Implementation of Kruskal's Algorithm to Optimizing Estimation Mega Project Cost in Indonesia (Sumatera, Java and Bali Focus)

Gazandi Cahyadarma, 13513078

Program Studi Teknik Informatika Sekolah Teknik Elektro dan Informatika Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia gazandic@students.itb.ac.id

Abstract— Optimization is one of the most important aspects in both engineering and sciences fields. There are many major problems happened regarding to the optimization problems. The implementation of the optimization process will result the most optimal solution.

The best optimized solution will have minimum or maximum values among set of possible solutions. In mega project cost, optimization problem also occurred on direct cost and time installation process. Therefore, this issue has to be highlighted and must implement efficient algorithm to find the minimum length of connected places that built in mega project. The project must connecting places because the equation for estimation project cost will be helped by Kruskal's. This paper will focus on the application of Kruskal's algorithm in finding the most optimized project path by representing it as a weighted graph in minimum spanning tree form.

Keywords—mega project cost, graph, minimum spanning tree, optimization

I. INTRODUCTION

Nowadays, there are numerous problems can be modeled by using graph. Graph is one of the most common data structures that can be represented in a set of vertices and edges. This data structure can be applied for finding the minimum spanning tree within a graph. One of the applications of spanning tree that would be discussed inside this paper is about optimizing estimation mega project cost.

Today, mega project will exist until Indonesia ends and its demand in Indonesia has rapidly increased due to increment of citizens. But, the demand has not completely fulfilled due to unstable trust and economy,. This is occurred because too much corruption and too less trust, so writer thinks this problem can be solved by searching an optimized equation to check the mega project is optimized in cost or not. Thus, it can be useful for government observers. The Observer can see the mega project is effective or not. And it useful too project holder to optimizing his project.

Therefore, the project holder must find an efficient way to optimizing estimation project cost by using efficient algorithm such as Kruskal's algorithm.

In other hand, this algorithm will support the project

holder to optimize the usage of cost in long-term period and "long distance" projects. Long distance means the project must be connecting places with long distance like island to island, province to province, etc. This beneficial fact must also be highlighted in order to reduce the cost of Mega Project. Thus, this paper will introduce a way to find the minimum spanning tree from minimum distance to connect places by implementing Kruskal's Algorithm.

Depend on wikipedia, there are very much mega project in Indonesia that proposed to government, like Trans-Sumatra toll road, but there are much project in Sumatera and Java. Therfore, Bali is popular in this world and very potential to build some mega project. Thus, that three island are most dense island in Indonesia. So, writers focus in Sumatera, Java, and Bali in this paper.

II. FUNDAMENTAL THEORIES

2.1 Definition of Graph

Graphs are discrete structures consisting of vertices and edges that connect these vertices. [1] There are different types of graphs, regarding whether edges have direction, same pair of vertices, or loops are allowed. We use this structure to model acquaintanceships between things for instance organizations, people, species, etc.

Mathematically, graph is defined as G = (V,E) where •V = a nonempty set of vertices (or nodes) •E = a set of edges. Each edge has either one or two vertices associated with it, called its endpoints. An edge is said to connect its endpoints. Here is the example of graph:



FIGURE 2.1 A GRAPH WITH 5 EDGES AND 5 VERTICES/NODES.

The graph above is undirected graph that has 5 edged and 5 vertices. The vertices are a,b,c,d, and e.

2.2 Weighted Graph

A weight graph is a graph G in which each edge e has been assigned a real number w(e), called weight (or length) of e. [1] This weight of each edge represents the distance between two vertex/node. Here is the example:



Figure 2.2 Weighted Graph

According to Figure 2.2, every edge between two vertices has weight that represents the distance between them.

2.3 Spanning Tree

A spanning tree in a graph G is a minimal subgraph connecting all the vertices of G. [3] The spanning tree is represented as an undirected graph. If we consider that the graph is a weighted graph, then the weight of a spanning tree of G is defined as the sum of weights of all branches.





2.4 Minimum Spanning Tree

A minimum spanning tree in an undirected connected weighted graph is a spanning tree of minimum weight. The minimum weight is the term given to a spanning tree which the sum of its edges is minimal. For instance, take a look at Figure 2.4.



FIGURE 2.4 WEIGHTED GRAPH

Figure 2.4 above is a weighted graph. The graph has a minimum spanning tree and illustrated in Figure 2.5.



Figure 2.5 Minimum Spanning Tree from Graph in Figure 2.3

The minimum spanning tree obtained by using Prim's Algorithm or Kauskal's Algorithm. In Figure 2.5, we can discover that the minimum value of the sum of its weight is 71.

2.5 Kruskal's Algorithm

Kruskal's algorithm is a greedy algorithm in graph theory that finds a minimum spanning tree for a connected weighted graph. This means it finds a subset of the edges that forms a tree that includes every vertex, where the total weight of all the edges in the tree is minimized. If the graph is not connected, then it finds a minimum spanning forest (a minimum spanning tree for each connected component).

This algorithm first appeared in Proceedings of the American Mathematical Society, pp. 48–50 in 1956, and was written by Joseph Kruskal.

Other algorithms for this problem include Prim's algorithm, Reverse-delete algorithm, and Borůvka's algorithm.

Now, take a look at pseudo-code for Kruskal's algorithm

KRUSKAL(G): $1 A = \emptyset$ 2 foreach v \in G.V: 3 MAKE-SET(v) 4 foreach (u, v) ordered by weight(u, v), increasing: $5 = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{i=1}^{N} \sum_{i=1}^$

5 if FIND-SET(u) \neq FIND-SET(v):

6 $A = A \cup \{(u, v)\}$ 7 UNION(u, v)8 return A

Here is the execution of Kruskal's Algorithm as the example:





Table I . Kruskal's Algorithm

2.6 Estimation Project Cost

In a world of limited funds, as a project manage you're constantly deciding how to get the most return for your investment. The more accurate your estimate of project cost is, the better able you will be to manage your project's budget. Therefore, estimating a project's costs is important for several reasons: It enables you to weigh anticipated benefits against anticipated costs to see whether the project makes sense. It allows you to see whether the necessary funds are available to support the project. It serves as a guideline to help ensure that you have sufficient funds to complete the project.

Although you may not develop and monitor detailed budgets for all your projects, knowing how to work with project costs can make you a better project manager and increase your chances of project success.

A project budget is a detailed, time-phased estimate of all resource costs for your project. You typically develop a budget in stages — from an initial rough estimate to a detailed estimate to a completed, approved project budget. On occasion, you may even revise your approved budget while your project is in progress.

Your project's budget includes both direct and indirect costs.

Direct costs include the following:

-Salaries for team members on your project

-Specific materials, supplies, and equipment for your project

-Travel to perform work on your project

-Subcontracts that provide support exclusively to your project

Indirect costs fall into the following three categories:

Overhead costs: Costs for products and services for your project that are difficult to subdivide and allocate directly. Examples include employee benefits, office space rent, general supplies, and the costs of furniture, fixtures, and equipment. You need an office to work on your project activities, and office space costs money. However, your organization has an annual lease for office space, the space has many individual offices and work areas, and people work on numerous projects throughout the year. Because you have no clear records that specify the dollar amount of the total rent that's just for the time you spend in your office working on just this project's activities, your office space is treated as an indirect project cost.

General and administrative costs: Expenditures that keep your organization operational (if your organization doesn't exist, you can't perform your project). Examples include salaries of your contracts department, finance department, and top management as well as fees for general accounting and legal services.

Contingency costs: additional cost that be alocated because there are maybe additional job (whenever not happen yet).

III. THE MEGA PROJECT AND THE RELATION OF ESTIMATION MEGAPROJECT COST AND KRUSKAL'S ALGORITHM IN INDONESIA

3.1 Mega project

A mega project is an extremely large-scale investment project. Mega projects are typically defined as costing more than US\$1 billion and attracting a lot of public attention because of substantial impacts on communities, environment, and budgets. Mega projects can also be defined as "initiatives that are physical, very expensive, and public". Care in the project development process is required to reduce any possible optimism bias and strategic misrepresentation, as a curious paradox exists in which more and more mega projects are being proposed despite their consistently poor performance against initial budget and schedule forecasts.

Mega projects include bridges, tunnels, highways, railways, airports, seaports, power plants, dams, wastewater projects, Special Economic Zones, oil and natural gas extraction projects, public buildings, information technology systems, aerospace projects, weapons systems, large-scale sporting events and, more recently, mixed use waterfront redevelopments; however, the most common mega projects are in the categories of hydroelectric facilities, nuclear power plants and large public transportation projects.

3.2 Relation Estimation Mega project Cost and Kruskal's Algorithm

In deciding estimation project cost, there are direct cost that include material cost or labor cost. Material cost sell by the material seller with per meter or kilometer price. It will easy to calculate the optimizing price if the project has small scoop. But, if it concern the large scoop like a country, example Indonesia, it will be hard to calculate the optimize cost. So, we need something to calculate the optimize. It not will be only for macro engineer or macro contractor, but it can be for honest politician or macro projects observer to checking the proposal of mega project is optimized or not. With kruskal's algorithm, we can make a spanning tree with optimized weight. It can be implemented in mega project that have business with material cost which has must be built and connecting the capital of province in Indonesia. Like, subway train that connecting all capital of province in Indonesia. Or mega project to connect all capital of province in Indonesia with fiber optics cable to make a new operator in Indonesia, or this is can be used to make optimized logistics estimation cost. So, kruskal's algorithm will be beneficial if kruskal's algorithm use to optimize estimation mega project cost that must be connecting the capital of province in Indonesia m or connecting places in Indonesia. And, Kruskal's Algorithm be used in deciding material cost, labor cost, and Land Cost.

3.3 Indonesia Map Analysis

The current Indonesia Map can be showed in Figure 3.1(taken from petacitra.blogspot.com). With the map, we will be analyzed is the three most dense and potensial island in Indonesia : Java, Sumatera, and Bali.



Figure 3.1 Indonesia Map

There are sunda strait that separate java and sumatera, and there are strait that separate java and bali, to make a effective spanning tree, we will make 3 spanning tree which consist java spanning tree, sumatera spanning tree, and bali spanning tree because there are very much cities. We will make regions of spanning tree for

better analyzing and implementation. Because bali just have one capital of Bali province then, bapeta indonesiali spanning tree just have one nodes, Denpasar nodes. And to connect three spanning three, we just searched on the internet which capital of province that connect sumatera and java with smallest distance and which capital of province that connect java and bali with the smallest distance. We found the smallest distance that connect capital of province in sumatera and java is Bandar Lampung-Serang. And we found the smallest distance that connect capital of province in bali and java is Surabaya-Denpasar.

Based on figure above, the vertices represent theintersection of each district and the edges represent the distance between each vertex. In every vertex is given coordinate as the representation of each location in real map and in each of edge is given weight that represents the length between each vertex within the scale of the map. the distance in kilometer in x y must be convert into length with this function :

length =
$$\sqrt{(x1 - x2)^2 + (y1 - y2)^2}$$

x1 = x coordinate of vertex 1

 $x^2 = x$ coordinate of vertex 2

y1 = y coordinate of vertex 1

 $y_2 = y$ coordinate of vertex 2

This equation (1) will be used to plot the length of

every edge. Every edge stands for length between two

nodes/vertices. And we must convert the degree with :

1 degree = 111.322 Km = 111.322 meters

- 1 degree = 60 minutes = 3600 seconds
- 1 minute = 60 seconds
- 1 minute = 1.885,37 meters 1 second = 30.9227 meters

IV. IMPLEMENTATION OF KRUSKAL'S ALGORITHM IN OPTIMIZING ESTIMATION MEGA PROJECT COST

4.1 Sumatera Spanning Tree

4.1.1 Cities Distance

Sumatera has 8 provinces with 8 capital in it. The location of each capital can be viewed in table below :

Cities	Coordinates
Banda Aceh	5°33′N95°19′E
Medan	3°35′N 98°40′E
Padang	0°57′0″S 100°21′11″E
Pekanbaru	0°32′0″N 101°27′0″E
Batam	1°05′N 104°02′E
Jambi	1°35′24″S 103°36′36″E
Palembang	2°59′10″S 104°45′20″E
Bandar Lampung	5°25′46,6″S 105°15′45,26″E

Table 4.1 Cities and Coordinates

Connected A-B	Cities	Distance in Coordinates	Distance in kilometer
Banda medan	aceh-	2° 02' , 3°21'	436 km
Banda padang	aceh-	6° 30' , 5°02'11" ,	916 km
Banda batam	aceh-	4° 28' , 8°43'	1092 km

Banda aceh- pekanbaru	5°01′0″, 6°8′0″	882 km
Banda aceh- jambi	7° 08' 24" , 8°17'36,36"	1218 km
Banda aceh- palembang	8° 32' 10" , 9°26'20"	1418 km
Banda aceh- bandar-lampung	0° 58' 46,6" , 9°56'45,26"	1651 km
Medan -padang	4° 32' , 1°41'11"	540 km
Medan - batam	2° 30' , 5°22'	660 km
Medan – pekanbaru	3°03′0″ 2°47′0″	460 km
Medan – jambi	5° 10' 24" , 4°56'36"	798 km
Medan – palembang	6° 32' 10" , 6°5'20"	995 km
Medan – bandar lampung	9° 00' 46,6" , 6°35'45,26"	1244 km
Padang – Pekanbaru	1° 29' , 1°05'49" ,	206 km
Padang – Batam	2° 02' , 3°40'49"	469 km
Padang – Jambi	0° 38' 24" , 3°15'25"	417 km
Padang – Palembang	2° 02' 10", 4°06'03"	509 km
Padang – Bandar Lampung	4° 28' 46,6", 4°54'34,26"	741 km
Pekanbaru – Batam	0° 33' , 2°40'49"	306 km
Pekanbaru – Jambi	2°07′24″ , 2°09′36″E	337 km
Batam – Jambi	2° 40' 24" , 0°25'14"	302 km
Batam – Palembang	4° 04' 10" , 0°43'20"	460 km
Batam – Bandar Lampung	6° 30' 46,6" , 1°13'45,26"	738 km
Jambi – Palembang	1° 23' 46", 1°08'44"	201 km
Jambi – Bandar Lampung	3° 50' 22,6", 1°39"9,26"	468 km
Palembang – Bandar Lampung	2° 26' 36,6" , 0°30'25,26" ,	278 km

Table 4.2 Distance of the cities

4.1.2 Sorting the Distance & Graph for Cities in Sumatera

After we get distance between cities, we model a graph that relate between that cities on figure 4.1 below



Figure 4.1

the distance in table 4.2 has not been sorted yet, so we must sort the distance because we make the Kruskal's Algorithm.

Smallest Distance (ascending)	Connected Cities A-B	Distance in kilometer
1	Jambi – Palembang	201 km
2	Padang – Pekanbaru	206 km
3	Palembang – Bandar Lampung	278 km
4	Batam – Jambi	302 km
5	Pekanbaru – Batam	306 km
6	Pekanbaru – Jambi	337 km
7	Padang – Jambi	417 km
8	Banda aceh- medan	436 km
9	Medan – pekanbaru	460 km
10	Batam – Palembang	460 km
11	Jambi – Bandar	468 km

	Lampung	
12	Padang – Batam	469 km
13	Padang – Palembang	509 km
14	Medan -padang	540 km
15	Medan - batam	660 km
16	Batam – Bandar Lampung	738 km
17	Padang – Bandar Lampung	741 km
18	Medan – jambi	798 km
19	Banda aceh- pekanbaru	882 km
20	Banda aceh- padang	916 km
21	Medan – palembang	995 km
22	Banda aceh- batam	1092 km
23	Banda aceh- jambi	1218 km
24	Medan – bandar lampung	1244 km
25	Banda aceh- palembang	1418 km
26	Banda aceh- bandar- lampung	1651 km

Table 4.3 Distance of the cities (sorted)

4.1.3 Kruskal's Algorithm in Cities in Sumatera

In this part, we will model Cities in Sumatera with Kruskal's Algorithm. The process of modelling can be seen on the table below :



4.2 Java Spanning Tree

4.2.1 Cities Distance

Java has 8 provinces with 8 capital in it. The location of each capital can be viewed in table below :

Cities	Coordinates
Serang	6°7'12"S 106°9'1"E
Jakarta	6°12′S 106°49′E
Bandung	6°54′53.08″S 107°36′35.32″E
Semarang	6°58′S 110°25′E
Yogyakarta	7°48′5″S 110°21′52″E
Surabaya	7°15′55″S 112°44′33″E

Table 4.5 Cities and Coordinates

Connected Cities A-B	Distance in Coordinates	Distance in kilometer
Serang- Jakarta	0°4′48″, 0°39'59″	76 km
Serang – Bandung	0°47′41.08″, 1°27′34.32″	186 km
Semarang – Serang	0°50'48" , 4°15'59"	485 km
Yogyakarta – Serang	1°40′53″, 4°12′51″	505 km
Surabaya – Serang	1°8′43″, 6°35′32″	746 km
Jakarta – Bandung	0°42′53.08″, 0°47′35.32″	121 km
Jakarta – Semarang	0°46′, 3°36′	411 km
Jakarta – Yogyakarta	1°36′5″, 3°32′52″	434 km
Jakarta – Surabaya	1°3′55″S 5°55′33″Е	672 km
Bandung – Semarang	0°3'6.92", 2°48'24.68"	314 km
Bandung – Yogyakarta	0°53'11.92", 2°45'16.68"	324 km
Bandung – Surabaya	0°11'1.92", 5°7'57.68"	572 km
Yogyakarta – Semarang	0°50′5″, 0°3′08″	94 km
Surabaya – Semarang	0°17′55″, 2°19′33″	261 km
Surabaya – Yogyakarta	0°32'10", 2°22'40"	272 km

Table 4.6 Distance of the cities

4.2.2 Sorting the Distance & Graph for Cities in Java

After we get distance between cities, we model a graph that relate between that cities on figure 4.2 below:



the distance in table 4.6 has not been sorted yet, so we must sort the distance because we make the Kruskal's Algorithm.

Smallest Distance (ascending)	Connected Cities A-B	Distance in kilometer
a	Serang- Jakarta	76 km
b	Yogyakarta – Semarang	94 km
с	Jakarta – Bandung	121 km
d	Serang – Bandung	186 km
e	Surabaya – Semarang	261 km
f	Surabaya – Yogyakarta	272 km
g	Bandung – Semarang	314 km
h	Bandung – Yogyakarta	324 km
i	Jakarta – Semarang	411 km
j	Jakarta – Yogyakarta	434 km
k	Semarang – Serang	485 km
1	Yogyakarta – Serang	505 km
m	Bandung – Surabaya	572 km

n	Jakarta – Surabaya	672 km
0	Surabaya – Serang	746 km

Table 4.7 Distance of the cities

4.2.3 Kruskal's Algorithm in Cities in Java

In this part, we will model Cities in Java with Kruskal's Algorithm. The process of modelling can be seen on the table below :





Table 4.8. Kruskal's Algorithm in graph at figure 4.2. the spanning tree has 866 km in distance.

4.3 Spanning Tree of Java, Sumatera and Bali

Denpasar has 8°39'S 115°13'E coordinates. And it has coordinates distance and kilometer distance with Surabaya : <1°23'5", 2°28'27">, 316 km. Therefore, Serang and Bandar Lampung has coordinates distance and kilometer distance : <0°41'25,4", 0°53'15,74 ">, 127 km. And it will produce the great spanning tree below :





4.4 The Estimation of Mega Project that Involve All Capital of Province in Sumatera, Java and Bali

If there is a megaproject that involve All Capital Province in Sumatera, Java and Bali like make a new operator of telecommunication that must built a fiber optic that must be connect all capital in province, the equation of the estimation of mega project cost will be :

Direct Fibre Optic Cost = 120/100 * 3500 * (Fiber Optics Cost/kilometer)

Direct Labour Cost = 120/100 * 3498 * (Labour Salary to install fibre optic/kilometer)

Direct Tools Rent Cost = 120/100 * 3498 * (Tools

rent/kilometer)

Direct Property Cost = 0 (because doesn't need buying a land)

Direct Other cost = 120/100 * 3498 * (Other

cost/kilometer)

Undirect Cost = Overhead cost + Administration Cost + other cost

there are 120/100 * 3498 cost because the spanning tree has 3498 kilometer and in real world, the road not gave the best distance because this earth has irregular terrain. So, we must give 20% additional to give compensation.

So the Cost will be = 150/100 (Direct Fibre Optic Cost + Direct Labour Cost + Direct Tools Rent Cost + Direct Other cost + Undirect Cost)

and there are 150/100 because, 10% additional is PPN, 20% is benefit for contractors, and 20% is contingency cost. Because in mega project there are will come much obstacle like to free a soil or rent soil, or whatever that must be need.

But, the equation above just example, to get the true equation we must turn the spesific cost into general cost. Like turn the fibre optic with material. So the optimized equation for megaproject that involve all capitals of province in java, sumatera, and bali will be like that :

Optimized equation cost = 150/100 ((120/100 *3498 * (Material cost/km + Labour Salary/km + Tools Rent/km + Property Cost/km + Other Cost/km))+ Undirect Cost)

the optimized equation cost above just not can be used if involve all capitals of province in java, sumatera, and bali. It can used too if 3498 will replaced by another total weight in spanning tree. Like if we want to megaproject just in sumatera, the optimazed equation will be :

Optimized equation cost = 150/100 ((120/100 *2189 * (Material cost/km + Labour Salary/km + Tools Rent/km + Property Cost/km + Other Cost/km))+ Undirect Cost)

3498 replaced 2189, because in Sumatera the total weight in Sumatera spanning tree is 2189 km. So, the equation can be universal for all project that connecting places or something that can be optimized to be spanning tree.

V. CONCLUSION

Graph theories can be applied in representing to optimizing estimation cost with map of certain areas. The vertex/node of the graph represents the intersection of each location and the edge represents the length between the vertices. In optimizing estimation cost is using the weight graph and each edge is given weight that represents the length between the vertices by using the formula for finding the length between two nodes.

The implementation of Kruskal's Algorithm is very useful in optimizing estimation cost. This algorithm would able to find the shortest spanning tree within the map that represented by graph. This would effectively reduce the cost and optimizing the the project The spanning tree obtained from this algorithm would be used as the optimized distance in mega project to produce optimized equation for search estimation mega project cost. This equation can help the project holder to optimize the cost and time. This equation can help too the observer of government beware government not to cheat and corrupt because the cost can be predict.

VII. ACKNOWLEDGMENT

First of all, I would thank to God for His perfect providence while preparing, writing, and editing the paper. Then, I deeply express my thanks to my advisors, Dr. Ir. Rinaldi Munir and Mrs. Harlili, whose help, advice, and support. I also thank to my parents and friends or the prayers and valuable courage given to me. Finally, I would thank to Institut Teknologi Bandung for the hospitality and support during the completion of the paper.

REFERENCES

- [1] John O. Clark and Derek Allan Holton, A First Look a Graph
- Theory. Singapore: World Scientific, 1991
- [2] K. H. Rosen, Discrete Mathematics and Its Applications 7 th . New York: McGraw-Hill, 2012
- [3] N. Deo, Graph Theory with Applications to Engineering and
- Computer Science. New Delhi: Prentice-Hall, 1974
 [4] Rinaldi Munir, Matematika Diskrit 3 rd , Bandung: Informatika , 2009
- [5] For Dummies, How to Estimate Project Costs, 10 Desember 2014 <u>http://www.dummies.com/how-to/content/how-to-estimate-projectcosts.html</u>

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 10Desember 2014

13513078/ Gazandi Cahyadarma

Makalah IF2120 Matematika Diskrit - Sem. I Tahun 2014/2015