## CHAPTER 4

## ANALYSIS AND DESIGN

### 4.1 Analysis

In this chapter, the authors analyze the data used in his research so that it can be used in the Greedy Algorithm and Genetic Algorithm. The purpose of this analysis is to find a more optimal algorithm between the Greedy Algorithm and the Genetic Algorithm from the distance away from the calculated 2-dimensional floor plan that has been created as a dataset. The dataset itself is obtained from the Semarang Polrestabes social media and online collection through Google Maps. Table is 50 regions taken as a dataset in this study. The order from the beginning to the end has been determined according to the number.

Table 4.1: Regional Data Table

| No | $\mathbf{X}, \mathbf{Y}$ | Region | No | X, Y | Region |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 31,30 | City Hall | 26 | 42,22 | Poncol Station |
| 2 | 30,28 | Jl. Mgr <br> Sugiyopranoto | 27 | 33,36 | Paragon Mall |
| 3 | 27,29 | Kalibanteng | 28 | 34,31 | Kampung Pelangi Kalisari |
| 4 | 12,28 | Jl. Pamularsih <br> Raya | 29 | 28,24 | William Booth Public Hospital |
| 5 | 18,25 | Sam Poo Kong | 30 | 27,22 | Jl. S.Parman |
| 6 | 22,20 | Jl. Kaligarang | 31 | 29,19 | Taman Diponegoro Semarang |
| 7 | 25,20 | Kariadi <br> Hospital | 32 | 31,18 | Akpol |
| 8 | 28,21 | Jl. Veteran | 33 | 32,15 | Karangrejo |
| 9 | 31,19 | Jl. Pahlawan | 34 | 33,12 | Jatingaleh |
| 10 | 37,21 | Simpang Lima | 35 | 32,10 | Gombel Baru |
| 11 | 38,24 | Jl. Ahmad Yani | 36 | 33,6 | Diponegoro Statue |


| 12 | 42,22 | Jl. Majapahit | 37 | 32,3 | Tembalang |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 13 | 54,17 | Jl. Gajah | 38 | 38,1 | Gombel Lama |
| 14 | 56,25 | Great Mosque <br> of Central Java | 39 | 32,8 | Karangrejo Raya |
| 15 | 54,28 | Jl. Jolotundo | 40 | 30,10 | Unika |
| 16 | 54,27 | MT. Haryono | 41 | 27,11 | Akpelni |
| 17 | 45,27 | Bubakan | 42 | 24,14 | Jl. Menoreh Raya |
| 18 | 44,38 | Kota Lama | 43 | 22,14 | Taman Sampangan |
| 19 | 42,39 | Tawang Station | 44 | 22,17 | Jl. Kelud Raya |
| 20 | 43,42 | Pasar Johar | 45 | 24,18 | Taman Madukoro |
| 21 | 40,37 | Sri Ratu <br> Pemuda | 46 | 26,20 | Jl. Indraprasta |
| 22 | 37,36 | BCA Pemuda | 47 | 26,27 | Jl. Imam Bonjol |
| 23 | 36,34 | Jl. Thamrin | 48 | 30,30 | Jl. Hasanudin |
| 24 | 35,29 | Kampung Kali | 49 | 30,32 | Jl. Peres |
| 25 | 37,27 | Jl. Gajahmada | 50 | 32,31 | Jl. Petek |

Then from this table is implemented into a 2-dimensional map or graph that shows the path used at the beginning before the algorithm analysis process. This graph will function as a dataset, because per point the area is drawn according to the coordinates of point X and point Y.

## Illustration 4.1: Initial Graph



Because in this study the dataset will be tested with the Greedy Algorithm and the Genetic Algorithm, the coordinates per point are taken respectively as a dataset that can be inputted in the coding of each algorithm. Then for the distance per point in Illustration 4.1, it is determined from the boxes that are crossed by a line from one point to another. From this, it can be calculated the overall distance of the graph which can later be compared with each algorithm.

Illustration 4.2: Initial Graph with Range


From Illustration 4.2 we can determine the total distance from the graph. Which serves as one of the final results that can be compared against the results that will be generated by the Greedy Algorithm and Genetic Algorithm.

Table 4.2: Initial Sequence and Distance

| Sequence (Normal) | Total Distance |
| :---: | :---: |
| $1-2-3-4-5-6-7-8-9-10-11-12-$ |  |
| $13-14-15-16-17-18-19-20-21-22$ |  |
| $-23-24-25-26-27-28-29-30-31-$ | 281 |
| $32-33-34-35-36-37-38-39-40-41$ |  |
| $-42-43-44-45-46-47-48-49-50-1$ |  |

After getting the initial data, proceed with searching for the results of the Greedy Algorithm analysis. In the Greedy Algorithm, only the coordinates of the X and Y points in the programming are needed. Unlike the Genetic Algorithm which requires a variety of variables and several trials to get optimal results.

Illustration 4.3: Graph after the calculation uses the Greedy Algorithm


In Illustration 4.3, it can be seen that the resulting graph is different from the initial graph. This is because Illustration 4.3 has already passed the programming of the Greedy Algorithm. Then the distance per point can be taken from the boxes that are crossed by each line so that the total distance of the entire trip is obtained.

Illustration 4.4: Graph with distance after using the Greedy Algorithm


| Sequence (Greedy) | Total Distance |
| :---: | :---: |
| $1-48-2-3-47-29-30-8-31-9-32$ |  |
| $-33-34-35-39-36-37-38-40-41-$ |  |
| $42-43-44-45-7-46-6-5-4-49-50$ | 225 |
| $-28-24-25-11-10-12-17-26-23-$ |  |
| $22-21-19-18-20-27-15-16-14-13$ |  |

From the data in table 4.3 it can be seen that for the time being, the total distance generated from the two datasets, the first is the initial data and the second is the data that has passed through the Greedy Algorithm process, while the total distance from the Greedy Algorithm is smaller than the initial data. Data from table 4.3 shows the route generated by the Greedy Algorithm. Routes taken (1) Balaikota, (48) Jl. Hasanudin, (2) Jl. Mgr Sugiyopranoto,
(3) Kalibanteng, (47) Jl. Imam Bonjol, (29) RSU William Booth, (30) Jl. S.Parman, (8) Jl. Veteran, (31) Taman Diponegoro Semarang, (9) Jl. Pahlawan, (32) Akpol, (33) Karangrejo, (34) Jatingaleh, (35) Gombel Baru, (39) Karangrejo Raya, (36) Patung Diponegoro, (37) Tembalang, (38) Gombel Lama, (40) Unika, (41) Akpelni, (42) Jl. Menoreh Raya, (43) Taman Sampangan, (44) Jl. Kelud Raya, (45) Taman Madukoro, (7) RS Kariadi, (46) Jl. Indraprasta, (6) Jl. Kaligarang, (5) Sam Poo Kong, (4) Jl. Pamularsih Raya, (49) Jl. Peres, (50) Jl. Petek, (28) Kampung Pelangi Kalisari, (24) Kampung Kali, (25) Jl. Gajahmada, (11) Jl. Ahmad Yani, (10) Simpang Lima, (12) Jl. Majapahit, (17) Bubakan, (26) Stasiun Poncol, (23) Jl. Thamrin, (22) BCA Pemuda, (21) Sri Ratu Pemuda, (19) Stasiun Tawang, (18) Kota Lama, (20) Pasar Johar, (27) Paragon Mall, (15) Jl. Jolotundo, (16) MT. Haryono, (14) Masjid Agung Jawa Tengah, (13) Jl. Gajah.

Henceforth, the data process using the Genetic algorithm also uses the same initial data but with the addition of the variables needed in the Genetic Algorithm.

Genetic Algorithm itself requires several supporting variables in order to be able to program its path and produce results that can be compared with the previous two data. In addition to the initial input data, the variables required by the Genetic Algorithm are generation probability, crossover probability, mutation probability, and population probability. Populations are collections of individuals that will be processed together in an evolutionary cycle. Crossover is a cross between two mothers to produce new offspring. Mutations are a way to randomly cross genes from data. Generation is a unit of the evolutionary cycle. Individuals are a collection of genes / chromosomes. Fitness is a value that states whether or not an individual.

Selection in Genetic Algorithms is the process of selecting to get a good parent so that the fitness value will be high and produce a new generation. In this study, the selection method used was elitist and tournament selection methods. Tournament selection is a selection that fills a new population of a pair of parents who win the tournament based on fitness scores. Elitist selection itself is a method in which individuals are selected in the order of fitness values. The better the fitness value, then it can be used for comparison with other individuals and produce a population.

For the data to be used by the Genetic Algorithm, the initial data is the same as the Greedy Algorithm, which is in graph 4.1 and the data table in table 4.1. For data specifically needed Genetic algorithms are shown in table 4.4.

Table 4.4: Genetic Algorithm's Data Variable

| No | Mutation Probability | Number of Generations | Population Size |
| :---: | :---: | :---: | :---: |
| 1 | 0.01 | 50 | 150 |
| 2 | 0.01 | 100 | 100 |
| 3 | 0.01 | 150 | 50 |
| 4 | 0.01 | 250 | 150 |
| 5 | 0.01 | 50 | 50 |

Of the five data variables that will be modified in the Genetic Algorithm the best end result will be used as a comparison. Then it will be processed in the Genetic Algorithm for the five data with the regional data in table 4.1. Each of the data in table 4.4 will be tried 5 times. And 5 trials that apply to the five data in table 4.4. Then out of the 5 trials, the best results will be taken.

Table 4.5: The calculation results of Genetic Algorithm with Mutation Probability 0.1 , Generation 150, Population 50

| No | Distan | Tim | Route Order |
| :---: | :---: | :---: | :---: |
| 1 | 277.23 | $29.0$ | $\begin{gathered} 35-34-40-33-1-48- \\ 47-4-5-44-43-41-32 \\ -9-11-25-24-28-27- \\ 22-21-19-20-18-26- \\ 23-50-49-3-2-29-10 \\ -12-13-14-16-15-17 \\ -30-8-31-42-45-6- \\ 7-46-37-38-36-39 \end{gathered}$ |
| 2 | 343.68 | 27.3 | $\begin{aligned} & 22-19-20-18-14-13- \\ & 38-34-11-17-26-21- \\ & 24-25-10-9-32-50- \end{aligned}$ |


|  |  |  | $\begin{aligned} & 23-27-1-8-29-47-4 \\ & -5-45-6-7-30-46- \\ & 31-33-37-36-39-40- \\ & 35-41-43-44-42-12- \\ & 16-15-2-3-48-49-28 \end{aligned}$ |
| :---: | :---: | :---: | :---: |
| 3 | $302.24$ | $31.6$ | $\begin{aligned} & 22-21-16-15-18-20- \\ & 19-23-50-29-8-45-6 \\ & -46-31-33-35-38-37 \\ & -41-30-5-4-3-2-49 \\ & -27-28-24-25-11-48 \\ & -1-47-7-44-43-42- \\ & 40-39-36-34-32-9- \\ & 10-12-13-14-17-26 \end{aligned}$ |
| 4 | 329.35 | 29.1 | $\begin{aligned} & 1-2-36-38-37-39-34 \\ & -50-49-3-4-5-44- \\ & 40-41-42-43-45-46- \\ & 30-7-6-28-24-26-17 \\ & -13-12-10-9-8-25- \\ & 11-14-16-15-18-20- \\ & 19-21-22-23-27-48- \\ & 32-33-35-31-29-47 \end{aligned}$ |
| 5 | $353.96$ | $28.1$ | $\begin{aligned} & 13-16-15-14-17-12- \\ & 11-30-46-7-3-24-27 \\ & -49-48-1-25-10-33- \\ & 34-35-32-9-31-8-19 \\ & -20-18-28-47-44-43 \\ & -45-6-5-4-42-37- \\ & 38-36-29-2-50-23- \\ & 22-21-26-39-40-41 \end{aligned}$ |

Table 4.5 shows the results of one of the modification experiments of the variable mutation probability, generation probability and population size conducted 5 times. The probability of mutation is 0.01 , the probability of generation 150 and population size 50 indicate that the results are not constant. The point is the results obtained, there is a process of finding a path that takes a long time and there is a fast. The calculation above is based on python programming which produces distance, time, and route order.

The data generated by the program is then drawn graphically using the sequence of routes produced and calculated the distance using point x and point y in the graph. Each of the 5 experiments was taken the fastest results and converted into a graph and calculated the distance.

Table 4.6: Modified data with the best results after using the Genetics algorithm


|  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :--- |

Table 4.6 shows the modification of the five data used and processed using the Genetic algorithm 5 times and the best data is taken by looking at the fastest distance from the 5 trials. From this table, the most optimal data using Genetic Algorithm is data with probability 0.01 , Number of Generations 250 and Population Size 150 with route sequence (46), J1. Indraprasta, (31) Taman Diponegoro Semarang, (32) Akpol, (9) Jl, Pahlawan, (8) Jl. Veteran, (30) Jl. S.Parman, (29) William Booth Public Hospital, (10) Simpang Lima, (11) Jalan Ahmad Yani, (12) Jl. Majapahit, (13) Jl. Gajah, (14) Great Mosque of Central Java, (16) MT. Haryono, (15) Jl. Jolotundo, (17) Bubakan, (25) Jl. Gajahmada, (26) Poncol Station, (23) Jl. Thamrin, (18) Kota Lama, (20) Pasar Johar, (19) Tawang Station, (21) Sri Ratu Pemuda, (22) BCA Pemuda, (27) Paragon Mall, (50) Jl. Petek, (49) Jl. Peres, (3) Kalibanteng, (47) Jl. Imam Bonjol, (4) Jl.

Pamularsih Raya, (5) Sam Poo Kong, (6) Jl. Kaligarang, (44) Jl. Kelud Raya, (43) Taman Sampangan, (45) Taman Madukoro, (7) Kariadi Hospital, (24) Kampung Kali, (28) Kampung Pelangi Kalisari, (1) City Hall, (48) Jl. Hasanudin, (2) Jl. Mgr Sugiyopranoto, (33) Karangrejo, (34) Jatingaleh, (38) Gombel Lama, (37) Tembalang, (36) Diponegoro Statue, (39) Karangrejo Raya, (35) Gombel Baru, (40) Unika, (41) Akpelni, (42) Jl. Menoreh Raya.

From the data from table 4.6 above then the total distance from the graph data that has been drawn using the sequence data from each of the best results is calculated in table 4.6.

Table 4.7: Data with Number of Generations 50 and Population Size 150

| Mp | $\mathbf{N g}$ | Ps | Distance | Time | Route Order | Distance from Graph |  |
| :--- | :--- | :--- | :---: | :---: | :---: | :---: | :---: |
| 0.01 | 50 | 150 |  |  |  |  |  |

Illustration 4.5: Graph of Number of Generations 50 and Population Size 150


Table 4.8: Data with Number of Generations 100 and Population Size 100


Illustration 4.6: Graph of Total Population 100 and Population Size 100


Table 4.9: Data with Number of Generations 150 and Population Size 50

| Mp | Ng | Ps | Distance | Time | Route Order | Distance from Graph |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.01 | 150 | 50 | 277.23 |  |  | $35-34-40-33-1-48-47-4-$ <br> $5-44-43-41-32-9-11-25-$ <br> $24-28-27-22-21-19-20-18$ |


|  |  |  |  |  | $8-31-42-45-6-7-46-37-$ <br> $38-36-39$ |  |
| :--- | :--- | :--- | :--- | :--- | :---: | :--- |

Illustration 4.7: Graph of Number of Generations 150 and Population Size 50


Table 4.10: Data with Number of Generations 50 and Population Size 50

| Mp | $\mathbf{N g}$ | Ps | Distance | Time | Route Order | Distance from Graph |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |


|  |  |  |  |  | $37-39-36-33-11-26-23-22$ <br> $-27-49-3-5-4-6-24-17-$ <br> $20-19-18-21-50-48-47-7$ |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 0.01 | 50 | 50 | 340.48 | 9.3 |  <br> $-46-43-44-29-45-42-41-$ <br> $40-30-28-1-2-8-31-9-32$ <br> $-25-13-14-16-15-10-12-$ <br> $35-34-38$ | 323 |

Illustration 4.8: Graph of Number of Generations 50 and Population Size 50


Table 4.11: Data with Number of Generations 250 and Population Size 150

| Mp | $\mathbf{N g}$ | Ps | Distance | Time | Route Order | Distance from Graph |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |


| 0.01 | 250 | 150 | 240.13 | 149.5 | $\begin{gathered} 46-31-32-9-8-30-29-10- \\ 11-12-13-14-16-15-17-25 \\ -26-23-18-20-19-21-22- \\ 27-50-49-3-47-4-5-6-44 \\ -43-45-7-24-28-1-48-2- \\ 33-34-38-37-36-39-35-40 \\ -41-42 \end{gathered}$ | 226 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |

Illustration 4.9: Graph of Number of Generations 250 and Population Size 150


From all data generated using Genetic Algorithm, it shows that each Generation Number and Population Size produce different Distance Length outputs. Because the Number of Generations and Population Size affect the processing time and also the length of the distance produced. The number of generations is the number of steps used by the Genetic algorithm to carry out the calculation. In general, if the Number of Generations is greater, then the steps taken by the Genetic algorithm to do the calculation will be more and more, and that will provide an opportunity for the Genetic algorithm to do even more calculations and be better
able to produce a more optimal / shorter Long Distance output instead of using a lower Number of Generations. From table 5.1 with Number of Generations 50 when compared to table 4.1.10 with Number of Generations 250, table 5.0 has a lower length of distance compared to table 4.11. That's because the Genetic algorithm has more room to do calculations with 250 Number Generations compared to 50 Number Generations. But as the Number of Generations increases, the time spent doing the calculations will fall longer than with lower Number Generations.

Population size is a set of possible routes for processing generated by the Genetic Algorithm from its calculations. Table 4.7 has Number of Generations 50 and Population Size 150, while Table 4.11 has Number of Generations 50 and Population Size 50. From these two data tables, they have the same Number of Generations but their population size is different. From the results of these two tables, namely the output Length Distance, Population Size also affects the shortest result produced by the algorithm and also the length of time required by the algorithm to calculate. Table 4.7 has a lower distance length compared to table 5.1 but the length of time in table 4.7 is higher than table 4.11, which the time is 28.5 seconds. This is because the Population Amount in Table 4.7 is higher and provides an opportunity for the Genetic Algorithm to find the possibility of the route which allows for the fastest route to be found.

Then from the five data from each of the Number of Generations and Population Size that have been determined, in Table 4.11 it is superior in the results of the Long Distance of all existing data. Then this data is compared with the results of the Greedy algorithm.

Table 4.12: Genetic Algorithm and Greedy Algorithm Data

|  | Route Order | Distance |
| :---: | :---: | :---: |
| Normal | $\begin{aligned} & 1-2-3-4-5-6-7-8-9-10-11-12-13-14- \\ & 15-16-17-18-19-20-21-22-23-24-25-26- \\ & 27-28-29-30-31-32-33-34-35-36-37-38- \\ & 39-40-41-42-43-44-45-46-47-48-49-50- \end{aligned}$ <br> 1 | 281 |
| Genetic <br> Algorithm | $\begin{aligned} & 46-31-32-9-8-30-29-10-11-12-13-14-16 \\ & -15-17-25-26-23-18-20-19-21-22-27-50 \\ & -49-3-47-4-5-6-44-43-45-7-24-28-1- \\ & 48-2-33-34-38-37-36-39-35-40-41-42-46 \end{aligned}$ | 226 |


|  | $1-48-2-3-47-29-30-8-31-9-32-33-34-$ |
| :---: | :---: | :---: |
| Greedy | $35-39-36-37-38-40-41-42-43-44-45-7-$ |
| Algorithm | $46-6-5-4-49-50-28-24-25-11-10-12-17-$ |
| $26-23-22-21-19-18-20-27-15-16-14-13$ | 225 |

Illustration 4.10: Graph of Initial, Genetic Algorithm and Greedy Algorithm



From table 4.12, it is shown that both algorithms both provide the fastest route solution compared to the initial route, the Greedy Algorithm and the Genetic Algorithm are able to find the distance of the route which is faster than the normal route. However, if the two algorithms are compared even more, the Genetic Algorithm is superior to the Greedy Algorithm.

### 4.2 Design

In this research, the research design used is a case study research design. Which means, the data taken and the topics discussed are taken from one problem as an example in the community. This research is qualitative, where the data used is taken by observation. Then the data obtained in the algorithm will be examined in this study. There are several processes that occur during data processing in the algorithm so that it produces the output that is sought.


From the diagram above shows that how the calculation process uses the Greedy Algorithm. From the dataset that has been determined, this algorithm looks for the final result which is the fastest and optimal route that can be generated by the Greedy Algorithm itself.


The diagram above shows how the genetic algorithm works. Using a predetermined dataset, this algorithm is able to find its own optimal results according to their abilities.

