follow the economic flow, more people are incentivized to work for them. An example that illustrates the phenomenon above is the figure below, which illustrates the generally increasing job demands in the United States.

Figure 1. Line Graph of the Total Unfilled Job Vacancies in the United States.
Source: (Organization for Economic Co-operation and Development, 2000)

A Decrease in Summer Jobs Taken by Young People

Generally, there has been a decrease in the number of young people taking up employment. There existed a pattern during the pre-2000's when they seemed to have continually high employment rates; however, after the 2000s, this employment rate decreased. It is also essential to know that even though these young people's employment increases over the summer, the percentage increase has generally decreased over the past decades. An example of data that illustrates the phenomenon above is the share of working teenagers (16 to 19-year-old people) in the United States.

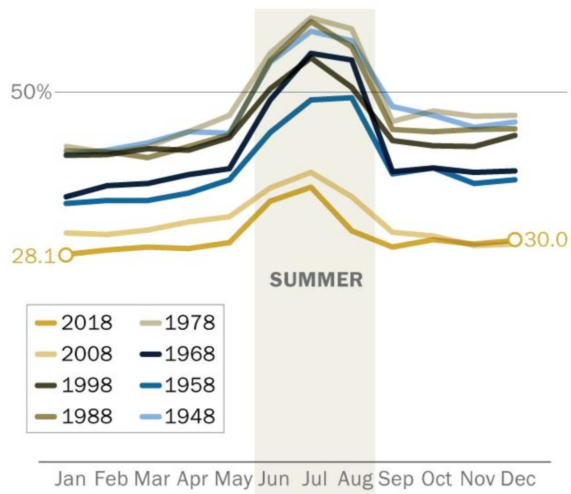


Figure 2. The share of working teens increases during the summer compared to the rest of the year. Source: (Pew Research Center, 2020)

Importance of Summer Jobs

As the people who will be the ones most likely to be employed soon, young people need to have mindsets in getting a job. Numerous research points out that this increases the likelihood of getting a job in the future. It is difficult for individuals, especially these young people who are smaller than large corporations, to get a job after graduation. Therefore, as young people, it is essential to have summer jobs.

Assumptions

We are going to assume the following scenarios about our model:

1. The place where these people will get their jobs is located in the city or any area where there is heavy traffic, which accounts for any significant change in time that it would take to reach the job. This city is also assumed to be a 30 by 30 square lattice (with side length 0.2 km) that would determine the location of individual establishments on lattice points, which replicates the structure of this hypothetical city that is organized into "blocks." This city also only has trains for public transportation.
2. The pay rate, cost of transportation, and any expenses going to and from the job that one will take will always stay constant (disregarding scenarios such as fuel economy, changes in the fuel rate, and changes in train rate). We also assume that there are trains to all destinations in the city without any delay.
3. The person can still select the job that they want from all of the options presented to them on the application, which is still a determining factor in the job they choose.

‘Variables

As the world right now continues to be automated, it is important to create a simple model that would allow young people to find jobs quickly through an online platform with an algorithm that determines the best job.

Factors in Determining Summer Jobs

Our model considers all of the variables that one might think of when applying for any job. This section explains each variable that one might take into account, including justifications for any variable that one would take into account.

- **Type of Job** - This includes what type of jobs one might want to take due to interests. For example, these are Agriculture, Work, Finance, Hotel, Business Management, and others.

- **The Intensity of the Job** - This considers what the person is willing to do (for example, if they do not want to run around, they can work for jobs such as being a concession attendant. However, if they want physical activity, they can take on jobs such as being a lifeguard in a local resort.)
- **Earnings in the Job** - This takes into account the pay rate, excluding any bonuses to make the model simpler. This also considers any expense that a person might need in the job, which includes transportations and meals.
- **Time in the Job** - This includes the number of hours that one has and the time needed to get to the job. For example, on an average day, one can travel 5 km from their home to the job in around 5 minutes. However, when there are deterrent factors such as traffic, this would be factored into the decision.
- **"Enjoy" Factor in the Job** - This determines the "enjoyment factor," which is rated by the people working and consuming goods from the establishment. This can be rated subjectively (with factors that are out of this paper's scope) in terms of a number, with 1 being a job that they dislike the most and 5 being a job they love the most.

Analysis of Probabilistic Variables

We will assume that trains' speed will always be constant to have more straightforward computations, assuming that the train's speed is 120kph. We will also assume that the speed it takes to walk is 5.0kph, and the speed it takes to ride a bicycle is 25.0kph (assuming that the city is entirely flat with no hills). Assume that every speed is constant.

We assume the time that one takes by vehicle depends on accidental circumstances. Therefore, we can use a dataset to determine the average time one can take on a bustling road. In this example, we will use a simulation of Traffic Noise Predictive Models by Iannone et al. (2013). A paper cited in the mentioned paper is consistent with acceptable levels in today's society.

The uncertainty of the time it takes will be randomly chosen among two standard deviations away from the mean of the distributed data (which is the average speed of the vehicle), which is justifiable to take data from 95% of a large data set. The time will then be determined by using the speed-distance-time formula. Analyzing the probability density curve below provides possible data for the vehicle's speed that the one might take. The data below provides a distribution of $\overline{X}$ = 70 kph and s =10 kph.

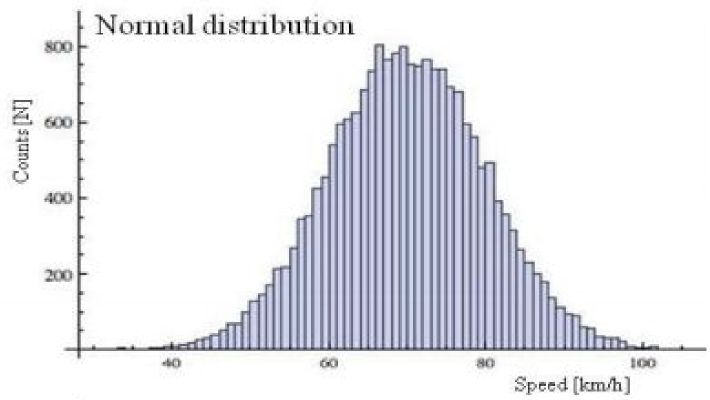


Figure 3. Probability Distribution Curve of Average Speeds in a Simulation (Source: Iannone et al., 2013)

Using this method, we are to take any value from 50 to 90 kph, which is reasonable due to the distribution of the average speeds in the data set.

Model

The model aims to reduce the number of questions that the application will ask to save time on both the user's end and the employer's end. However, our model also provides sufficient information from both users to determine an accurate ranking of the city's jobs. To implement the problem presented above, we would use the following algorithm (on the user's end) to determine the best jobs that the person might take.

User-End Questions

The user will answer a series of questions that would determine quantitative and qualitative factors in ranking the best jobs for them. Questions marked with an asterisk (*) would be recommended to answer to weigh their choices properly. All of the Questions from items 2.1, 2.2, 2.4, and 2.5 are required, while Questions 2.3 and 2.6 are not required.

1. Where do you live? (*) *(The input would be two integers that illustrate the location of their area of residence in this lattice city.)* The results will be stored in two integers x_1 and y_1.
2. What is your mode of transportation for this job? (*) *(The input would be one of four options: walking/running, riding a bike, riding a train, or driving a vehicle. This covers variables like the affordability of modes of transportation, among others.)* The result will be stored as M.
3. What type of job do you prefer? *(The input would involve taking at most three options, which the user will rank from a checkbox list. This would ensure that they get the type of job they want the best.)* The results will be stored as indices J_1, J_2, J_3. We will assume that category J_1 is their most preferred option.
4. How heavy would you like to work? (*) (*The input would be a scale from 1 to 5, where 1 is the most sedentary type of activity and 5 is the most active/least sedentary type of activity.*) The result will be stored as H.
5. Which days of the week would you like to work? (*) (*This ensures that personal spaces such as religious customs are respected.*) The result will be stored in seven variables D_1 to D_7, corresponding to days Sunday to Saturday. If a person is willing to work on a specific day, the variable's value will be equal to 1; otherwise, the value would be equal to 0.
6. How many hours per week would you like to dedicate to this summer job? *(This would ensure that there would be sufficient time allocated to both the job and social activities.)* The input will be stored as the variable T.

Employer-End Questions

The employer and employees can answer a series of questions that would determine quantitative and qualitative factors in ranking the ideal employees for them. Questions marked with an asterisk (*) would be recommended to have an answer to weigh their choices properly. Questions marked with an exponent symbol (^) are questions that only the employer can answer. We assume that there are no other job requirements (e.g., gender, height).

1. How many slots are available for the position? (*^) (*The input would be a non-negative integer.*) The result would be stored in an integer N.
2. How would you describe your work? (*) (*For each category in the data set, they will rank the relevance of their job to that category on a scale from 1 to 5, where 1 is the least relevant and 5 is the most relevant. For example, a cashier position would be ranked as "5" in business management.*) The result would be stored in a dictionary, which we will call W, with string keys and values that correspond to the mean of all respondents' answers.
3. Where is the job workplace located? (*^) *(The input would be two integers that illustrate the location of the employer's location in this lattice city.)* The results will be stored in two integers x_2 and y_2.
4. How much salary do the employees earn? (*^) *(This input would be answered in dollars per hour.)* The results will be stored as r_s.
5. How rigorous would the job be for the employees? (*) (*The input would be a scale from 1 to 5, where 1 is the most sedentary type of activity and 5 is the most active/least sedentary type of activity.*) The mean of the result of all will be stored as H_e.
6. Which days of the week is the workplace open? (*^) *(This ensures that the personal spaces of the teenager are respected.)* The result will be stored in an array of seven variables, E_1 to E_7, corresponding to days from Sunday to Saturday. If a workplace works on a particular day, that variable's value will be equal to 1; otherwise, it would be equal to 0.

7. How many hours of work would be the minimum requirement for the employees per week? (*^) *(This input would be answered preferably as an integer.)* The result will be stored as t.

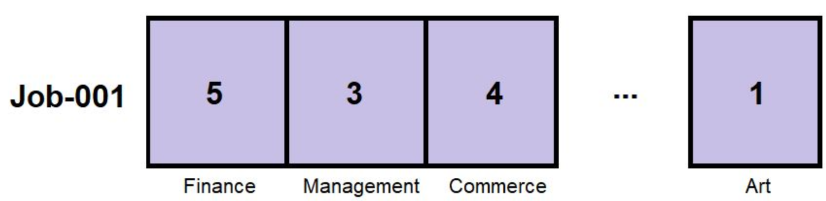
Ranking Algorithm

The algorithm would provide priorities to the different factors of interest. The data for each factor came from the answers that the users and employers provided by answering the questions. These priorities were decided by how much the factor could influence the job choice for the user. The ranking would start in the first step. and all ties and no values in the ranking would be broken using the next step. All jobs that have no slots open would not be included in the ranking.

Step 1. We will rank the relevance of the job you are looking for. To be precise, we will rank the "relevance coefficient" Q_1 below:

$$Q_1 = 3 \times W_{J,1} + 2 \times W_{J,2} + W_{J,3} \ldots (1)$$

If they do not supply a value for any subsequent choice, it will be set to 0. For example, for a user that selects Finance as their first choice, Art as their second choice, and Management as their third choice, the Q_1 score of Job-001 below would be equal to $5 \times 3 + 2 \times 1 + 3 = 20$.



The Relevance Coefficient will guarantee that the jobs they choose are the most suited to their interest. which is important for incentivizing the young generation to increase the number of summer jobs. When one has a passion for work, they are more likely to orient themselves towards being more entrepreneurial (De Clercq et al., 2012).

If all jobs have Q_1 scores that are less than or equal to 8 (which is a reasonable average value for $1 \le W_{J,1},\ W_{J,2},\ W_{J,3} \le 5$), the second choice will become the first choice, and the third choice will become the first. This would repeat until there is a job with $Q_1 \geq 8$. If there are no jobs left, the ranking will proceed to step 2.

Step 2. We will rank the rigor of the job they want, which is variable H (supplied by the user), while H_e is the mean of the respondents' answers. The value that we would use to rank Q_2 would be:

$$Q_2 = |H - H_e| \ldots (2)$$

If a person does not supply a value for this step, they do not have any preference and can proceed to the next step.

Step 3. We will rank the money that a person could get from the job, deducting any expenses that one might have. Computing the total money per week, we get our Q_3 score:

$$-Q_3 = r_s \times t - (r_d \times d + 54) \ldots (3)$$

where r_s is the pay per hour of the job,

t is the number of hours of work per week,

r_d is the cost per kilometer of the vehicle you are using, and

d is the distance from your home to the job, where $d = |x_1 - x_2| + |y_1 - y_2|$.

The vehicle's cost per kilometer includes costs like fuel economy of the vehicle that one is using and the transportation cost for train fare. We assumed that a person spends \$54 per week on food based on research by Hamm (2018).

Step 4. We will rank the time in the job, including travel time and work time. Computing the total time per week, we get the value:

$$Q_4 = 24 - (2 \cdot (\sum d/v)/5 + t/k) \ldots (4)$$

where d is the distance from your home to the job in km,

v can be five randomly generated speeds from 50 kph to 90 kph,

t is the number of hours of work per week, and

k is the number of days that the workplace is open (i.e., the number of integers $1 \leq i \leq 7$ where $E_i = 1$).

The first term accounts for travel time (to and from the workplace), where traffic each day is highly uncertain. The randomness of the value v can take into account the certain average speed in the city considering the unpredictability of traffic in the area. We also assume that they work the same number of hours per day.

Step 5. We will rank the availability of the person, which is essentially:

$$Q_5 = a \ldots \text{(5)}$$

where a is the number of integers $1 \le j \le 7$ that satisfy $D_j = E_j = 1$. This accounts for any time that a person might become available, doing activities such as religion and family recreation.

Step 6. We will rank the vicinity of the job from their home, which is essentially:

$$Q_6 = d = |x_1 - x_2| + |y_1 - y_2| \ldots \text{(6)}$$

This will be the distance of their home from the job.

Step 7. We will rank the number of hours that one would want to work. Essentially, we will rank the value:

$$Q_7 = |T - t| \ldots \text{(7)}$$

where T is the number of hours the employer requires per week, and t is the number of hours that the user wants to work per week.

Application Design

The application model would be straightforward, which is a mobile application that would determine the client's best job. Provided below and the next pages are the flowcharts of the application's user-end, employer-end, and back-end designs.

User-End Design

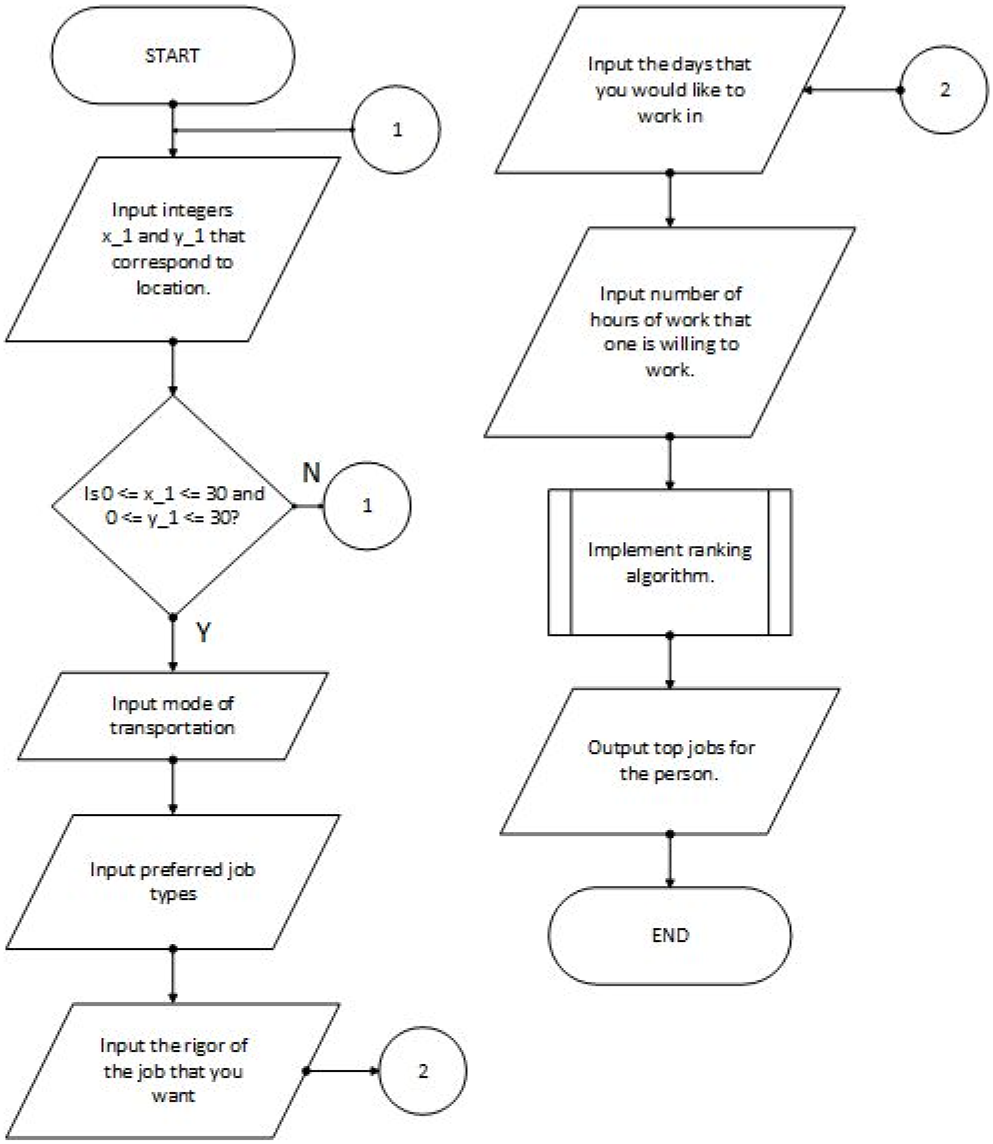


Figure 4. User-End Algorithm of the Application.
(Source: A member of Team #11394)

Employee-End Value Updating Design

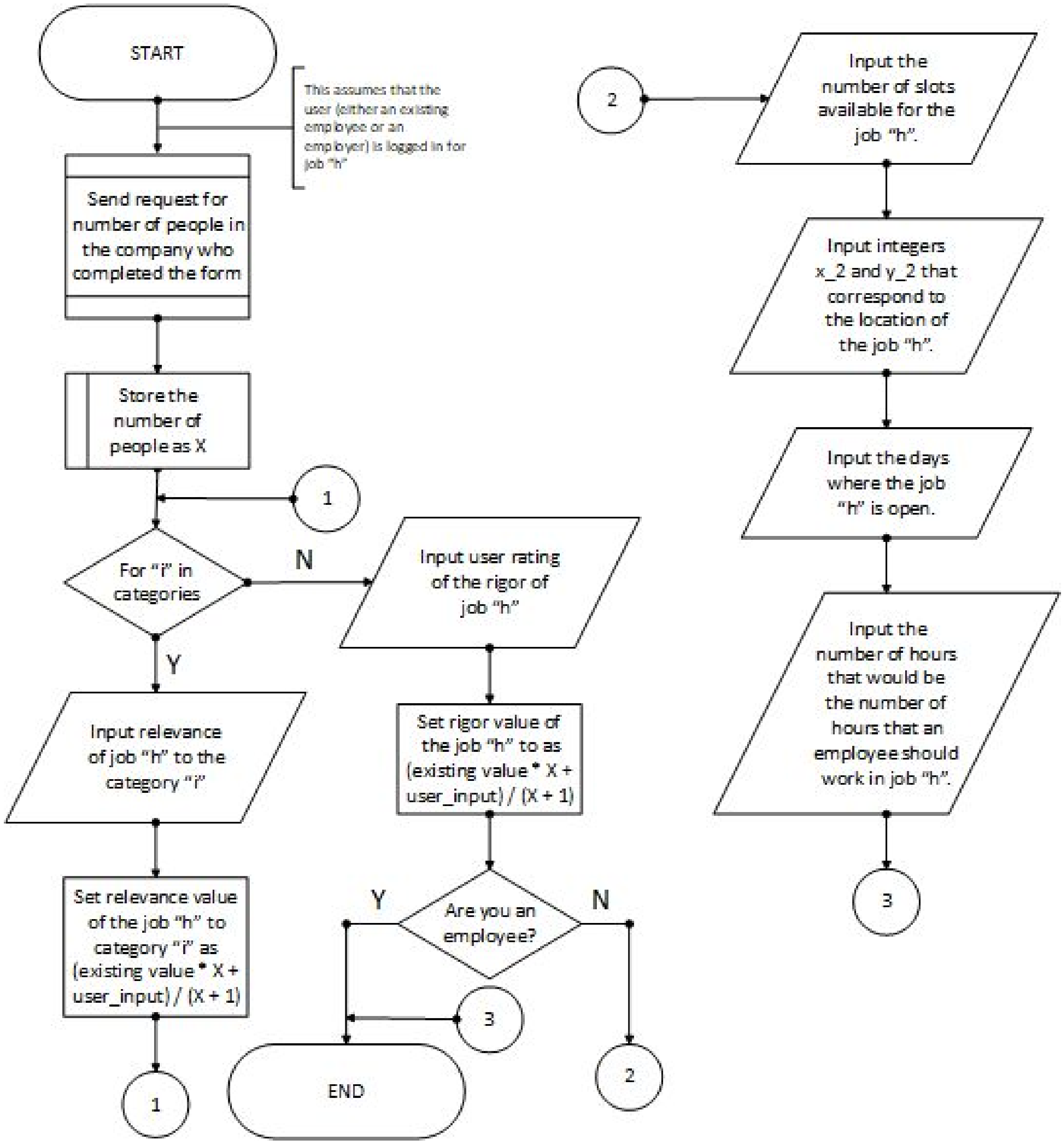


Figure 5. Employee-End Value Updating Design Algorithm.

Source: (A member of Team #11394)

Other Value-Updating Designs

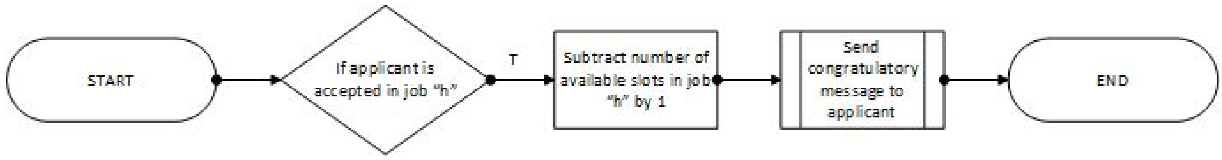


Figure 6. Other Value-Updating Designs

Source: (A member of Team #11394)

Testing and Results

We, the authors, follow the prescribed method provided by the problem authors. We will be testing the results on 10 fictional individuals (we will label them Person-001 to Person-010). We will also be testing the data on 20 fictional jobs (we will label them Job-001 to Job-020). Both lines of data have been made using reasonable reasons, as stated in the problem.

Method

The testing would involve 10 fictional individuals whose answers are fictional and randomly generated by this modeling team members. The answers will describe these people that they answer in the "User-End Section" of this paper. Similarly, the 20 fictional jobs are defined by the answers in the section "Employee-End Section" of this paper. We derived our job categories from Recruiter (2019). In order to analyze the table, the legend is in Appendix A.

Person Database

				job_preference				work_days							
person_id	**loc_x**	**loc_y**	**transpo_mode**	*pref_1*	*pref_2*	*pref_3*	**rigor**	*S*	*M*	*T*	*W*	*T*	*F*	*S*	**hours_work**
Person-001	6	9	bike	info_tech	marketing	arts_atc	1.1	0	1	1	1	1	1	0	35
Person-002	12	13	vehicle	marketing	business_ma	finance	2.3	1	1	0	1	0	1	1	30
Person-003	26	4	train	archi_c	stem	edu_train	2.2	0	1	1	0	1	0	1	32
Person-004	17	19	train	arts_atc	archi_c	agri_fnr	3.4	0	1	0	1	0	1	0	30
Person-005	29	21	bike	stem	info_tech	edu_train	1.8	0	1	1	1	1	0	1	40
Person-006	8	27	walk	hosp_tour	transpo	marketing	4.1	1	1	1	1	1	0	1	36
Person-007	2	15	vehicle	info_tech	finance	business_ma	1.2	1	0	1	0	1	0	1	32
Person-008	23	25	walk	arts_atc	archi_c	edu_train	2.4	0	1	1	0	1	1	1	30
Person-009	10	24	bike	manufac	hosp_tour	transpo	4.8	0	1	1	1	1	1	0	35
Person-010	11	14	train	transpo	agri_fnr	info_tech	3.9	0	1	1	1	1	1	1	42

Table 1. Database of the people for testing in the problem.

Job Database

		relevance																open_days							
job_id	slots	agri_fnr	archi_c	arts_atc	business_ma	edu_train	finance	hosp_tour	info_tech	manufac	marketing	stem	transpo	loc_x	loc_y	salary_h	rigor	S	M	T	W	T	F	S	hours_work
Job-001	1	1	1.1	1	1.8	1	5	5	2.7	1	1	1	1	10	19	13	4.3	0	1	1	1	1	1	0	45
Job-002	0	2.5	1.4	5	1.6	1.1	1.8	2.2	4.3	1	1.5	1	1	5	10	14	2.7	1	1	1	1	1	1	0	40
Job-003	3	1	1.7	1.4	1.9	1.8	1.2	2.4	1	5	2.1	1	3	29	4	10	1.3	0	1	1	1	1	1	1	30
Job-004	5	1	3.2	2.6	4.3	3.3	3.7	5	3.1	1.8	4.9	2.4	1	15	19	12	1.9	0	1	1	1	1	1	0	30
Job-005	3	1	1.4	1	3.5	5	2.2	1	5	4.5	1.4	2.5	1	23	14	10	1.9	1	1	1	1	1	1	0	45
Job-006	2	5	1	1	1.5	1	1	1	1	3.5	1	1.5	1	28	4	13	4.8	0	1	1	0	1	1	1	40
Job-007	2	1	4.7	5	3.1	3.9	2.8	4.1	3.2	1.3	4.3	3.5	1	16	6	10	1.9	1	1	1	1	1	1	1	30
Job-008	2	1	3.9	2.7	3.6	4.7	3.2	2.9	4.9	1.1	2.5	5	1	25	5	15	1.9	1	1	1	0	0	1	1	35
Job-009	4	1.8	3.2	2.1	2.4	3.8	2.9	1.2	5	1	2.7	4.5	1	27	20	14	4.2	0	1	1	1	1	1	0	30
Job-010	3	1	1	1	1	5	1	1	1.5	1	1	5	1	5	16	10	1.7	0	1	1	1	1	1	0	35
Job-011	2	1	2.2	5	1	2.1	1	1	1	1	1	1	1	29	4	18	1.8	1	1	0	0	0	1	1	30
Job-012	4	1	1	1	4.2	1.2	2.8	5	1	1	5	1.9	4	6	20	8	3.8	0	1	1	1	1	1	1	45
Job-013	3	1	1.1	1.1	1	1	1	4.8	3.2	4.3	1	4.2	5	1	5	11	3.5	1	1	1	1	1	1	1	45
Job-014	2	5	1	1	1	1	1	1	1	1	1	1	1	25	3	13	4.9	0	1	1	1	1	1	1	45
Job-015	4	1	3.6	1.1	2.4	1.2	2.3	4.1	1.7	3.8	3.2	2.7	5	28	8	17	2.7	0	1	1	1	1	1	0	40
Job-016	0	3.4	4.3	1	1.2	1.7	4	1.2	1	2.6	1	3.5	1.5	15	20	14	1.6	1	1	1	0	1	1	0	45
Job-017	4	1	1	4.5	3.4	5	1	1	1	1	1	1	1	8	24	18	1.2	1	1	1	1	1	1	0	35
Job-018	0	1.9	1	1	1	1.9	1.6	1.2	1	4.9	1.2	2.7	2.7	14	11	14	4.3	0	1	1	1	1	1	0	45
Job-019	5	1	5	1	1.5	1	1	1	1	1	1	1	1	10	19	12	2.1	0	1	1	1	1	1	0	40
Job-020	2	1	2.2	3.1	5	3.9	5	2.4	3.4	3.1	4.6	3.8	1	16	20	9	2.8	0	1	1	1	1	1	1	45

Table 2. Database of the jobs for testing in the problem.

Results

The results have been ranked according to the ranking algorithm presented in an earlier section in this paper. For the fare, we assumed that running a vehicle is at \$0.76/km (which was derived from personal records) and that the cost of running a train is at \$0.23/km (which was also derived from personal records). For example, for Person-001, the ranking of the person is as below.

job_id	remaining slots	relevance
Job-001	1	11.1
Job-002	0	20.9
Job-003	3	8.6
Job-004	5	21.7
Job-005	3	18.8
Job-006	2	6
Job-007	2	23.2
Job-008	2	22.4
Job-009	4	22.5
Job-010	3	7.5
Job-011	2	10
Job-012	4	14
Job-013	3	12.7
Job-014	2	6
Job-015	4	12.6
Job-016	0	6
Job-017	4	9.5
Job-018	0	6.4
Job-019	5	6
Job-020	2	22.5

Table 3. Ranking of Person-001, with relevance referring to the Q_1 score

Therefore, we conclude that by the algorithm, the best jobs for the people were as follows. We authors, considering ourselves in these people's feet, would believe that this would necessitate our needs.

Person ID	Job ID of the Best Job
Person-001	Job-007
Person-002	Job-020
Person-003	Job-008
Person-004	Job-007
Person-005	Job-008
Person-006	Job-015
Person-007	Job-020
Person-008	Job-011
Person-009	Job-013
Person-010	Job-013

Table 4. Best jobs according to the ranking algorithm for the people tested

Strengths and Limitations

Strengths. Our model's biggest strength is its adaptability to larger datasets (either by the number of people who are using the application or by the number of jobs available to people on-the-spot) in the application. If there are more categories that the jobs can fall under, the model is highly flexible and can still meet these people's demands upon selection. The model is also highly adaptable to other data types (such as decimal values for distance and variable fuel economy, salaries, and hours for workers). This can also be flexible for non-constant traffic time and fuel economy, This is also easily implementable in libraries such as Python "Pandas", the JavaScript D3.js and Aperture, and the C++ GSL Statistics.

Even though the model is minimal and can have little to no contention for scores Q_2 onwards, the model is highly efficient in breaking contentions for larger datasets in more jobs, which can be demonstrated in larger cities or places looking for more jobs. The ranking algorithm also meets the economic and psychological demands that cause the decrease of people taking up summer jobs, providing more incentives to people through prioritizing interests and salary.

Limitations. Our model does not factor in any variable costs such as fuel economy and train fare (which we have set to a constant value for the sake of demonstration purposes and are very small to account for in this problem). These are because the fuel economy for each value remains different between different car models, terrain, and other factors that affect the performance of the vehicle. However, due to phenomena like the city’s demand and the flatness of terrain in a city, these things could be negligible.

Our model also does not factor in any other factors that could affect a person’s performance and acceptance into the job. These are factors like the sex of the person among others. However, only the person looking for the job is the only one who can rate their performance and aptitude in the job; therefore, these could be ignored.

Our model also does not take into account a most efficient method in determining the interests of the people on the ground (i.e. those that are using the application). Even though there are more advanced searching algorithms for relevance (e.g. the relevance categorization of the Google search engine), this ranking algorithm is already decent enough for a small application. All other advanced relevance searches can be discussed in a more advanced model.

Conclusion

We believe that our model is reliably accurate, justified in its assumptions, and flexible to data from the real world. It can provide a ranking algorithm that gives a more comfortable choice for people using the application and provides the person’s economic, physiological, and psychological demands of the person using the application. We believe that this model can provide incentives that would counter the decreasing employment trend in younger generations. Finally, we believe that this model would be more important to keep up the trend of employment in young generations, which provides a better incentive for people to be employed.

References

De Clercq, D., Honig, B., & Martin, B. (2012). The roles of learning orientation and passion for work in the formation of entrepreneurial intention. *International Small Business Journal: Researching Entrepreneurship*, *31*(6), 652–676. https://doi.org/10.1177/0266242611432360

Hamm, T. (2018, September 7). *Lessons from the Average American's Food Expenses*. The Simple Dollar. Retrieved 2020, November 10, from https://www.thesimpledollar.com/save-money/lessons-from-the-average-americans-food-expense/

Iannone, G., Guarnaccia, C., & Quartieri, J. (2013). Speed Distribution Influence in Road Traffic Noise Prediction. *Environmental Engineering and Management Journal, 12*(3), 493–501. https://doi.org/10.30638/eemj.2013.061

Khayitov, A. & Giyosova, G. (2020). *Increasing Population Employment Under Conditions of Modernization Country Economy*. Архив Научных Исследований. https://journal.tsue.uz/index.php/archive/article/view/1349

Organization for Economic Co-operation and Development. (2000, December 1). *Total Unfilled Job Vacancies for the United States*. FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/LMJVTTUVUSM647S

Pew Research Center. (2019, June 25). *Share of working U.S. teens increases in summer, but less so than in past decades*. Retrieved 2020, November 7, from https://www.pewresearch.org/fact-tank/2019/06/27/teen-summer-jobs-in-us/ft_19-06-20_teensummerjobs_2/

Recruiter. (2019). *List of Careers and Occupations. Find Local Jobs in 40,000 Careers*. Recruiter. https://www.recruiter.com/careers/

Appendix A

Legend for the Databases

The two databases use almost similar legends. For the person database, the legend is as follows:

- ***person_id*** refers to the ID of the person that one is taking,
- ***loc_x*** refers to the x-coordinate of the location of the person,
- ***loc_y*** refers to the y-coordinate of the location of the person,
- ***transpo_mode*** refers to the transportation mode the person,
- ***job_preference*** refers to the preferred category of the person taking the job; where ***pref_n*** refers to the n^{th} choice of the person,
- ***rigor*** refers to the rigor of the job that they want to take up,
- ***work_days*** refer to their availability on a certain day, which corresponds to 1 if they are available and 0 if they are not available on that day,
- ***hours_work*** refers to the number of hours that they are willing to work.

For the job database, the legend is as follows:

- ***job_id*** refers to the ID of the job that one is taking,
- ***slots*** refers to the number of slots remaining in the job,
- ***relevance*** refers to the relevance of the job to a certain category, where agri_fnr refers to Agriculture, Food, and Natural Resources, archi_c refers to Architecture and Construction, business_ma refers to Business Management and Administration, edu_train refers to Education and Training, finance refers to Finance, hosp_tour refers to Hospital and Tourism, info_tech refers to Information Technology, manufac refers to Manufacturing, marketing refers to Marketing, stem refers to Science, Technology, Engineering and Mathematics, and transpo refers to Transportation,
- ***loc_x*** refers to the x-coordinate of the location of the job,
- ***loc_y*** refers to the y-coordinate of the location of the job,
- ***salary_h*** refers to the salary of the job per hour,

- ***rigor*** refers to the rigor of the job that they want to take up according to the rigor,
- ***open_days*** refers to the days that the job is opening,
- ***hours_work*** refers to the number of hours that they are working.