Searching Optimal Route for Public Transportation Of Palembang City Using A*Algorithm

Fithri Selva Jumeilah
Computer Science, STMIK GI MDP
Palembang Indonesia
fithri.selva@yahoo.com

Abstract— **Palembang is the capital city South Sumatra Province as well as big city. For a big city, transportation is a critical issue for society as well as its tourist. This study develops an optimal route search system using A * algorithm. A * algorithm using nearest distance estimates to achieve goals and have a heuristic value that is used as a basis for consideration. The resulting software is capable of finding the optimal transportation solution in the Palembang area so that people can get information easily and the accuracy of the application by 91.8%.**

***Keywords*— Algorithm A*, Transportation route search, Windows Phone.**

I. INTRODUCTION

Palembang is the capital city of South Sumatra Province. Palembang is one of the big city in Indonesia. For a big city, transportation is important for society as well as its tourist. Because of the wide area of Palembang and many roads cause difficulty to find the most optimal route in terms of distance and cost.

Conventional city maps often do not provide the fullest. This is due not everyone can read a map properly, making it difficult to determine the most optimal route. Currently, many digital maps are easy to use but the information provided is not equipped public transport information. For tourists, public transport information is very important to be help them to the location of destination cheaply and quickly without having to use a taxi.

The purpose of this research is to build an application that can facilitate the tourists to find public transportation that they need to get to the destination. To produce the optimal solution from the cost and distance it takes a search algorithm is A * [1].

II. LITERATURE REVIEW

A. Algorithm Search

Search algorithm is an algorithm that considers how to choose the optimal solution in the search [1,2]. At each search algorithm there is a difference in terms of development nodes to reach the goal state. There are four criteria for measurements search algorithm[2]:

1. Completeness: can algorithms certainly find a solution?
2. Time Complexity: how long algorithm takes time to find a solution?
3. Space Complexity: how much memory it takes to do a search?
4. Optimality: can the algorithm find the best solution?

Even the best search algorithm is not necessarily appropriate for all types of cases or problems, should be selected search algorithms appropriate to the needs of cases handled.

There are two search techniques based on the way developing the node that is uninformed Search or Blind Search and Informed Search or Heuristic Search [3,4,5]. Uninformed is searching for a solution algorithm without information that can direct the search to reach the goal state from the current state, the algorithm only can identify a goal state[1,2,5].

Informed Search is a strategy to make the process of a state space search problem selectively, which guide the search process of finding the best solution[1,2,3]. Informed Search often also called Heuristic Search as to reach the goal state uses rules to select the branches are most likely to resolve the problem that is received or often called heuristic function[4,6,7].

B. Algorithm A*

A * algorithm generates an optimal path from the initial state and then through the graph towards the destination. The algorithm is classified as path finding algorithm which uses a technique Informed Search. In addition to calculating the path cost of the current state to the goal state with heuristic function, A * algorithm is also considering the cost of the path that has been taken so far from the initial state to the current state [1]. So when a road has been taken and there is another way that has a lower cost but provide the same position seen from the goal state, the lower the new road that will be chosen. This algorithm calculates the heuristic value based on equation 1[1,4,2].

$$f(x) = g(x) + h(x) \quad (1)$$

Description :

f(x): evaluation function,
g(x): costs that was issued from the initial state to node x so far,
h(x): estimated cost from the node x to reach the goal.

Principles of A * algorithm is to find the shortest path from an initial state to the destination node by finding the node that has the value f (x) lowest. At each step of the process A * search, select the node that has the highest priority. This is done by applying the evaluation function f (x) which is adequate at each node. The lower value of f (x), the higher priority. If one node is a goal then stop. If not, do the selection of the node again.

C. Geographic Coordinates

On the map there is a conventional longitudinal stripes (vertical) and transverse (horizontal) which will help to determine the position of a place on earth. The intersection between the lines of longitude and latitude named maps coordinates. Longitude-latitude coordinate system consists of two components that determine, namely:

1. The lines from top to bottom (vertical) which connecting north to the south pole of the earth, also known as latitude.
2. Horizontal line parallel to the equator, also called longitude

To calculate the distance between two coordinates using the formula Haversine as in equation 2[8].

$$d = R \times c \tag{2}$$

Description:
d: the distance between two points selected according R unit
R: determination the radius of the earth is 6,371 (km)[8].

To get value of c must use the equation 3

$$c = 2 \times atan2\ (\sqrt{a}, \sqrt{1-a} \tag{3}$$

where A is obtained from equation 4, 5 and 6

$$a = \sin^2 \frac{\Delta Lat}{2} + \cos\left(\frac{\pi\ x\ Lat1}{180}\right) \times \cos\left(\frac{\pi\ x\ Lat2}{180}\right) \times \sin^2\left(\frac{\Delta Lon}{2}\right) \tag{4}$$

Where ΔLat get from:

$$\Delta Lat = \pi \times (Lat2 - Lat1)/180 \tag{5}$$

Description:
Lat2 : Latitude of the destination point,
Lat1 : Latitude of the initial point,

$$\Delta Lon = \pi \times (Lon2 - Lon1)/180 \tag{6}$$

Description:
Lon1 : Longitude of the destination point
Lon2 : Longitude of the initial point

III. Algorithms and Implementation

A. A An Artificial Inelegant Algorithm for finding Directions*

The main problem of this research is how to find the most optimal solution from the starting point to the destination point entered by users in the region of Palembang. The resulting solution is the result of the consideration of aspects of distance and cost to implement aspects of the A * algorithm, but it is more focused on the aspects of distance.

One example of a search problem using public transport is the routing from the Start point to the PS point as shown in the figure below:

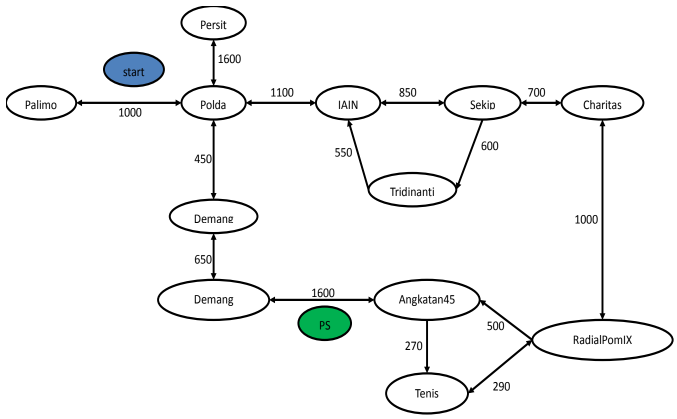

Fig. 1 Example of Map Search

To determine value of h of the problems in Fig. 1 can be seen in Table I which is the length of the straight line obtained from the calculation of the equation 1 calculation by using latitude and longitude coordinates at each node.

TABLE III
Straight Line Length of Each Intersection toward Destination (PS)

No	Node	Distance to Goal (h(x))(m)
1	Start	1887
2	PS	0
3	Polda	1739
4	Palimo	2665
5	DemangAryodila	1470
6	Persit	2.594
7	IAIN	1210
8	Tridinanti	686.7
9	DemangAngktan45	1237
10	Sekip Pangkal	1067
11	Charitas	1309
12	Angkatan45Tenis	282.4
13	TenisPomIX	416.7
14	RadialPomIX RivaiAngkatan45	658

Figure 1 is an example of a settlement made by A * start from Polda to the PS node. Starting with the opening Demang aryodila branch and consider the best nodes. Node that has the smallest value of f will open all of its successors and will look for a branch that has the smallest value of F to find the destination node. A * algorithm can optimize the search, because the A * algorithm to consider the two costs, that is the cost estimates to the destination and the cost incurred.

After obtaining the last intersection, then do BackTrack to check Parent node, to find the Parent equal initial Intersection. To calculate the required fees must be checked by means of a series of sequentially tracing the intersection of the solution and count how many times a change of public transport. At issue this time is not a change of public transport, so fare is Rp. 3.000. While the optimum of this problem is: Start Start→Polda→DemangAryodila→DemangAngkat45→PS.

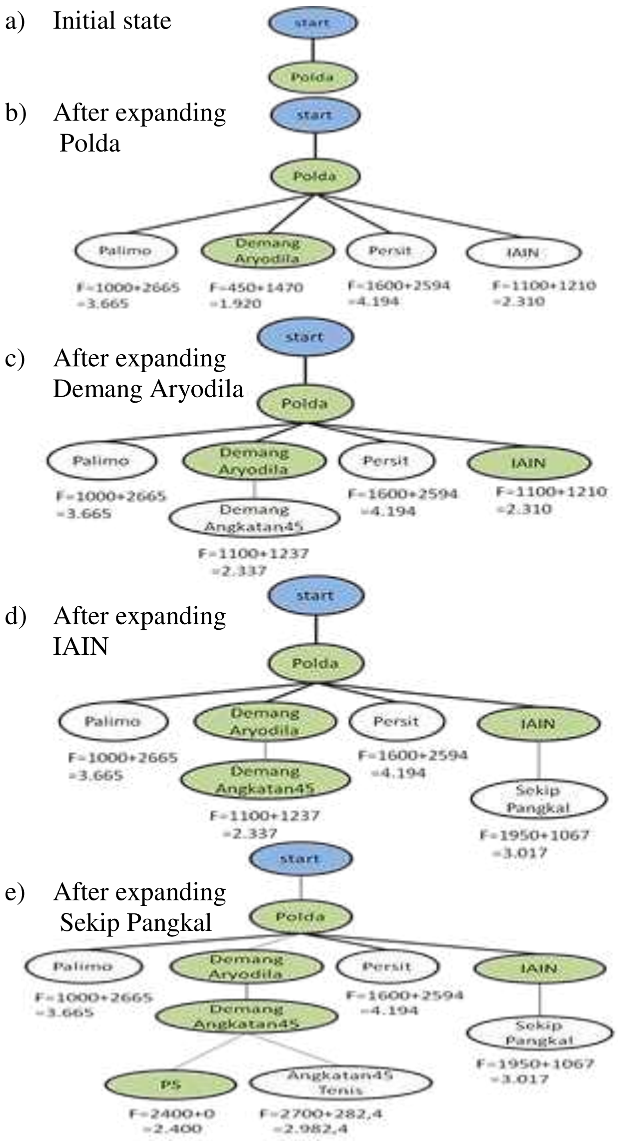

Fig. 2 Steps for resolve the problem Fig 1

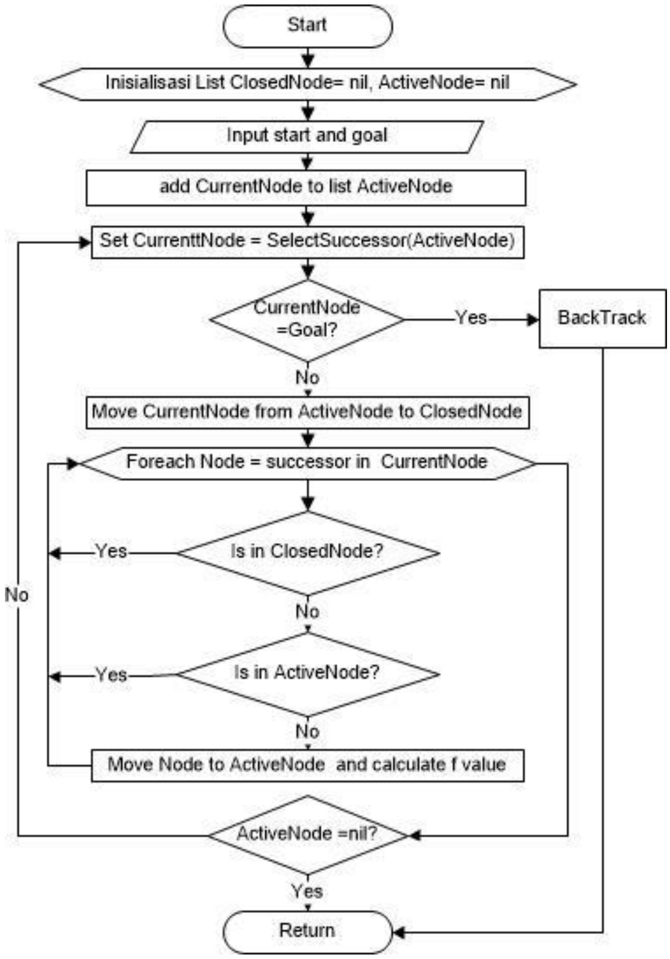

Fig. 2 Flowchart of algorithm A*

B. Implementation

Palembang city Public transportation consists of bus medium, small bus, and generally urban transportation. For the public transportation routes used in this application was obtained from the Department of Transportation of Palembang City in South Sumatra. In this study, every intersection is traversed by public transport will be set as a node. For the cost incurred from the initial state to a node x is the actual distance and transportation rates. For the estimate the cost of a node x is the distance to the destination obtained from equation 2. For the map used is a digital map of the Bing map. All the coordinates of the intersection will be stored in the database.

Figure 3 flowchart above describes how search software for public transportation routes. Search results very dependent on the A * algorithm, in which the algorithm works from the starting point to the destination point by finding the intersection with the estimated weight of the lowest intersection (denoted by the notation g) summed with the lowest estimate of the intersection of weight to achieve the goal (denoted by the notation h). The route has a junction with value of f (x) is the smallest and can be reached by public transport in the Parent nodes will be developed first.

If the destination node has been found, the program will perform backtrack to the Parent of each node to get a series of intersections that form the most optimum.

IV. Analysis and Result

Here's the interface of the application if we have input the origin and destination, it will show the route of solutions that have been generated by the A * algorithm. For the management of the map just send a command to Bing maps so that the solution generated by the A * algorithm can be color.

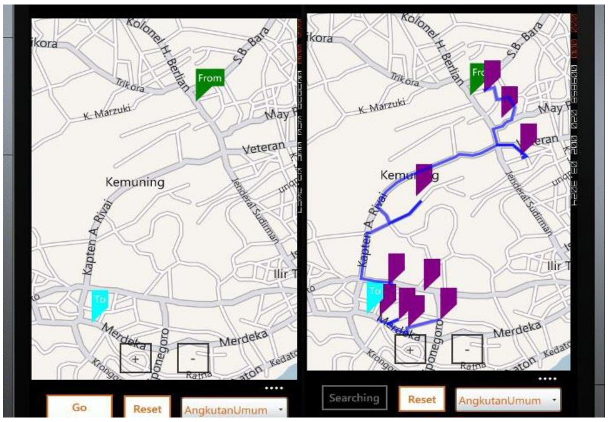

Fig. 4 Example of route interface

Besides route, users also get of public transportation information what should they drive and they will also need to know the cost of that spend to achieve the destination.

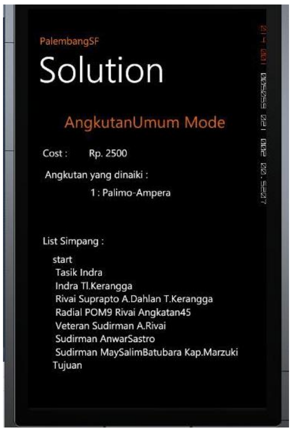

Fig. 5 Example of solution interface

From Origin coordinates, find the intersection at the start the closest to the intersection point and that intersection is Polda. Figure 6 is the interface of the application testing program is a modified application of the actual application. The solution shown is three solutions that has the smallest f value of all the possibilities, and we can see that the application is able to choose the solution which has the smallest f value.

Solution 1 is the best solution that is has the smallest F value, so the actual application that is shown only solution 1. the optimum solution that is a solution which has the smallest f value. After seeing the solution obtained is the optimum solution, it can be concluded that the analysis and the design is correct, so that the result in the implementation of the optimum solution.

For the testing, samples taken from all of public transportation route random. But for a sample can have more than one of public transportation so that sample taken is less than the number of public transportation. In the application of digital maps used are Bing Maps, therefore testing is done by comparing the distance generated by the application and the results obtained Bing Maps. Error is obtained from the equation:

$$Error = \frac{|Distance\ from\ Bing\ Maps - Distance\ from\ Aplikasi|}{Distance\ from\ Bing\ Maps} \times 100\% \qquad (7)$$

Fig.6 Example if shown many solutions and include with value of f

The average value of the percentage error is 8.2%, the percentage of accuracy: 100% -8.2% = 91.8%. Then from the results of Table IV-6 test, the accuracy of the identification of this application reaches 91.8%. This is caused by:

1 Applications take the distance between the intersection with a straight line and count them by using equation 2 is not using the actual distance. It is caused by the existing map Palembang on Bing Maps many have the difference.
2 Results calculated by Bing Maps already rounded, so that the data obtained had changed.
3 There is a public transportation routes that does not exist on Bing Maps, while the application is not in route Bing Maps coordinates taken from Google Maps so as to give effect to the accuracy of the application.

TABLE II
RESULTS OF TESTING SAMPLES

No	Coordinat		Distance		Error
	Start	Goal	BingMap	Application	
1	-2.999477, 104.768149	-3.019877, 104.749095	3300	3101	6%
2	-3.00804, 104.75990	-2.99073, 104.78522	3500	3370	3.7%
3	-2.995706, 104.76471	-2.991506, 104.74866	2400	2327	3%
4	-2.971020, 104.749266	-2.9816379, 104.7598	2300	2211	4%
5	-2.969252, 104.729776	-2.952537, 104.755611	3600	3512	2.4%
6	-2.9636810, 104.74187	-2.938823, 104.7205924	3800	4249	11.8%
7	-2.987338, 104.739561	-2.991881, 104.7264289	2200	2080	5.4%
8	-2.978509, 104.74531	-2.967366, 104.73149	4100	2426	40.8%
9	-2.953111, 104.767228	-2.933654, 104.76731	2200	2489	4%
10	-2.9308253, 104.768344	-2.918653, 104.78199	2300	2128	7.4%
11	-2.9822552, 104.762164	-2.979769, 104.77675	1800	1957	8.7%
12	-2.977283, 104.752723	-2.964169, 104.767658	2800	2682	4.2%
13	-2.976683, 104.75117	-2.975140, 104.77117	2100	1990	5.2%

V. CONCLUSIONS

The conclusions of this research are:

1. PalembangSF application is a search application using the A * algorithm has been successfully searching the optimal solutions of public transportation route based on distance and cost solutions with high accuracy percentage that is 91.8%.
2. This application can provide a list of what transportation should be used along with the location down comes the closest to the destination.

Suggestions for further development of this research include the following:

1. The data used by PalembangSF for of public transportation travel mode is the data Palembang City Department of Transportation in 2002, so that many transport routes that do not correspond to the current route.
2. Use a digital map of Palembang more complete, so the search is not only based on the coordinates but also by address.
3. The data used is the latest data, the solutions produced in accordance with current conditions. Data should be stored in a centralized data storage server so that the data processing easier.

REFERENCES

[1] Desiani, Anita and arhami. 2007. Konsep Kecerdasan Buatan. Andi Publiser. Yogyakarta.
[2] Russel, S. and P. Norvig. 1995. Artificial Intelligence: A Modern Approach. Prentice Hall, New Jersey.
[3] Chen, C. and Y. Cheng. 2008. Road Digital Map Generation with Multi-track GPS data. School of Electronics and Information Engineering, Beijing Jiaotong University. China.
[4] Suyanto.2011. Artificial Intelegence: Serching, Reasoning, Planing, Learning. Informatika. Bandung.
[5] Nilsson, Nils.1998. Artificial Intelligence: A New Synthesis. Morgan Kaufmann. China.
[6] Kusumadewi, Sri. 2003. Artificial Intelligence (Teknik dan Aplikasinya). Graha Ilmu. Yogyakarta.
[7] Krishnamoorthy; S and Rajeev. 1996. Artificial Intelligence and Expert Systems for Engineers. CRC Press. LLC.
[8] Pescaru, Dan and Curiac Daniel-Ioan.2014. Anchor Node Localization for Wireless Sensor Networks Using Video and Compass Information Fusion. Sensors 2014, 14, 4211-4224; doi:10.3390/s140304211.