Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

COMBINATION OF GALE-SHAPLEY AND PESA-II ALGORITHM IN STUDENT-ACCOMMODATION MATCH

TRINH BAO NGOC[1], DANG NHAT QUANG[2], DANG TIEN DAT[3], LE ANH PHAN[4], NGUYEN NGOC PHUONG KHANH[5], HA THI THANH THAO[6], HOANG PHUONG THAO[7], LE THI CHUNG[8]

[1,2,3,4,6] Department of Information System, Hanoi University, Vietnam
[5, 7,8] Department of Computer Applications, Hanoi University, Vietnam
Email: [1]ngoctb@hanu.edu.vn, [2]2001040173@s.hanu.edu.vn, [3]2001040023@s.hanu.edu.vn, [4]1901040155@ms.hanu.edu.vn, [5]2207090033@ms.hanu.edu.vn, 61801040206@s.hanu.edu.vn, [7]thaohp@hanu.edu.vn, [8]lethichunghtc@gmail.com

ABSTRACT

The inability of new students to secure suitable accommodation not only diminishes their quality of life and academic performance but also under- mines the reputation of the university. This paper tackles the problem by using Stable matching theory, a mathematical framework facilitating mutually beneficial relationships over time. By combining the Gale-Shapley algorithm and the PESA-II algorithm which is a multi-objective evolutionary optimization method, our approach systematically evaluates each student's requirements for accommodation, checking based on the recommendation list provided by the university for each student, striving to create a stable and fulfilling match. This guarantees a fair result that matches the preferences of both students and accommodations, with two main outcomes: students gain access to optimal accommodations, meeting all requirements, fostering an improved academic environment.

Keywords: *Student Selection, Boarding House Arrangement, Stable Matching Theory, Gale-Shapley Algorithm, PESA-II Algorithm.*

1. INTRODUCTION

Finding an appropriate accommodation is a crucial aspect of student life, particularly for those attending universities away from their hometowns. Therefore, the mismatch between students and their chosen accommodation can lead to a variety of issues, including decreased academic performance, waste of money, and reduced overall satisfaction with the university experience [1][[2] has compiled that almost 91% of students in Sydney, Australia either rent accommodation privately or reside with their families. In contrast, 5% opt for commercial accommodations, while 4.1% choose university-provided housing *Figure 1*, highlighting the importance of developing methods for assigning students to accommodations.

Student accommodation refers to multiply-occupied housing that students can choose [1], including University dormitories, Purpose-built student accommodation (PBSA), Student houses in multiple occupation (HMOs). The problem of student-accommodation matching involves two distinct but interconnected processes: Student Selection and Accommodation Arrangement. On the student side, individuals must navigate a myriad of factors such as location, cost, amenities, security, privacy, and roommate attributes [3]. Conversely, universities must optimize their accommodation offerings to attract and retain students while ensuring a diverse and inclusive community [4]. Current

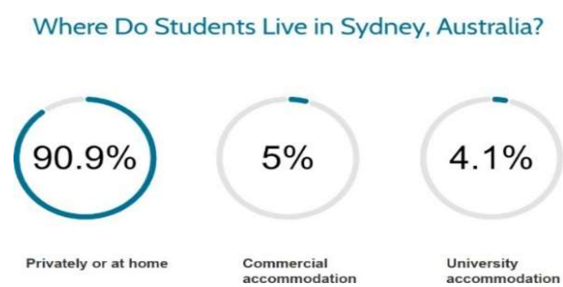

Figure 1: Student Accommodation In Sydney [2]

studies are limited by these optimization factors, and a specific solution to solve the problem automatically by advanced algorithms has not been proposed. We will continue to analyze the problems in current studies in the following section later.

To address matching problems, our paper utilized the Stable matching theory (SMT), the Gale-Shapley algorithm (GSA) [6], and the PESA-II algorithm [7]. SMT provides a mathematical framework for modeling mutually beneficial relationships, often applied to the stable matching problem through the GSA. The GSA, introduced by Gale and Shapley in 1962, has been utilized to address matching problems across various domains, ranging from housing markets to college admissions. It has been proved that the algorithm provides matching not only stable but also optimal in matching problems [5]. In addition to the GSA, we have integrated the Pareto Envelope-based Selection Algorithm II (PESA-II), which tackles multi-objective challenges by identifying Pareto optimal solutions [8]. And PESA-II, renowned for its grid-based fitness assignment mechanisms, offers a robust solution that balances competing criteria [9]. To guide our research, we posed three key questions:

1. What criteria do students- accommodations consider in the matching process?

2. How to find the most suitable pairs of student-accommodation?

3. How to optimize a pairing quality in the context of matching students and accommodations?

This solution takes advantage of one of the most famous theories today (SMT) that won the Nobel Prize, and flexibly combines it with useful technical tools to solve the problem (GSA and PESA-II), creating a novel solution named Stable Matching for Student Selection and Accommodation Arrangement (SMSSAA), which will be explained more in the upcoming sections.

2. RELATED WORK

Throughout the years, researchers have delved into different aspects of this subject, tackling both theoretical concepts and practical applications. [3] proposed a mechanism along with two solving algorithms to match current tenants with new housing, but there was no practical solution to find out an optimal point. [10] successfully assigned Technion university dormitories to students, considering many interesting factor, this paper lacks a mathematical model and a recommendation system to assist both parties in making decisions. Harris [11] proposed a good mechanism for allocating on-campus housing to students, dividing them into two groups: those with a preference for a specific room and those preferring a particular roommate over a room but the characteristics of both parties was a bit simple.

Bristi [12] tackled the problem of both homeowners and tenants by proposing a system that fulfilled all their desires, facilitating the development of society, economy and education in developing countries. Wang [13] provided insights into matching babysitters with suitable households, assessing based on the details of babysitters and household information. [10] addressed the match between student and dormitories based on the socio-economic background, academic performance, as well as their roommates and room conditions. The study of Gupta developed an application pairing roommates by evaluating the compatibility in personality and lifestyle preferences, but there are many other important aspects need to take into account [14]. Then Bornare [15] integrated multiple algorithms to develop Troommate, a solution for matching ideal roommates based on criteria such as social traits, dietary habits, and sleeping schedules, aiming to find the perfect roommate. [16] introduced a mechanism system which ensures the compatibility and fairness in allocating students to dorms, incorporating the demand priority and college ranking of students. Huang [17] allocated healthcare facilities to the elderly population residing in residential neighbor- hoods in Hangzhou, China's main urban districts, prioritizing proximity based on dis- tance. In short, the relevant paper along with their prior contributions to the topic will be listed in the $Table$ 1.

Table 1. Summary Of Relevant Literature.

Year	Author	Problem	Methods
1999	Abdulkadiroğlu et al	House Allocation with Existing Tenants	Top trading cycles mechanisms
2008	Perach et al.	Assigning dormitories to students	Modified version of Gale-Shapley
2018	Harris et al.	Allocating on-campus housing to students	Modified version of SDM
2019	Bristi et al.	Matching between House Owner and Tenant	Gale-Shapley algorithm
2022	Wang	Matching babysitters to households	Gale-Shapley algorithm
2022	Perach and Anily	Matching student-groups to dormitories	Stable matching model, Gale-Shapley algorithm
2022	Gupta et al.	Matching roommates based on personality	K-Means Clustering algorithm
2023	Bornare et al.	Finding a perfect roommate	Gale-Shapley, Elo rating
2024	Afacan	Dorm allocation	Modified version of Gale-Shapley
2024	Huang et al.	Matching healthcare facilities to elderly population	Gale-Shapley algorithm

The majority of relevant literature provides extensive solutions for suitable dorms, houses, and roommates, mainly centered on applying Stable matching theory and the Gale-Shapley algorithm. However, a few lingering issues remain unresolved: (i) needed factors that effect the quality or matching process for student-dorm or student-roommate has not been fully defined; (ii) solutions for matching problems largely rely on the Gale-Shapley algorithm – an old fashion technique.

To address this gap, our paper introduces a novel approach, Stable matching for student selection and accommodation arrangement in university, applying the Gale-Shapley algorithm (GSA) which received the Nobel Prize in Economic Sciences in 2012 [6]. We broaden the scope of matching to include student-accommodation pairs, considering additional factors to facilitate students in finding suitable accommodations.

3. PROBLEM DEFINITION

In SMSSAA, students and accommodations are two kind of players who are separated into restricted sets S and A, respectively. The set of limited students is $S = \{s_1, . . . , s_P\}$ where s_1 is the first specific student, $P \in \mathbb{N}$ represents the total number of students registering for accommodation. Similarly, the set of limited accommodations is $A = \{a_1, . . . , a_Q\}$ where a_1 is the first specific university, $Q \in \mathbb{N}$ stands for the total number of selected accommodations recommended by the university. Every accommodation $a \in A$ has its preference list over the students, and each student $s \in S$ also has its own preference list over the accommodations. The preference list of student m, denoted as PS_m , is defined as $PS_m = [ps_1, . . , ps_{P_m}]$ where P_m is the number of accommodations that student m wants to assess ($0 \leq P_m \leq P$), ps_1 represents the most preferred accommodation for student m, ps_{P_m} represents the least preferred. The preference list of accommodation n, denoted as PA_n, is defined as $PA_n = [pa_1, . . , pa_{Q_n}]$ where Q_n is the number of students that accommodation n wants to select ($0 \leq Q_n \leq Q$), pa_1 rep- resents the most preferred student for accommodation n, pa_{Q_n} represents the least preferred.

Table 2 contains a dataset comprising information about five students {Bob, Alice, Mike, John, Tuan} and their corresponding properties.

Table 2. Examples Of Students' Properties About The Problem To Be Solved

Student	Income	Academic performance	International	Length of stay
Bob	122.2	1	0	3.5
Alice	2113.3	5	0	6
Mike	4334.4	2	0	12
John	54.5	3	0	8
z	73.5	5	1	6.5

Meanwhile, *Table* 3 represents the two accommodations being recruited {Heavan, Citizen} and the corresponding accommodation properties.

Table 3. Examples Of Accommodation Properties About The Problem To Be Solved

Accommodation	Renting price	Private facilities	Air Conditioner	Many roomates
Heavan	200.5	0	1	0
Citizen	10.15	1	1	1
Dorm	10.155	1	1	1

After thorough investigation and research, in this paper, we have determined the accommodation properties such as *location*, *amenity*, *security*, *privacy*, *quantity of room- mates*, *roommate attributes*, *renting price* based on studies conducted by [18][19][20][21], to serve as the requirement for students in selecting suit- able accommodation. Moreover, authors like Devlin in [22] or [23] have defined properties such as *gender*, *income*, *academic performance*, *length of stay*, and *international status* as essential requirements for accommodations when selecting suit- able students.

4. MODEL

Students will be asked to rank the importance of various accommodation selection [25] variables in SMSSAA. Next, we match students with the best accommodations based on this rating. We divided the table in $Table$ 4 below so that we could precisely determine how important each group is.

Table 4. Accommodation Factors

Factor		Formula
Renting price	Renting price a	$T1 = a.wt$
Location	Proximity to university l_1 Proximity to public transport l_2 Proximity to shopping l_3	$T2 = \sqrt{\sum_{g=1}^{3} lg.w}$
Amenity	Free internet access f_1 Hot water in bathroom f_2 Air conditional f_3	$T3 = 2.\sum_{h=1}^{3} fh.wt$
Security	Exterior doors are locked at night s_1 Cameras monitor common areas $s2$ Control of those entering/existing $s3$	$T4 = \sum_{k=1}^{3} \ln(sk).wt$
Privacy	Private facilities p_1 Many roommates p_2	$T5 = \frac{(P1 = P2).w1}{2}$ m
Roommate attributes	Academic performance r_1 Daily study time r_2	$T6 = r1.r2.wt$

Where:
$T_{1,2,3,4,5,6}$ corresponding to renting price, location, amenity and other factors which was mentioned in $Table$ 4; w_t is the weight index of the corresponding factors of a student.

After that, we discover a evaluation formula, $Equation$ 1, to determine the preference score of students m towards accommodation n:

$$PS_m(a_n) = \sum_{i=1}^{6} T_i \quad (1)$$

In a similar vein, we also developed the subsequent formula in the following $Table$ 5 to establish the order of importance of accommodation when choosing students:

Table 5. Each Student's Factor Is Important To The Accommodation

Factor	Formula
Income c	$S_1 = \sqrt{c}.w_v$
Academic performance d	$S_2 = 2d.w_v$
International n	$S_3 = n^e.w_v$
Length of stay u	$S_4 = 5u.w_v$

Where:
$S_{1,2,3,4}$ corresponding to income, academic performance, international, length of stay;
w_v is the weight index of the corresponding factors towards accommodation.

From there, we find a evaluation formula to calculate the preference score of accommodation n towards students m as $Equation$ 2 follows:

$$PA_n(s_m) = \sum_{j=1}^{4} S_j \quad (2)$$

From the two evaluation functions of both agent, $Equation$ 3 is the pairing pay-off function is formed to illustrate the quality of a pair between student m and accommodation n:

$$P(s_m, a_n) = |W^{PS}.PS_m(a_n) + W^{PS}.PA_n(s_m)| \quad (3)$$

Where:

$PS_m(a_n)$ is the preference score of student m toward accommodation n, ($Equation$ 1); $PA_n(s_m)$ is the preference score of accommodation n toward student m; W^{PS} is the weight coefficient of preference score PS of student m; W^{PS} is the weight coefficient of preference score PS of accommodation n.

5. ALGORITHM APPLIED

5.1 PESA-II and Gale-Shapley

SMSSAA is recognized as an example of an NP-

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Hard problem due to the complexity involved in verifying a stable matching in polynomial time $\{O_{SMSSAA} = (x^k) \mid k \in \mathbb{N}^*,\ x = P + Q\}$, where P represents the total number of students and Q represents the total number of accommodations [24]. Particularly, optimizing this problem becomes difficult as the number of students $S = \{s_1, .., s_P\}$ and accommodations $A = \{a_1, .., a_Q\}$ increases exponentially, resulting in heightened computational requirements and exponential time complexity $\{O_{SMSSAA} = (k^x) \mid (k > 1)\}$. To address this challenge, PESA-II is combined with GSA, aiming to explore a large Pareto space (PS) that contains a set of non-dominated solutions $\{PS = \{SL_1, .., SL_W\} \mid W \in \mathbb{N}^*$,

$SL = S \cup A\}$. This paper introduces a model of the Gale-Shapley algorithm for resolving stable matching between students and accommodations (Gale & Shapley, 1962):

$$GSA = \{S, A, PS_m, PA_n, Q_j, MS\} \quad (4)$$

Where:

$S = \{s_1, .., s_P\}$ is a set of students, and P is the number of students ($P \in \mathbb{N}^*$);

$A = \{a_1, .., a_Q\}$ is a set of accommodations, and Q is the number of accommodations ($Q \in \mathbb{N}^*$);

$PS_m = [ps_1, .., ps_{P_m}]$ is a preference list of student m, and P_m is the number of accommodations that the student m want to assess ($0 < P_m \le P$);

$PA_n = [pa_1, .., pa_{Q_n}]$ is a preference list of accommodation n, and Q_n is the number of students that the accommodation n want to assess ($0 < Q_n \le Q$);

SL_n is the number of available slots of the accommodation n;

$p = (s_m, a_n)$ is a pair, matching student m with accommodation n;

$MS = \{p_1, .., p_V\}$ is a successful matching set, where:

$p_i = (s_m, a_n)$ is the i_{th} successful matching pair within the matching set MS;

V is the number of successful matching pair in MS ($0 < V \le P$).

Let student m applies to the accommodation n, where n is the highest ranked accommodation in preference list PS_m of

```
Algorithm 1: Gale-Shapley Algorithm & PESA-II
PESA-II (Initialize Population P = {SL_1, ..,SL_x})
For gen = 1 to MaxGen do
  | # Evaluate fitness value of SL_i by GSA (using Model 4)
  | For each (SL_i ∈ P) do
  |   | GSA (Solution SL_i)
  |   | While Waiting list W != Ø do
  |   |   | # Match student S_m → accommodation A_n
  |   |   | While student S_m is free do
  |   |   |   | if (S_m've proposed to ∀A_n ∈ PA_n) then
  |   |   |   |   | S_m can't match any more, and L = L + S_m;
  |   |   |   | else if (A_n's capacity != full) then
  |   |   |   |   | S_m matches with A_n, and MS = MS + (S_m,A_n);
  |   |   |   | else if (A_n prefers S_g over S_m in PA_n) then
  |   |   |   |   | S_m remains free;
  |   |   |   | else:
  |   |   |   |   | remove (S_g,A_n), and add (S_m,A_n) to MS;
  |   |   |   | end
  |   |   | end
  |   | end
  |   | Return fitness score F(SL_i) (using Equation 8)
  | end
  | Devide P into k Hypercubes H_k = {SL_i | SL_i ∈ P};
  | Selection: Add non-dominated ND solutions SL_i to E
  | Crossover: Transplant two ND solutions SL_i, SL_j ∈ E;
  | Mutation: Mutatate offspring solutions SL_i', SL_j';
  | # Evaluate fitness value of SL_i ∈ E by GSA (using Model 4)
  | For each (SL_i ∈ E) do
  |   | Add non-dominated ND solutions to Archive list AR for the next generation;
  | end
end
Return solution SL_i with highest fitness value in AR
end

* Symbol:
•  Set student: S = {s_1, ..,s_P},
•  Set accommodation: A = (a_1, ..,a_Q),
•  Student preferences: PS_m = (ps_1, ..,ps_{P_m}),
•  Accommodation preferences: PA_n = (pa_1, ..,pa_{Q_n}),
•  Matching set: MS = {{p_1, ..,p_v} | p_i = (s_m,a_n)},
•  Waiting list: W = S = {s_1, ..,s_P},
•  Leftover list: L
•  External population: E
•  Archive list: AR
```

Figure 2: Pseudocode Of Gale-Shapley And PESA-II Algorithm

student m, expressed as $\{a_n = max\,(PS_m)| \, a_n \in A\}$, using $Equation$ 1. If the slot Sl_n of the accommodation n is overapplied (full slot) and the student m is ranked over student g in preference list of accommodation n, represented as ($m \succ_n g$), using $Equation$ 2. Then the student g is deleted from the current match and the student m will match to the accommodation n. For the student g who is rejected becomes unmatched and will keep apply to their next favorite accommodation a_n, ($a_n \in PS_g$). The algorithm concludes when every student has proposed to all the accommodations in their preference list and returns a successful matching set MS. However, it is important to note that there is at least one matching pair $\{\exists p = (s_m, a_n) \mid PS_m \in S, \forall a_n \in PS_m\}$ is unsuccessful, when none of the accommodations an in student's s_m preference list PS_m are willing to accept student an as a match.

GSA, while proficient at matching preferences, is limited to providing only one solution per execution, which could lead to missing the best possible result [26]. To address this constraint, a hybrid approach combining GSA with PESA-II is utilized.

SMSSAA = $\{P,\ SL_i,\ H_k,\ d_i(SL_i),\ p_k(H_k),\ CR,\ MR\}$,(5) Initially, a population $P = \{(SL_1, . . , SL_W) \mid SL_i = (S \cup A), x \in \mathbb{N}^*\}$, consisting W number of solutions (chromosomes) SL_i, each consist of individuals from sets S and A, seeing $Figure$ 3. After matching each individual student from set S to each individual accommodation from set A by using GSA $Model$ 4, successful matching pairs within the solution SL_i are assigned a payoff value to evaluate the quality of each pair, using

$Equation$ 3. In order to select promising solutions, each solution SL_i is selected from k hypercubes $H_k = \{SL_i \in P \mid 1 < i \leq W\}$ based on a fitness value $F(SL_i)$ from $Equation$ 6, crowding distance $d_i(SL_i)$ and hypercube's selection probability $p_k(H_k)$. Selected solutions undergo crossover and mutation at specific rates, known as the cross- over rate CR and mutation rate MR, to generate offspring. Offspring with higher fitness and lower crowding distance are chosen for the next generation. The algorithm iterates to find the satisfactory matching solutions and stops when a specified number of iterations are reached. The integration of PESA-II with GSA is shown in $Figure$ 2 with pseudocode.

5.2 Model of chromosome

Chromosome in a SMT, is represented as an array of individuals from both sets, akin to chromosomes in biology, with each individual (gen) carrying their preferences list toward their partner. For instance, considering 5 students $S = \{s_1, .., s_5\}$

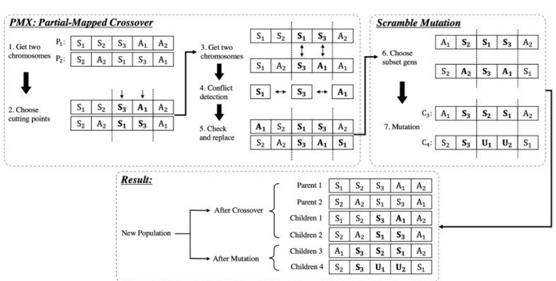

Figure 3: Representation of chromosome

and 3 accommodations $A = \{a_1, .., a_3\}$, the chromosome for stable matching is illustrated in $Figure$ 3:

Where:

The first 5 elements denote the order of students $\{s_1, .., s_5\}$; The second 4 elements denote the order of accommodations $\{a_1, .., a_4\}$; The order of students in set $\{s_1, .., s_5\}$ and accommodation in set $\{a_1, .., a_4\}$ determines their matching priority in STMT; In the chromosome, each individual preference list of S_m and A_n is ex- pressed below:

- S_m preference list: $PS_m = [ps_1, .., ps_{P_m}]$
- A_n preference list: $PA_n = [pa_1, .., pa_{Q_n}]$.

Generally, when P students and Q accommodations participate in the SMSSAA, the population typically contain $(P + Q)^{log(P + Q)}$ chromosomes, each composed of $(P + Q)$ gens. The population size and the number of population are adapted based on the complexity of the input data and the available computational resources.

Fitness function in solving the SMSSAA problem assesses the overall stability score of matching solutions (chromosomes) SL_i between students and accommodation:

$$F = \sum_{k=1} P(s_m, a_n) \qquad (6)$$

Where:

L is the number of successfully matched students ($0 < L \leq P$); $P(s_m, a_n)$ is the payoff function that

calculating the compatibility score of a successful matching pair between student m and accommodation n, referring to $Equation$ 3.

Since $PS_m(a_n) > 0$ $Equation$ 1 and $PA_n(s_m) > 0$ $Equation$ 2, the fitness function F is maximized when the pairing pay-off P is maximized. Therefore, the research problem aims to optimize the value of fitness function F.

Crossover: During crossover, two parent

S_1	S_2	S_3	S_4	S_5	A_1	A_2	A_3	A_4

chromosomes swap segments to create Child 1 and Child 2. A mapping table resolves gene duplicates by replacing them with corresponding genes from the other parent, ensuring genetic uniqueness.

Mutation: In the refining process, genetic mutations occur with a probability of mutation px in each generation. The overall procedure of crossover and mutation is described in the following $Figure$ 4:

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

6. COMPUTATIONAL ANALYSIS

6.1 System and Data Diagnostics

System configuration A web-based application called "Matching theory" has been de- veloped to assess the operational effectiveness of the proposed algorithm in tackling the SMSSAA problem. *Table* 6 below presents system details and key parameters for the algorithms utilized in the experiment of the project.

Figure 4: Flow chart of Crossover and Mutation

Table 6. System And Experiment Information

Experiment parameters	Description
Number of students (P)	2000
Number of accommodations (Q)	20
Number of capacities	Select randomly within [1, P]
Number of characteristics	18
Pairing requirement	Each student can propose to any accommodation Each accommodation can accept students up to its capacity
Edge weight	Assign the characteristics priority in the range of [1, 10]
Preference weigth of student (W^{PS})	Assign in the range of [0, 1]
Preference weigth of accommodation (W^{PS})	Assign in the range of [0, 1]
Algorithm parameters	**Description**
Number of generation	1000
Number of chromosome	100
Crossover rate (CR)	70%
Mutation rate (MR)	5%

The server's system configuration for conducting the experiment includes an Intel® Core™ Ultra 9 processor 185H, 16GB of RAM, 1TB of Enterprise SSD Storage, a GPU NVIDIA Quadro P1000, and a network speed fluctuating around 1Gbps.

Data analysis

The small part of the SMSSAA's 2020 individuals dataset presented in *Table* 7 provides comprehensive information pertaining to the Student and Accommodation sets. However, due to the space constraint, the number of characteristics has been narrowed down from 18 to 6 characteristics. The Accommodation set includes renting price (USD/ month) (a), proximity to university (0 if it's No, 1 if it's Yes) (l_1), proximity to public transport (0 if it's No, 1 if it's Yes) (l_2), and the Student set includes income (USD/ month) (c), academic performance (1 if it's Poor, 2 if it's Average, 3 if it's Good- Average, 4 if it's Good, 5 if it's Excellent) (d), international (0 if it's No, 1 if it's Yes) (n). Each individual will specify their requirements and assign weight levels to their partner's characteristic columns. Additionally, each individual will define their own properties within their characteristic columns. Any undefined cells left will be assigned a value of 0.

Table 7. Small Part Of The 2020 Individuals Dataset.

Characteristics / Players	Accommodation (A)			Student (S)		
	a	l_1	l_2	c	d	n
Student 1						
Requirements	2855.0:6968.6	0	0	0	0	0
Weight	4	2	1	0	0	0
Properties	0	0	0	5680.8	5	0
Student 2						
Requirements	3021.8:9127	0	1	0	0	0
Weight	3	7	2	0	0	0
Properties	0	0	0	5068	3	0
Accom 1						
Requirements	0	0	0	5677	4	1
Weight	0	0	0	8	1	6
Properties	4591.3	0	1	0	0	0
Accom 2						
Requirements	0	0	0	4732	2	1
Weight	0	0	0	8	1	6
Properties	3193.9	0	1	0	0	0

Experiment and Evaluation

Table 8 summarizes the results of subset

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

data from *Table* 2 and *Table* 3 in *Section III*, each row provides the fitness value of the optimization process and the running time of each algorithm. For example, in the first run with PESA-II, the value is represented as "6496/3.2s", where "6496" means the fitness value of the optimized process, and "3.2s" means the completion time of PESA-II is 3.2 seconds.

Table 8. Testing Results Of 8 Individuals Experiment

No.	NSGA-II	NSGA-III	SMS-EMOA	PESA-II	VEGA
1	6496/5.1s	6496/5.2s	6496/5.0s	6496/3.2s	6496/4.0s
2	6496/5.0s	6496/5.1s	6496/5.0s	6496/3.1s	6496/4.2s
3	6496/5.2s	6496/5.5s	6496/5.0s	6496/3.0s	6496/4.3s
4	6496/5.1s	6496/5.4s	6496/5.1s	6496/3.2s	6496/4.2s
5	6496/5.1s	6496/5.3s	6496/5.0s	6496/3.5s	6496/4.1s
6	6496/5.1s	6496/5.2s	6496/5.0s	6496/3.1s	6496/4.5s
7	6496/5.1s	6496/5.1s	6496/5.0s	6496/3.2s	6496/5.0s
8	6496/5.0s	6496/5.3s	6496/5.0s	6496/3.3s	6496/4.1s
9	6496/5.0s	6496/5.3s	6496/5.0s	6496/3.3s	6496/4.5s
10	6496/5.2s	6496/5.3s	6496/5.1s	6496/3.3s	6496/4.2s

The second experiment has a medium dataset with 2000 students and 20 accommodations. The outcomes are generated after 10 times of running and presented as *Table* 9 below.

Table 9. Testing Results Of 2020 Individuals Experiment

No.	NSGA-II	NSGA-III	SMS-EMOA	PESA-II	VEGA
1	76877094/196s	76863947/195s	76858157/331s	76863989/83s	76873688/180s
2	76864036/236s	76876857/195s	76814864/390s	76827856/89s	76869273/179s
3	76859663/195s	76843375/194s	76877382/313s	76867813/91s	77485693/179s
4	77427304/207s	76864189/196s	76871869/284s	76858460/89s	76822702/175s
5	76842998/202s	76818620/195s	76850961/280s	77436183/89s	76860425/173s
6	76864267/198s	76868163/195s	76846930/336s	76837528/90s	76880984/175s
7	76862215/196s	76863759/195s	76822283/265s	76864075/89s	76868828/176s
8	76830151/197s	76846813/194s	76848799/371s	76852667/85s	76877308/172s
9	76867847/195s	76848790/195s	76864312/268s	76877267/85s	76822590/171s
10	76873199/195s	76873141/195s	76818461/267s	77553392/86s	76822241/171s

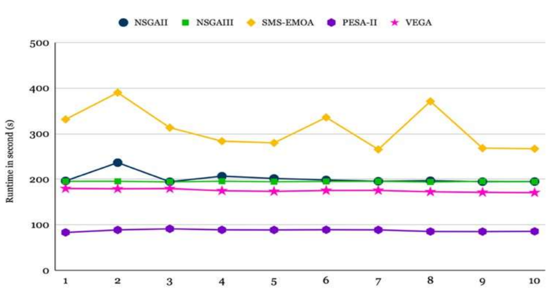

Figure 5: Run-Time Experiment Of 2020 Individuals In Second (S)

Despite minor fluctuations in fitness values among different running times, the highest value in the tenth run of PESA-II with "77553392" showing positive matching results when integrating MOEA frameworks with GSA. Different from the first data bucket' results, the fitness value fluctuates across all of 5 algorithms. Notably, the substantial difference is observed in the running times in the *Figure* 5 below, with PESA- II emerging as the fastest algorithms, running averagely at 87.6 seconds and operating approximately 58% faster than NSGA-II, NSGA-III, VEGA and 80% faster than SMS-EMOA.

Furthermore, a comparison of the fitness scores of 2020 individuals has been depicted in *Figure* 6 below. Specifically, PESA-II achieves the highest score of "77.55 mil- lion" in the tenth run of the experiment, surpassing the average scores of all the other algorithms by a significant margin of "600.000", indicating the notable superiority performance of PESA-II.

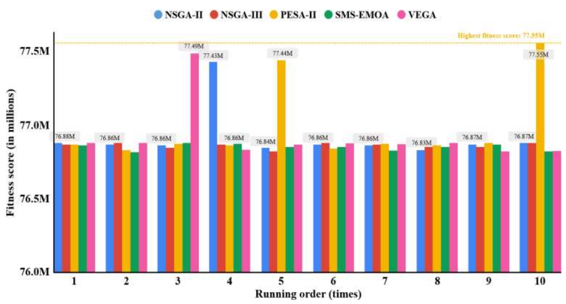

Figure 6: Fitness Scores Of 2020 Individuals Experiment In Million (S)

The *Figure* 7 below presents a small part of the matching result from the experiment involving 2020 individuals. Considering space limitations, it can only illustrate the matching results of the first 20 students and the last student alongside their corresponding matching accommodations. However, the actual results of the experiment will be larger and will be attached in the *Excel* file after finishing the matching processes. From the experiment, it's evident that most students successfully applied to their top favorite accommodations.

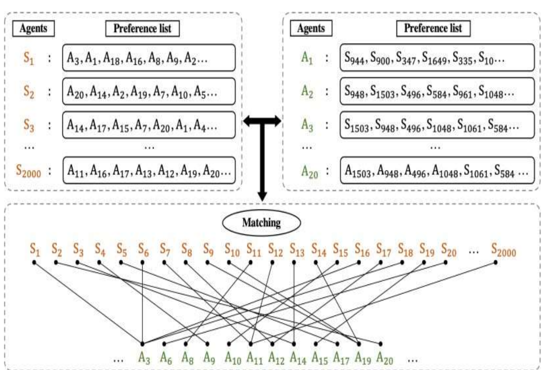

Figure 7: An Excerpt Matching Results Of 2020 Individuals Experiment

In summary, the algorithms have been provisionally assessed to demonstrate their effecttiveness, maintaining consistent speed and yielding reasonably satisfactory results. However, this method only aids in identifying viable options rather than confirming the optimal solution for the problem.

6. CONCLUSION

In this paper, we addressed the challenge of allocating accommodations to students by proposing numerous factors to optimize the pairing process. To tackle this problem, we introduced a novel approach that integrated the Gale-Shapley algorithm and the PESA-II algorithm, deriving an optimal solution from a mathematical model based on SMT - a Nobel Prize-winning theory. This method ensures fairness in outcomes, aligning with the preferences of both students and accommodations. Additionally, our paper highlights the efficiency of the PESA-II algorithm in reducing infrastructure costs, comparing it with other methods such as NSGA-II, NSGA-III, VEGA, and SMS-EMOA. From there, it opens up a promising direction for solving similar problems.

REFERENCES:

[1] Jones, S, Blakey, Student accommodation: The facts. *HEPI Analytical Report 2*, 2020, ISBN: 978-1-908240-70-5, pp. 62

[2] Bouchrika, I. (2024, November 1). Most affordable online colleges.: https://research.com/education/most-affordable-online-colleges

[3] Abdulkadiroğlu, A., Sönmez, T., "House allocation with existing tenants", *Journal of Economic Theory*, 2019, vol 88 (2), pp. 233–260. DOI: 10.1006/jeth.1999.2553

[4] Afacan, M. O. , "Dorm augmented college assignments". Social Choice and Welfare, 1–19.

[5] Klijnsma, W. (2019), *Matching algorithms for the secondary school admission problem in Amsterdam* [Doctoral dissertation, University of Amsterdam].

[6] Gale, D., & Shapley, L. S. (1962). College admissions and the stability of marriage. The American Mathematical Monthly, Vol. 69, No. 1 (Jan., 1962), pp. 9-15.

[7] Li, M., Yang, S., Liu, X., Wang, K., PESA-II: Improved Pareto Envelope-Based Selection Algorithm II. *Lecture Notes in Computer Science*, 2013, vol 7811. Springer. DOI: 10.1007/978-3-642-37140-0_14

[8] Sajjad Shokouhyar, Seed Mansour, Simulation-based optimization of ecological leasing: A step toward extended producer responsibility. *The International Journal of Advanced Manufacturing Technology*, 2012, vol 66(1-4). DOI: 10.1007/s00170-012-4315-9

[9]Fan Zli, Qdeb, KGoodman E., Difficulty Adjustable and Scalable Constrained Multiobjective Test Problem Toolkit. *Evolutionary* Computation, 2020. DOI: 10.1162/evco_a_0025928:3

[10]Perach, N., & Anily, S., Stable matching of student-groups to dormitories. *European Journal of Operational Research*, 2022, vol 302 (1), pp 50–61. DOI: 10.1016/j.ejor.2021.12.048
[11]Harris, W. C., et al. (2018). *Housing problem with multiple preference types*. 2018, [Master thesis, University of NC State]
[12]Bristi, W. R., Chowdhury, F., & Sharmin, S., Stable matching between house owner and tenant for developing countries. *2019 10th International Conference on Computing, Communication and Networking Technologies*, pp 1–6. DOI: 10.1109/ICCCNT45670.2019.8944625
[13]Wang, D., & Li, S.-m., Socio-economic differentials and stated housing preferences in Guangzhou, China. Habitat International, 2006, vol 30 (2), pp 305–326. DOI: 10.1016/j.habitatint.2004.02.009
[14]Gupta, A., Almeida, I., Balaji, H., & Tiwari, M., Dormmate-a room-mate personality matching application. *Second International Conference on CS, Engineering and Applications* , Gunupur, India, 2022, DOI: 10.1109/ICCSEA54677.2022.9936173
[15]Bornare, A., Dubey, Troomate—finding a perfect roommate a literature survey. *International Conference on Data Management*, 2023, Analytics & Innovation, 3–17. DOI: 10.1007/978-981-99-1414-2_1
[16]Afacan, M.O. Dorm augmented college assignments. Soc Choice Welf 62, 609–627 (2024). DOI: 10.1007/s00355-024-01510-9.
[17]Huang, L., Zhang, K., Sun, Coursey, D., Application of Gale-Shapley algorithm in optimal matching for healthcare facilities to elderly population: The case of Hangzhou, China. *Applied Economics*, 2024, pp 1–12. DOI: 10.1080/00036846.2024.2320175
[18]Khozaei, F., Hassan, A., Development and validation of the student accommodation preferences instrument (SAPI). *Journal of Building Appraisal*, 6 (3), 299– 313.
[19]Mullen, B., & Felleman, V., Tripling in the dorms: A meta-analytic integration. *Basic and applied social psychology*, 1992, DOI: 10.1207/s15324834basp1101_3
[20]O'Connell, M., Rosenheck, R., Kasprow, W., & Frisman, L. (2006). An examination of ful filled housing preferences and quality of life among homeless persons with mental illness and/or substance use disorders. *The Journal of Behavioral Health Services & Research*, 33 (3), 354–365.
[21]Oppewal, H., Poria, Y., Ravenscroft, N., & Speller, G. (2017). Student preferences for university accommodation: An application of the stated preference approach. *Housing, space and quality of life*, 113–124.
[22]Devlin, A. S., Gender-role and housing preferences. *Journal of Environmental Psychology*, 1994, vol 14 (3), pp 225–235.
[23]Wang, L. (2022). A potential application of gale-shapley stable marriage model in the babysitter-households matching in China. *Proceedings of the 2021 International Conference on Culture, Design and Social Development,* 2022, DOI: 10.2991/assehr.k.220109.011
[24]Meeks, K., & Rastegari, B., Solving hard stable matching problems involving groups of similar agents. *Theoretical Computer Science*, 2020, vol 844, pp 171–194.
[24]Perach, N., Polak, J., & Rothblum, U. G. (2008). A stable matching model with an entrance cri terion applied to the assignment of students to dormitories at the technion. International Journal of Game Theory, 36, 519–535.
[25]Hall, R. L., & Willerman, B. (1963). The educational influence of dormitory roommates. *Soci ometry*, 294–318.
[26]Sajjad, F., Rashid, M., Zafar, A., Rodrigues, J. J. (2023). An efficient hybrid approach for optimization using simulated annealing and grasshopper algorithm for iot applications. *Discover Internet of Things*, DOI: 10.1007/s43926-023-00036-3