Heuristic based model for groceries shopping navigator

Muhammad Wardi bin Peeyee[1], Shuzlina Abdul-Rahman[2],
Nurzeatul Hamimah Abdul Hamid[2], MohdZaki Zakaria[2]
[1,2]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Malaysia
[2]Research Initiative Group of Intelligent Systems, Malaysia

Article Info

Article history:

Received Jan 11, 2019
Revised Apr 15, 2019
Accepted May 24, 2019

Keywords:

Genetic algorithm
Groceries shopping
Heuristic model
Shopping navigator
Travelling salesman problem

ABSTRACT

This paper presents a heuristic based model for groceries shopping navigator that attempts to improve the navigation problem that usually face by customers while doing their shopping. A system known as Shopping Navigator or shortly SHoNa was developed to give the optimal sequence of shelves to be visited by the customer and the total estimated shopping time so that the user can plan their shopping task earlier. Genetic algorithm was employed and implemented in a web-based platform that is compatible with other devices such as smartphones and tablets. SHoNA can minimize the shopping time by identifying the most optimal order of shelves inside the supermarket that needs to be visited by the customer. A series of experimental was performed in producing the optimum model. Our findings showed that the combination of order one crossover and inverse mutation produced a better optimal performance, which is the minimum total amount of groceries shopping time. SHoNA can be further enhanced with visualization features for a better shopping experience.

Corresponding Author:

Shuzlina Abdul-Rahman,
Faculty of Computer and Mathematical Sciences,
UniversitiTeknologi MARA, Shah Alam, Selangor, Malaysia.
Email: shuzlina@tmsk.uitm.edu.my

1. INTRODUCTION

Grocery shopping is one of the most fundamental activities which usually perform at the supermarket or hypermarket regularly [1]. Supermarket or hypermarket is a self-service store with a lot of products stacked along the aisles such as foods, clothes and pharmacy products [2]. Today, the growth of the supermarket is not only in terms of an increase in its numbers but also its size. Unfortunately, the expansion in the size of the supermarkets has resulted in customers having to walk longer than before. Nevertheless, it is one of the strategies used by supermarkets to increase their profits and sales. Grocery shopping looks trivial, but it can become a challenging task once a person cannot afford the time given their busy schedule. In Malaysia for example, there are thousands of products inside a huge supermarket. Every supermarket has a different shelf layout and the shoppers often find it difficult and time-consuming to find their desired products [3]. It is imperative for shoppers to know the exact location of their products to plan their shopping activity. Sometimes, the shoppers must ask the supermarket staff about the position of the products, which adds more time to the shopping task [4]. Therefore, we need to find an optimal shopping route for such shoppers, which not only can navigate themselves in finding their products inside the supermarket but also can reduce their shopping times.A heuristic technique is an approach used for problem-solving, learning, or discovery that employs a practical method not guaranteed to be optimal or perfect, but sufficient for the immediate goals [5-6]. It is designed to solve a specific problem quickly when the classic methods are too slow [7]. It also designed for finding the approximate solution when traditional methods fail to find any exact solution [8]. Groceries shoppers are having difficulty to find their desired products from the supermarket with the traditional approach since it requires them to visit almost all racks to locate their list of products, and this seems to be time-consuming. Thus, a heuristic based technique seems a good solution to assist those shoppers in being quick to complete their groceries shopping.

This paper presents a heuristic based model for groceries shopping navigator that attempts to improve the navigation problem that usually face by customers while doing their shopping. The findings show that the combination of order one crossover and inverse mutation produced better optimal results, which is the minimum total amount of groceries shopping time. The results gave a good comparison between the changed parameter tuning to determine which setting can yield better results. Section 2 describes about Genetic Algorithm while section 3 presents the Shopping Navigator (SHoNa) development. The results and findings of the experiments are presented in Section 4. Finally, Section 5 concludes the research with the recommendations of the future works.

2. GENETIC ALGORITHM

In solving navigation problems, Cela *et al*. [9] stated that polynominal time algorithm could become a solution and the results can be obtained in a quicker time if the algorithm is combined with time complexity [9]. Mladenović *et al.* [10] claim that the neighbourhood search technique can also solve navigation problem. The technique can improve the overall solution around seven percent using huge instances of data and it is also can reduce the computational time [9]. In addition, Ant Colony Optimization (ACO) hyperheuristic technique is also a better solution in handling navigation problem [11]. According to Aziz [11], ACO solution is better than several other methods such as simulated annealing, annealing genetic, Tabu search, and adaptive Tabu. Beyond that, Roberti *et al*. [12] found that general variable neighbourhood search and dynamic programming can solve 20 customers’ instances in a short amount of computing time. Another interesting algorithm that is commonly used in navigation problem is Genetic Algorithm (GA) and has been applied in many similar applications due to its simplicity and robustness [13-14]. Genetic Algorithm (GA) simulates the natural process of natural evolution, which follows the law of survival of the fittest. It is a method of repeating the genetic operators, which are selection, crossover and mutation, based on the total number of individuals in the population and it will evolve continuously for each individual in a population to generate more excellent population [14-17]. GA is a stochastic search method. It shows high robustness, global optimality, implicit parallelism and adaptability in solving the continuous optimization problems over a long period of time.

Figure 1 shows the pseudocode of a standard GA. It shows the iterative process of GA in finding optimal solution based on a certain fitness function. First, a simple GA process should initialize its population randomly before it can be computed into its fitness formula. After each individual has its own fitness value, some of the individuals will be selected to obtain the parent individuals through selection operator. Next, depending on the crossover rate, the crossover operator is applied to the individuals in the population to get the new generation. Then, according to some certain mutation rate, the mutation operator is performed. The steps from the evaluation process will be repeated until it meets the stopping condition. The possible termination situation is when either the fitness value reaches the predefined value, or the iteration number beats the maximum [18]. GA can successfully solve the supermarket shopping route problem because the problem they try to solve is similar to TSP [3]. Other than that, Hardi [19] showed in his study that the optimal solution for TSP can also be found just in one generation by using GA because of near optimal solution. Another study proposed by Xiao and Chen [20], found that the research where the adaptive GA and can increase the efficiency of searching and was able to reduce the time consumption of the searching. The standard basic GA usually costing more time to find the optimal solution compared to the adaptive GA.

To receive the best results when dealing with multi-objective GA that are dominated by a mono-objective GA, for instance, the single objective GA is needed to be parameterized in a different way which is by fitting to the identified problem [21]. In another perspective, the implementation of GA together with Geographic Information Systems (GIS) in a genetic programming has produced an unexpected output. The results from the integration of GA and GIS are better because DNA distance display the shortest path compared to geographical distance [22]. Other artificial intelligent techniques also can be hybridized with GA such as a hybrid between GA and the variable neighborhood search. Based on the study conducted by Pham and Huynh [23], the combination from those two techniques can perform much better in terms of running time and the quality of solutions in many data instances. They added [23], the performance of GA also can be improved with the application of nearest neighbor tour construction and second optimal mutation.

```
Genetic_Algorithm() {
               Initialize random population;
               Evaluate the population;
               Generation = 0;
     While termination criterion is not
     satisfied {
            Generation = Generation + 1;
            Select good chromosomes by
            reproduction procedure;
            Perform crossover with probability
            crossover (Pc);
            Perform mutation with probability
            of mutation (Pm);
            Evaluate the population;
     }
}
```

Figure 1. Standard pseudocode of genetic algorithm

Generally, the researchers were required to obtain the best combination order of cities and discover every possible solutions and alternatives. The solutions will cost some time, depending on the number of the cities they visit. However, heuristics method such as GA can deal with the time consumption problem because the method involves some rules to select the next city. GA gets the most people's favour and it is one of the most effective methods to solve navigation problem not just because of its wide adaptability, it is also because of its character in which there is no need to gain further insight into the problems and depends less on the specific fields of the problem [24-25]. In this research, the problem is occurs when the shoppers need to visit some shelf to find their desired groceries products in the shortest time possible. We use GA to get the optimal combination order of shelves in order to navigate the shoppers to purchase their groceries products efficiently and faster. Furthermore, GA has a simple calculation method, strong search capability that capable to adapt to the changes of nature, self-organization and parallelism.

3. SHOPPING NAVIGATOR (SHONA) DEVELOPMENT

This section presents the development of SHoNA that includes data acquisition, system architecture and system design. The process begins by collecting the information from the supermarket's grocery shelves, which includes the shelves' name, position and the products available on the shelves. We limit the sample to 33 groceries shelves and record the distance and the time taken between thesegroceries' shelves. Figure 2 shows the plan of the supermarket.

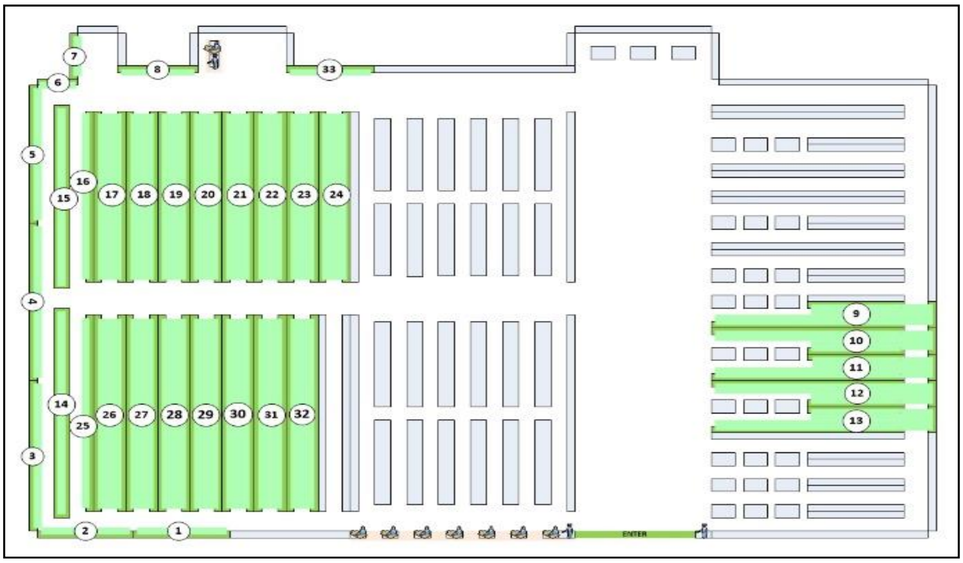


Figure 2. Supermarket plan layout

Figure 3 shows the architecture of SHoNa's system. As can be seen, the system will receive inputs from the user which are the products that the user wants to purchase and the user data. The information will be stored in the database before it can be used by the GA engine. From the database, the engine will retrieve the information about the products, which are the shelves that locates the products, the items that the user wants to purchase and the time between each shelf. After the process is completed, the system will display the results to the user, which are the sequence of shelves that the user needs to visit, and the total minimum of shopping time processed by the engine. For the GA engine, it starts by retrieving some data from the database, which are the list of the products that the user wants to purchase and the time as well as the distance between each shelf. After that, the system initializes the parameters such as the number of candidates in a population, iteration number and the type of genetic operator used. Next, the system generates a random population based on the initialized population number before the system includes the starting point. The starting point is the entrance of the supermarket in which its position is fixed along the process while the gene after the starting point, which is the shelves, will be the subject for the genetic operators later.

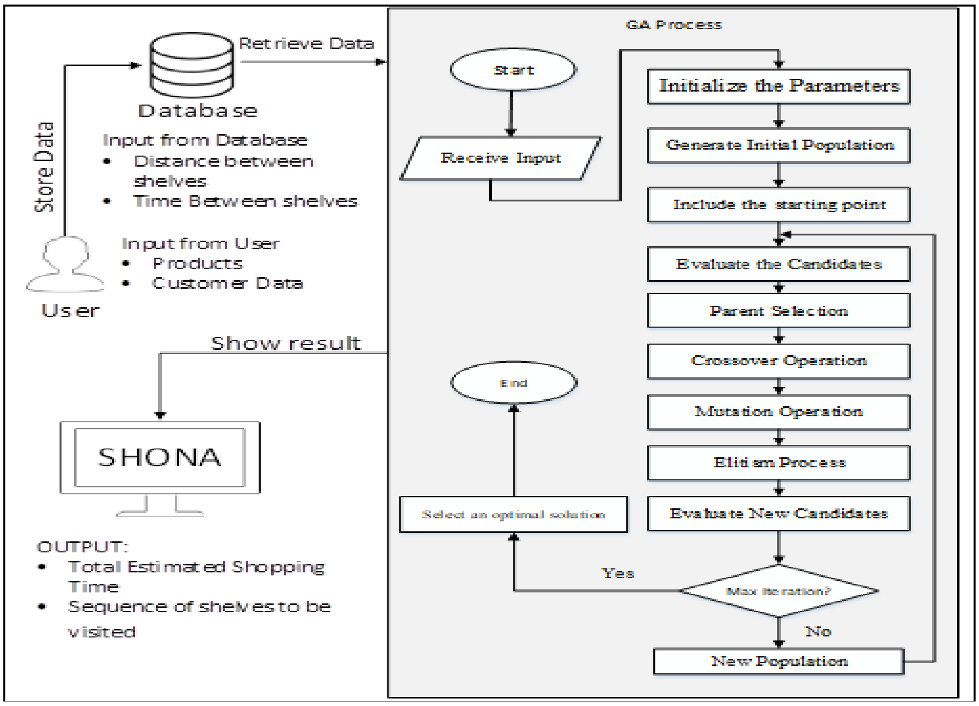


Figure3. System architecture of SHoNa

Figure 4 shows the example of a chromosome representation that represents "shopper" who wants to purchase groceries products from four different shelves. The first gene of the chromosome is the entrance of the supermarket and the rest of the genes are the shelves that the shopper needs to visit. In each gene, there are shelf's label, distance between shelves and time taken between shelves that the shopper needs to visit. The entrance will be fixed along the process while the rest of the genes will be manipulated by the genetic operators.

Equation 1 shows the fitness function of this study that is used to find the total minimal of shopping time. The two best candidates, which are the candidates with the shortest time among the other candidates, will be selected. These two best candidates are called as parent. Then, the selected parent will be applied with crossover and mutation operators. The process will generate two new candidates to proceed with the elitism process. There were two crossover methods and four mutation methods experimented in this study. The crossover methods are Order One Crossover and Cycle Crossover, while the mutation methods are Swap Mutation, Insert Mutation, Inverse Mutation, and Scramble Mutation. Order One Crossover is a simple permutation crossover. Basically, a random swath of consecutive alleles from the first parent drops down, and the remaining values are placed in the child in the order which they appear in the second parent. The Cycle Crossover operator identifies several cycles between the two parent chromosomes. Then, to form child 1, cycle one is copied from parent 1, cycle 2 from parent 2, cycle 3 from parent 1, and so on.

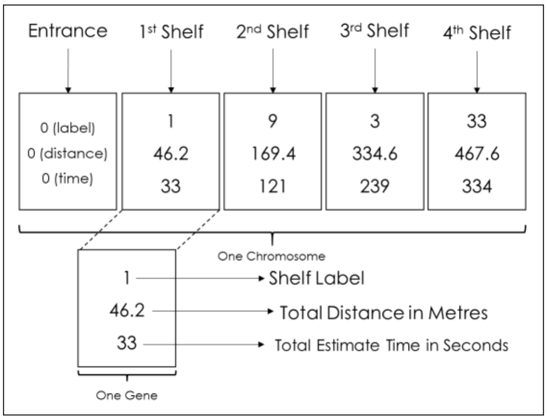


Figure 4. The example of a chromosome representation

$$min\ f(x) = \sum_{i=0}^{n-1} x_i \quad (1)$$

For the swap mutation, this technique picks two points at random while inversion method picks two alleles at random and then inverts the substring between them. For the scramble mutation, it picks a subset of genes at random before it randomly rearranges the alleles in those positions. Two new offspring will be generated after the crossover and mutation process. Next, elitism process will occur where the two worst candidates in the current population will be replaced with the two best candidates among the parent and the offspring. Lastly, the system will evaluate the candidates in the current population to be the final optimal results. If the system meets the stopping condition, which is the maximum number of iterations, the best solution from the last iteration will be the solution or else, the current candidates will repeat the process as the next new population until it reaches the stopping condition. Software such as Sublime Text Version 3 and Mobirise are the tools used in designing the system interface. The interfaces of the system are carefully designed in order to ease the user to interact with the system. A bootstrap template was applied for developing the interface as it has the responsive features and compatible to mobile devices. Figure 5 shows the difference between regular system interface and the responsive interface. The integration process combines the engine and the interfaces, and this process is the last activity in the system development phase. After these two components have been integrated, the system is ready to be tested and validated.

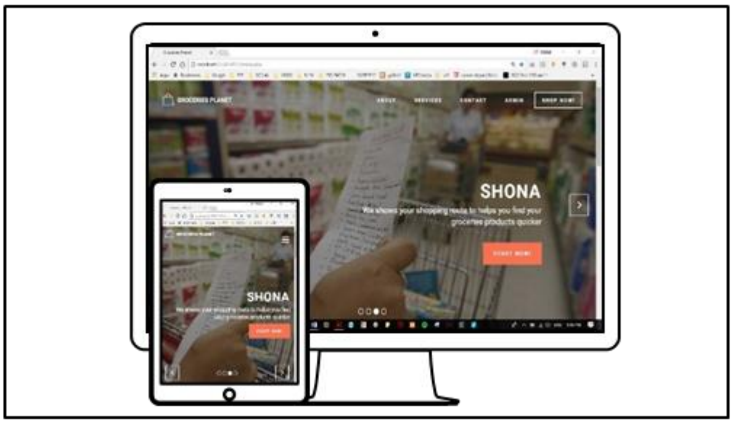


Figure 5. Regular webpage view and responsive webpage view

4. RESULTS AND ANALYSIS

This section presents the results of the experiments in two different cases, Case 1 and Case 2. The goal of this experiment is to find the best combination of genetic operators and to identify the convergence point of the optimal graph. There were two experiments involved in both cases. There were 50 iterations for experiment 1 and 150 iterations for experiment 2 and similar number of populations which is

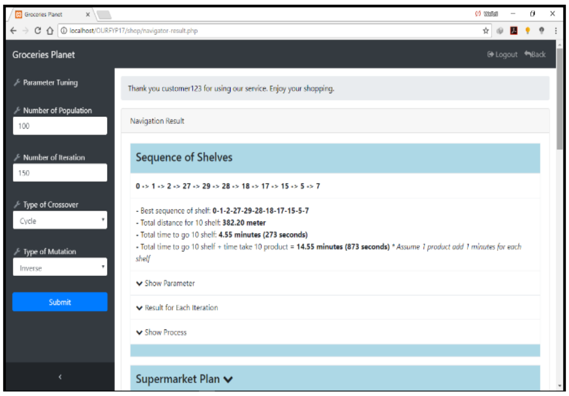
100 were tested in both cases. Next, for each experiment, there were two crossover types (order one and cycle) and four mutation types (swap, insert, inverse and scramble) in which the goal is to find the best combination of genetic operators.

4.1. Case 1

In the first case, SHoNa was tested to find ten random groceries products. For the first experiment, the system was tested with 100 number of population and 50 number of iterations. For the second experiment, the system was tested with 100 number of population and 150 number of iterations. Table 1 summarizes the results of case 1.From the first case, the combination between cycle crossover and inverse mutation was the best combination. The reason is because it produced better results compared to other genetic operators' combinations. Cycle crossover is considered as the best crossover operator in the second experiment since it produced more best results than order one crossover. On the other hand, inverse mutation was considered as the best mutation operator since it produced more best results in both experiments. The result from the best combination is shown in the Figures 6 and 7.

Table 1. Case 1 Results

Experiment	Best Crossover	Best Mutation	Average Shopping Time (min.)
1	Order One	Inverse	15.09
2	Cycle	Inverse	14.58



Number of Population: 100

Number of Iteration: 150

Type of Crossover: Cycle

Type of Mutation: Inverse

Thank you customer123 for using our service. Enjoy your shopping.

Navigation Result

Sequence of Shelves

0 -> 1 -> 2 -> 27 -> 29 -> 28 -> 18 -> 17 -> 15 -> 5 -> 7

- Best sequence of shelf: **0-1-2-27-29-28-18-17-15-5-7**
- Total distance for 10 shelf: **382.20 meter**
- Total time to go 10 shelf: **4.55 minutes (273 seconds)**
- Total time to go 10 shelf + time take 10 product = **14.55 minutes (873 seconds)** * *Assume 1 product add 1 minutes for each shelf*

Figure 6. Case 1 best results

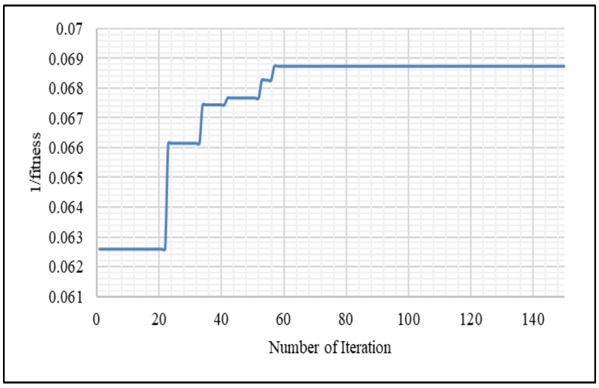


Figure 7. Optimal graph produced from the best results of case 1

From Figure 6, the user was estimated to complete the shopping task in 14.55 minutes. By using this system, user can find their desired products much easier compared to the conventional way of shopping. This is because the user can navigate themselves with the sequence of shelves generated from the system. From Figure 6, it shows that the best sequence of shelves for case 1 is (0-1-2-27-29-28-18-17-15-5-7). The convergence point was found in the 57th iteration as shown in Figure 7.

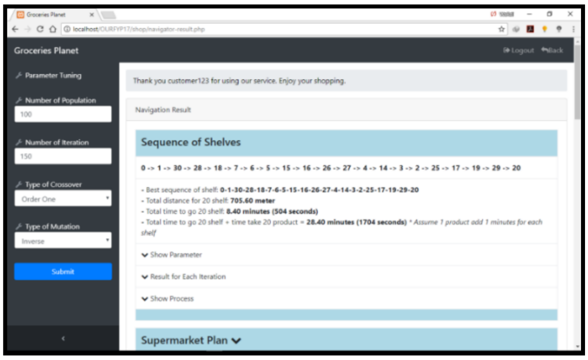
4.2. Case 2

In the second case, **SHoNa** was tested to find twenty random groceries products. For the first experiment, the system was tested with 100 number of population and 50 number of iterations. For the second experiment, the system was tested with 100 number of population and 150 number of iterations. Table 2 summarizes the results of case 2.

Table 2. Case 2 Results

Experiment	Best Crossover	Best Mutation	Average Shopping Time (min.)
1	Order One	Inverse	29.54
2	Order One	Inverse	28.40

From the second case, it was concluded that the combination between order one crossover and inverse mutation was the best combination. The combination had generated the best results for both experiments in Case 2. Its constancy in producing the best results was the key that leads to its best performance. The result from the best combination is shown in the Figures 8 and 9. From Figure 8, the user was estimated to complete the shopping task in 28.40 minutes. By using this system, user can find their desired products much easier com-pared to the conventional way of shopping. This is because the user can navigate themselves with the sequence of shelves generated from the system. Figure 8 shows that, the best sequence of shelves for case 2 is (0-1-30-28-18-7-6-5-15-16-26-27-4-14-3-2-25-17-19-29-20). The convergence point for the second case was found in the 140th iteration as shown in Figure 9. After the system was tested, we were managed to identify the convergence point for each case and we had also found that the best genetic operator's combination was order one crossover and inverse mutation. As we can refer to Tables 1 and 2, from the four experiments that have been done, three optimal results were generated from the order one crossover and all experiments have shown that the inverse mutation produced the best results. By using this system, user can plan their time before they can do the shopping task since the system can estimated an optimum shopping time for the user. This system also can improve the experience of grocery shopping into a better way. The user can navigate themselves in finding their grocery products which saves their time and energy.



Number of Population: 100

Number of Iteration: 150

Type of Crossover: Order One

Type of Mutation: Inverse

Thank you customer123 for using our service. Enjoy your shopping.

Navigation Result

Sequence of Shelves

0 -> 1 -> 30 -> 28 -> 18 -> 7 -> 6 -> 5 -> 15 -> 16 -> 26 -> 27 -> 4 -> 14 -> 3 -> 2 -> 25 -> 17 -> 19 -> 29 -> 20

- Best sequence of shelf: **0-1-30-28-18-7-6-5-15-16-26-27-4-14-3-2-25-17-19-29-20**
- Total distance for 20 shelf: **705.60 meter**
- Total time to go 20 shelf: **8.40 minutes (504 seconds)**
- Total time to go 20 shelf + time take 20 product = **28.40 minutes (1704 seconds)** ** Assume 1 product add 1 minutes for each shelf*

Figure 8. Case 2 best results

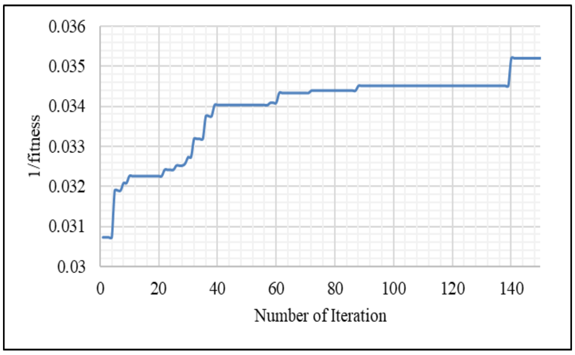


Figure 9. Optimal graph produced from the best results of case 2

5. CONCLUSION AND FUTURE WORKS

This system was designed to help the grocery shoppers in finding their desired grocery products while shopping at the supermarket. A heuristic based system, SHONA, was developed to help the shoppers to navigate themselves inside the supermarket and finish their shopping with a minimal total shopping time. Genetic Algorithm was used in the navigation system for finding the optimal sequence of shelves that the user needs to visit. Thus, it can help the shoppers to do their groceries shopping in a faster and efficient way. The results from our findings show that it is a reliable method for shoppers since it can produce better solutions than the classic shopping method. Our findings showed that the combination of order one crossover and inverse mutation produced a better optimal performance, which is the minimum total amount of groceries shopping time. SHoNA can be further enhanced with visualization features for a better shopping experience.

ACKNOWLEDGMENT

The authors would like to thank the Institute of Quality & Knowledge Advancement (InQKA), Universiti Teknologi MARA, Malaysia for the publication support.

REFERENCES

[1] Bhattacharya, S., *et al*. "*Ma$$iv€ an intelligent mobile grocery assistant*," 2012 Eighth International Conference on Intelligent Environments. IEEE, 2012.

[2] Chamhuri, N., and Batt, P., "Exploring the factors influencing consumers' choice of retail store when purchasing fresh meat in Malaysia," *International Food and Agribusiness Management Review*; 16(3) pp. 99-122, 2013.

[3] Chen, X., Li, Y., and Hu, T., "*Solving the supermarket shopping route planning problem based on genetic algorithm*," 2015 IEEE/ACIS 14th International Conference on Computer and Information Science (ICIS). IEEE, 2015.

[4] Rupanagudi, S. R., *et al*. "*A novel video processing based cost effective smart trolley system for supermarkets using FPGA*," 2015 International Conference on Communication, Information & Computing Technology (ICCICT). IEEE, 2015.

[5] Ippoliti, E. (Ed.), (2014), Heuristic reasoning (Vol. 16), Springer.

[6] Masrom, S., *et al*. "The implementation frameworks of meta-heuristics hybridization with dynamic parameterization," *Journal of Fundamental and Applied Sciences*; 9(6S) pp. 558-576, 2017.

[7] Salhi, S., (2017), Heuristic search: The emerging science of problem solving, Springer.

[8] Polya, G., (2014), How to solve it: A new aspect of mathematical method, Princeton university press.

[9] Cela, E., Deineko, V., and Woeginger, G. J., "The x-and-y-axes travelling salesman problem," *European journal of operational research*; 223(2) pp. 333-345, 2012.

[10] Mladenović, N., Urošević, D. and Ilić, A., "A general variable neighborhood search for the one-commodity pickup-and-delivery travelling salesman problem," *European Journal of Operational Research*; 220(1) pp. 270-285, 2012.

[11] Aziz, Z. Abd., "*Ant colony hyper-heuristics for travelling salesman problem*," Procedia Computer Science; 76 pp. 534-538, 2015.

[12] Roberti, R., and Wen, M., "The electric traveling salesman problem with time windows," *Transportation Research Part E: Logistics and Transportation Review*; 89 pp. 32-52, 2016. doi: https://doi.org/10.1016/j.tre.2016.01.010

[13] Oliver, (2017), Genetic algorithm essentials. (Vol. 679), Springer.

[14] Yu, Y., Chen, Y., and Li, T., "*A new design of genetic algorithm for solving TSP*," 2011 Fourth International Joint Conference on Computational Sciences and Optimization. IEEE, 2011.

[15] Jebari, K. and Madiafi, M., "Selection methods for genetic algorithms," *International Journal of Emerging Sciences*; 3(4) pp. 333-344, 2013.

[16] Robyr, J-L., *et al*., "*Convergence of multi-criteria optimization of a building energetic resources by genetic algorithm*," Proceedings of 2018 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE), 29 May-1 June 2018, Kajang, Malaysia. No. CONFERENCE. 29 May-1 June 2018, 2018.

[17] Yusoff, M., Ikram, M. N. S. B. M., and Janom, N., "Task assignment optimization for crowd sourcing using Genetic Algorithm," *Advanced Science Letters*; 24(11) pp. 8205-8208, 2018.

[18] Zakuan, A. Z., Rahman, S. A., and Jantan, H., "Towards academic successor selection modelling with Genetic Algorithm in multi-criteria problems," *International Journal of Engineering and Technology (UAE)*; 7(4) pp. 130-133, 2018.

[19] Hardi, R., "Genetic algorithm in solving the TSP on these mineral water," *2015 International Seminar on Intelligent Technology and Its Applications (ISITIA). IEEE*, 2015.

[20] Xiao, Z., and Chen, J., "*Research on path optimization based on improved adaptive Genetic Algorithm*," 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, 2015, pp. 207-209.

[21] Alves, R. M. F. and Lopes, C. L., "Using genetic algorithms to minimize the distance and balance the routes for the multiple traveling salesman problem," *2015 IEEE Congress on Evolutionary Computation (CEC). IEEE*, 2015.

[22] Sridhar, R. and S. Balasubramaniam, "GIS integrated DNA computing for solving travelling salesman problem," *2011 IEEE Symposium on Computers & Informatics. IEEE*, 2011.

[23] Pham, D. T. and Huynh, T. T. B., "*An effective combination of genetic algorithms and the variable neighborhood search for solving travelling salesman problem*," 2015 Conference on Technologies and Applications of Artificial Intelligence (TAAI). IEEE, 2015.

[24] Forhad, M. S. A., Hossain, M. S., Rahman, M. O., Rahaman, M. M., Haque, M. M., and Patwary, M. K. H., "An improved fitness function for automated cryptanalysis using Genetic Algorithm," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, 13(2) pp. 643-648, 2019.

[25] Permana, I., Rozanda, N.E., Syafria, F., and Salisah, F. N., "Optimization learning vector quantization using Genetic Algorithm for detection of diabetics," *Indonesian Journal of Electrical Engineering and Computer Science*, 12(3) pp. 1111-1116, 2018.

BIOGRAPHIES OF AUTHORS

Muhammad Wardi bin Peeyee graduated from Universiti Teknologi MARA in 2019 with a bachelor degree in Information Sytems (Intelligent Systems Engineering). He developed an optimzation model for groceries shopping navigator for his degree thesis.

Shuzlina Abdul Rahman received her bachelor’s degree in Computer Science from Universiti Sains Malaysia in 1996, Master of Science in Information Technology from Universiti Utara Malaysia in 2000 and PhD in Science and System Management from Universiti Kebangsaan Malaysia in 2012. She is currently working as Associate Professor at Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, Malaysia. Her research interest includes computational intelligence, machine learning and data analytics & optimization

Nurzeatul Hamimah Abdul Hamid graduated from Universiti Tenaga Nasional with a bachelor degree in Information Technology on 2001. She later obtained her Master of Science in Intelligent Systems from University of Sussex in 2005. She is currently serves as senior lecturer at the Center of Information System Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Malaysia. Currently she is an active researcher in the area of soft computing, computational Intelligence and multi-agent systems.

MohdZaki Zakaria holds a PhD in Electrical, Electronics and Systems Engineering by the National University of Malaysia. He is currently serves as senior lecturer at the Center of Information System Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Malaysia His research and publication interests include optimization, computational intelligence and data analytics.