CHAPTER 2

STUDIES IN SCHOOL BUS ROUTING AND NETWORK ANALYSIS

2.1 Introduction

This chapter, which covers two main discussions, will explain several aspects which one needs to fully understand before proceeding with research. The first part addresses school bus routing and findings from previous researches, and then moves to an examination of network analyst concept that will be applied in this particular research with respect to school bus routing. Apart from these two main discussions, there are also short explanations of several related areas, namely school buses, needy families, VRP, and accessibility.

2.2 School Bus Transport

There is no school bus operated in Surabaya. Surabaya student usually using a city bus or a lyn, together with others people in different age and occupation. Lyn is a modified station wagon. It removes usual seats in the second and third row and replaces them with long seats patched at the left and right sides. The capacity of the city bus is around 50 while the lyn is 10 passengers. Figure 2.1 shows the general look of these vehicles.

Figure 2.1 The city bus and the Lyn

A school bus is commonly identified as a yellow bus designed and manufactured specifically for student transport. It carries children and teenagers to and from school or to any outside location for school events. It is standard for a school bus to be equipped with specific warning and safety devices. The first school bus ever used to carry 25 children was noted in 1827 in the north-east of London (UK). Today, according to modern nomenclature (NASDPTS 2000) , there are 4 types of school buses; type A, type B, type C, and type D.

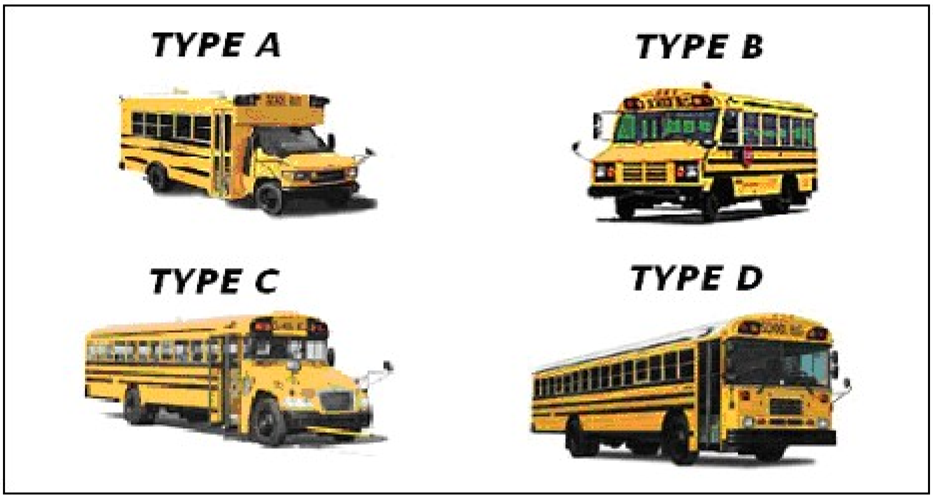


Figure 2.2 Four different types of school bus

Type A school buses are the smallest of its kind. This type is usually a modified van-type vehicle and the driver is located in a different area than where students are seated. The student is placed on a cutaway van chassis with one door. There are single or dual rear wheels on drive axle. This type of school bus carries more than 10 passengers and typical passenger capacity ranges from 10-16

Type B school buses are larger and heavier than Type A. Today, this type is less commonly produced than in the past. The driver and student sections placed in one body. There is one door placed behind the front wheels. Commonly, the engine compartment is located partially inside the passenger compartment next to the driver. This type of bus usually carries more than 20 passengers while typical passenger capacity ranges from 30-36.

The Type C school bus is also known as a "conventional" school bus. It has a passenger compartment which is installed in the back of a flat-back cowl. The gross vehicle weight is more than 10,000 pounds. The engine is placed in front of the windshield and the entrance door is behind the front wheels. This type of bus usually carries more than 20 passengers while typical passenger capacity ranges from 36-78.

The Type D school bus is known as a transit-style school bus. The form of this type is like the usual transportation bus. The gross vehicle weight is more than 10,000 pounds and the location of the engine is variable. It may be behind the windshield or beside the driver's seat. It may also be located at the back of the bus behind the rear wheels over even between the front and rear axles. The entrance door can be found ahead of the front wheels. This type of bus usually carries more than 20 passengers while typical passenger capacity ranges from 54-90.

As a basic standard, school buses have at least one emergency exit door or window in addition to a primary entry door. This additional door may be located in the roof, window exits, or side emergency exit. The number of emergency exits on a school bus depends on the size of the bus, and all exits must be clearly visible and opened immediately upon the activation of an emergency alarm.

For the purposes of this study, school bus type D will be referred to in all research. The dimensions of this particular type are considered and used to justify the street data in the ground survey. Details of dimension will help determine such factors as whether street traffic can pass by the school bus or not. Also, the size of the vehicle will be taken into consideration when examining its ability to handle specific turns and junctions on proposed bus routes. In terms of passenger capacity, this research uses the median number 70. This number consists of 50 sitting passengers and 20 standing passengers.

2.3 Definition of Needy Children

According to the Oxford Dictionary, the term "needy" means lacking the necessities of life or one is very poor. Using this definition, a needy family implies a family which is very poor. The definition of poor and the specific living conditions considered poverty level vary with each country. In Indonesia for example, needy specification is formulated by Centre of Statistical Bureau.

The Surabaya city government has implemented the formula from Centre of Statistical Bureau (BPS) about how to classify the needy citizen. This formula determines whether a citizen is at or below the poverty line. The formula has 4 factors: housing, health, education, and economic. These factors are divided into 14 variables and indicators. If the characteristic of the citizen meets the criteria, this variable equals 1 point, and if not this variable shows zero and is thus not applicable. The total value of these variables indicates the poverty level of the citizen so that a citizen with 14 points is categorized as extremely poor, a citizen with 11 to 13 points is categorized as poor, and a score of 9 to 10 points indicates almost poor. Citizens with points fewer than 9 are not included in the needy category. Table 2.1 shows the 14 variables.

Table 2.1 Needy family criteria

NO	Sector	Variable	Indication
1	Housing	Floor area per family member	Less than 8 m2 per family member
2	Housing	Type of the floor of the house	Most of the part is sand or cement
3	Housing	Type of the wall of the house	Most of the part is bamboo or low quality plywood
4	Housing	Toilet facilities	Have no private toilet
5	Housing	Drinking water source	Not from government water supply or not a good water
6	Housing	Lighting source	Non-electric
7	Housing	Fuel used	wood or charcoal
8	Health	Frequency of Meals in One Day	2 times or less
9	Health	Ability to buy Meat, Chicken, Fish, Eggs, and Milk in a week	Have no ability
10	Health	The ability get a treatment in the hospital / clinic / other health center	Have no ability
11	Education	Education level of the householder	never schooling or not passing elementary level
12	Economy	The ability to buy new clothes for every family member	One or less per year
13	Economy	The occupation of the householder	Farmers, Fishermen, Daily Labor, or other jobs with incomes below Rp. 600,000 per month
14	Economy	Ownership of assets or goods worth IDR 500.000, - (houses, cars, motorcycles, cattle, etc)	Does not have

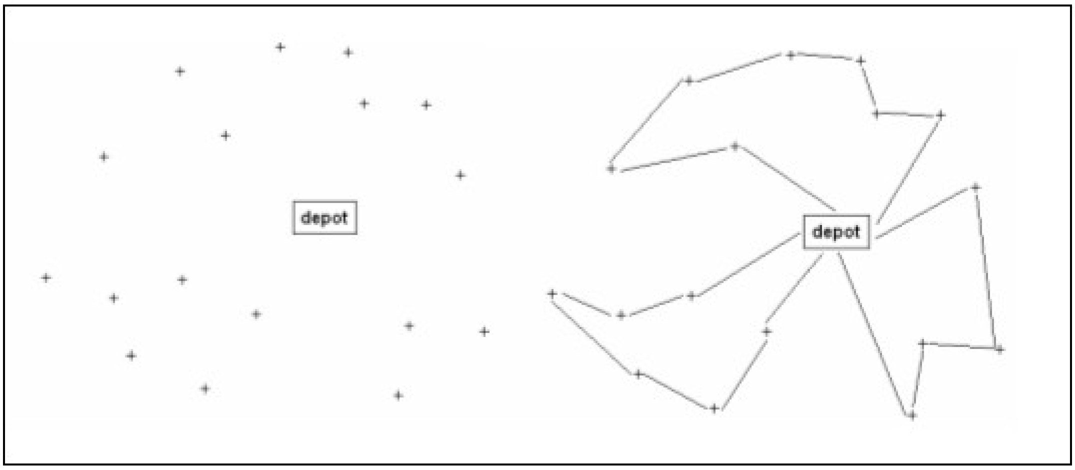
According to table 2.1, the following example of living conditions demonstrate the characteristics of someone categorized as extremely poor: very little house with a cement floor, bamboo walls, lacking a toilet and electricity, water supply is taken from a river, wood or charcoal used for cooking. This type of individual likely has a meal two times or less per day, and cannot afford to eat meat, chicken, fish, egg, or milk on a weekly basis. If suffering any illness, he is unable to afford hospital treatment. In terms of education, he has not completed elementary school and generates little income from his job and few assets. His clothes are very limited, and he can purchase just once during the year.

2.4 Finding Optimal Route

This section will introduce how to find the optimal route. Finding an optimal route is the goal of the well known term Vehicle Routing Problem (VRP). Vehicle Routing

Problem (VRP) is widely considered a success story among operation research techniques. The large number of real-world applications have widely shown that the use of computerized procedures for the delivery system planning produces significant savings in the overall transportation costs (Toth & Vigo, 2002; Golden et al, 2008).

VRP is a generic name given to a whole class of problems whose purpose is to determine a number of routes for a fleet of vehicles that originated and terminated from one or several depots, and have to make a visit in a number of geographically dispersed cities or customers (Bernabe 2007). The example of VRP input and output is shown in Figure 2.3. The objective of the VRP is efficiency---to deliver a set of goods or customers with known demands at a minimal cost. The main components of VRP include the network (usually in form of road networks), and the demand (usually goods or customers) as well as depots, vehicles, and drivers.

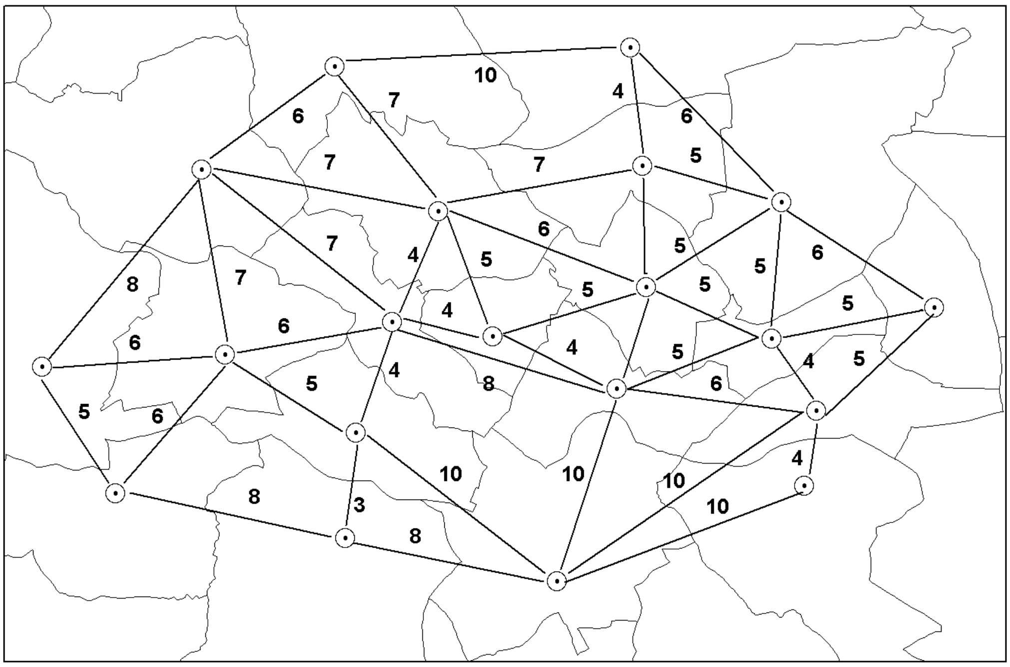


(source: Bernabe 2007)

Figure 2.3. A.VRP Input: Depot and location ; B. VRP Output : Vehicle route

Commonly the road network is described in a graph with the arcs representing the road sections. The vertices correspond to the road junctions and to the depot and customer locations. Figure 2.4 shows the example of this network. The arcs can be directed or undirected, and this depends on whether they can be traversed in only one direction, or in both directions, respectively. The one direction can be caused by several factors such as the presence of one-way streets. Every arc is equipped with a cost. This cost is often represented with respect to its length and its travel time. The cost is dependent on the vehicle type or on the period during which the arc is traversed (Toth & Vigo 2002).

Customer location is symbolized with vertex of the road graph. Customers can have any amount of goods which must be delivered or collected. The goods could be uniform or vary in description. Customers can be served in specific periods throughout the day. The vehicle can deliver or collect at various times as well, and can be used as a subset of the available vehicles. Sometimes it's not possible to fully satisfy the demands of each customer, and for this reason, a subset of customers can be left without service or the amounts to be delivered or collected can be reduced. In these situations, a strategy has to be adapted. Different customers priorities can be assigned to determine which customers can be safely left without service. Penalties associated with the partial or total lack of service can be assigned as a response to the loss generated from omitting or not delivering service for customers and/or goods.



(source : http://jasss.soc.surrey.ac.uk)

Figure 2.4. Representing route in graph

Depot is located at the vertices of the road graph. Vehicle routes begin and end at this designation. Depot can serve one or more customers and has a certain number of vehicles attached to it. The attached vehicles may be uniform or varying types. Depot is also determined by certain goods. Types of goods and the amount of goods that can be handled are a common characteristic. Depot locations can be partitioning (found in different locations) within the customer's area. A depot has a responsibility for servicing customers from a certain area and a vehicle has to return to a specified home depot at the end of each route. For this multi-depot case, the overall VRP can be deconstructed into several independent problems, each associated with a different depot (Toth & Vigo 2002).

Vehicles that transport the goods can have fixed features or vary in composition and size. As mentioned, vehicles must be attached to the depots and be associated with a specific capacity which could be in the form of weight, height, width, volume, or number of pallets. It is a common to subdivide the vehicle into compartments. Each compartment is characterized by capacity and by the type of goods that can be carried. Vehicle can perform loading operations while traversing in the subset of arcs of the road network. In its utilization, vehicle is associated with the cost. The cost is the sum of the distance unit, or time unit, or number of routes, or any other pertinent variables.

A driver's characteristics are sometimes included in the routing tasks. Drivers operate vehicles which introduces several points to consider in analysis. When determining the routing calculation, the cost may be affected by the number of drivers, overtime costs, the maximum duration of driving periods, number and duration of breaks during service, or even the working periods during the day.

In the routing task, the routes must satisfy several constraints. This constraint can be caused by the previous aspects discussed (vehicle, depot, customers, goods, and driver). Along with dealing with these constraints, vehicle routing has to optimize one or more objectives. Toth and Vigo (Toth & Vigo 2002) have listed several typical objectives of vehicle routing which include:

1) Minimization of the transportation cost which is generated by the traveled distance, the consumed time, and fixed costs. Fixed costs are often associated with the use of the vehicle and its driver.
2) Minimization of the number of vehicles required to serve customers.
3) Balancing the route so that travel times and cost are in balance.
4) Minimization of penalties which are generated by lack of service for customers.

2.4.1 VRP Types

Within its characteristics, VRP can be divided into several models. The following are collected from several sources (Diaz, 2007; Toth and Vigo, 2002; Minic, 1998; Wassan & Osman, 2002) and include :

1) **Capacitated VRP (CVRP)**

   Capacitated VRP is the basic version. CVRP is VRP with a fixed fleet of delivery vehicles with identical capacity. They must service customer demands for goods from a common depot at minimum transport cost. CVRP is like VRP with an additional constraint, the uniform capacity. The objective of CVRP is to minimize the travel time, or even better minimize the vehicle fleet, with the total demand of goods for each route not exceeding the capacity of the vehicle which serves it. The solution is called feasible when the total quantity of goods which are demanded do not exceed the capacity of the vehicle associated in that route.

2) **Distance Constrained VRP (DVRP)**

   Distance Constrained VRP is the variant of CVRP. The Capacity term is replaced with Distance term. Distance is associated with maximum length or time constraint. The objective of the problem is to minimize the total length of the routes or of their duration when the service time is included in the travel time of the arcs. The solution is called feasible when the total length of passed street or total time in servicing the customer in each route does not exceed the distance constraint of the vehicle associated in that route. The case in which both the vehicle capacity and the maximum distance constraints are present is called Distance-Constrained CVRP (DCVRP).

3) **Multiple Depots VRP (MDVRP)**

This is a VRP model for a company which has more than one depot. Vehicles can begin and end the services for their customer in these depots. This model is not applied in the depots where customers are clustered rather than centralized. If the customers are clustered around depots, the distribution problem should be modeled as a set of independent VRPs. On the other hand, if the customer and the depots are intermingled, then this MDVRP model should be used. Assignment of customers to depots is required in this model. A fleet of vehicles is based at each depot. Each vehicle starts from one depot, services the assigned customers, and then returns to that depot. The objectives of the problem involve servicing all customers while minimizing the number of vehicles and total travel distance.

4) **Periodic VRP (PVRP)**

In this VRP model, serving time can be planned in a period of days, unlike the classical VRP that focuses on the serving time over the span of a single day. There is parameter planning period M. If M = 1 then PVRP becomes processed like classical VRP. The objective is minimizing the sum of travel time and the number of vehicle fleets in a certain period which are needed to supply all customers. The vehicle may stay in specific place during the period and not return to the depot in the same day if it departs. A solution is feasible if all constraint of VRP is satisfied.

5) **Split Delivery VRP (SDVRP)**

SDVRP is a variation of the VRP in which the same customer is allowed to be served by different vehicles. This model fits in the case where goods from the customer are as big as or bigger than the capacity of the vehicle. For this situation, the order will split into smaller orders and the new sized order may be serviced in different vehicles. This splitting make the original VRP become SDVRP. The objective is to minimize the sum of travel time and the number of the vehicle fleet needed to supply all customers. A solution is feasible if all constraints of VRP are satisfied with allowing some customers to be supplied by more than one vehicle.

6) **Stochastic VRP (SVRP)**

In the basic VRP, all components have fixed characteristic measurements. Stochastic VRP (SVRP) is VRP in which one or more components of the problem is random. It can be stochastic in customers, stochastic in demand, or stochastic in time. Each customer can present or diminish in any probability. The demand of each customer can be expressed as a random variable. Travel times and service times for every path or customer is not fixed but represented in random variables. The objective is to minimize the sum of travel time and the number of vehicles needed to supply all customers with random values in their present, demands, and service times. The random data makes it is almost impossible to be satisfied. It normally needs adjustment from the decision maker to determine, with a given probability, which constraints have to be fully satisfied and less satisfied. When constraints are violated, corrective actions need to be determined.

7) **VRP with Pick-up and Delivering (VRPPD)**

VRP with pickup and delivering presents a problem of finding a set of optimal routes for a fleet of vehicles, in order to serve a set of transport action requests. Vehicles in this kind of VRP need to visit all pickup and delivery locations and fulfill the precedence and pairing constraints. Precedence constraints means there are pickup locations which have to be visited before visiting the corresponding delivery locations. Pairing constraints mean there are legal routes for one vehicle to do both the pickup and the delivery of the load of one transport action request. This restriction makes the planning problem more difficult to solve. For example, it can lead to complex calculation of the vehicle's capacities. It may increase travel distances or increase requirements for more fleets than another VRP types. To simplify the complexity, it is normal to make a situation in which all pick-up demands have to be brought to the depot first before continuing to deliver to customers. All delivery demands start from the depot. It creates a scenario in which there are no interchanges of goods from customer to customer directly along the routes. A typical or usual simplification involves requiring every vehicle to deliver all the goods in delivery orders before picking up any goods in picking orders.

8) **VRP with Backhauls (VRPB)**

One of the simplifications of the VRRPD is also known as VRP with Backhauls (VRPB). The problem with this type of VRP arises from the fact that the vehicles are rear-loaded, making the rearrangement of the loads on any place in the journey of the vehicle unfeasible. All deliveries must be made on each route before any pickups can be made. The customers are divided into two groups. The first group contains Linehaul customers which have demands that certain quantity of goods must be delivered. The second group contains Backhaul customers which have demands that specific quantities of goods are to be picked up. The quantities to be delivered and picked-up are fixed and are known before the vehicle travels along its route.

9) **VRP with Time Windows (VRPTW)**

The VRP with Time Windows (VRPTW) is the extension of the VRP in which customers are associated with a time interval in which they can be serviced. This time interval is called a time window. The time is generated when the vehicles leave the depot, and added in the traveling process and when servicing each customer. The service of each customer must start within the associated time window, and the vehicle must stop at the customer location for start time periods. If the vehicle arrives early, it has to wait until the start of the time window.

10) **Mix-Fleet VRP (MFVRP)**

The Mix-Fleet Vehicle Routing Problem (MFVRP) is a variant VRP in which the vehicle is not uniform but has a number of types. Vehicles are characterized by their capacity, fixed costs, and the variable cost per distance unit. The objective is to determine the optimal mix of heterogeneous vehicles in order to minimize the sum of travel time and the number of vehicles. The total demand of the routes must in line with the capacity of vehicle assigned to it, rather than exceed it or fall short.

11) **Dynamic VRP (DVRP)**

The traditional VRP can be said to be static or deterministic. In other words, the demand of the customer does not change over the course of the day in which goods are delivered. In contrast, the dynamic vehicle routing problem (DVRP) takes into account the possibility of the changing demands from customers after the day of operation has begun. The DVRP has an ability to include new requests into the already designed routes. Ideally, the new customers should be inserted into the already planned routes with minimal delay and without changing the order of the non-visited customers. However, in practice, the insertion of new customers can generate complicated tasks and can disturb the scenario of the non-visited part of the route. The insertion has to meet the all constraint of the routing. Generally, the more restricted and complex the routing problem is, the more complicated the insertion of new dynamic customers will be. For example, the insertion of new customers in a time window constrained or capacitated constrained routing problem will be much more difficult than in a non- constrained problem.

Finally, to gain a better understanding or bigger picture of VRP research,, more information can be obtained by reviewing a survey by Burak Eksioglu and friends (Eksioglu 2009). They have a voluminous amount of taxonomies VRP research in a well-respected listing.

2.5 School Bus Routing

Several researchers have projects related to bus routing. Some of them focus-on the use of a new algorithm or advancing an existing algorithm, while others implement existing algorithms to real world problems (Dorronsoro Díaz, 2007; Toth and Vigo, 2002; Bruce Golden et al, 2008). In advancing algorithm, Robert Bowerman introduces a multi-objective approach to make a model of school bus routes in an urban area (Robert Bowerman, 1995). A heuristic algorithm, based on their formula tested with data from a sample school board location in Wellington County, Ontario, Canada, is used as reference for the model study. Some aspects in his work can be adopted in this research as well. Firstly, a comparison can be made in terms of creating clusters of students before generating a school bus route and bus stops for each cluster. Secondly, there are a defined set of optimization criteria used to evaluate the set of school bus routes. These include:

1. **Number of routes**

   Number of routes generated should be kept to a minimum in order to minimize the operational cost of the vehicles and routes.

2. **Total bus route length**

   The less the route length is associated with the less time consumed.

3. **Load balancing**

   One of the qualities aspects of the routes is the similarity between the numbers of the loads in each route. This load balance results in services balances for the student.

4. **Length balancing**

   Like previous criteria, length balancing leads to the services balance for the student as well as the balance of operational costs.

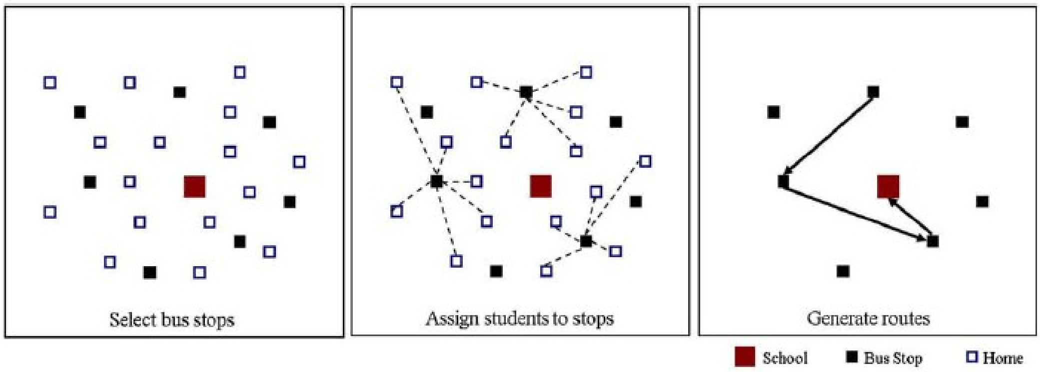
5. **Student walking distance**

   In order to make the route feasible to be used by the student, the student walking distance from home or from school along the route must be achievable.

Based on the number of schools, bus routing can be divided into many-to-one and many-to-several (Spada et al, 2005). An Example of many-to-one can be viewed in the work of M Fatih Demiral with colleague (Demiral et al, 2008) and Nayati Mohammed (Nayati, 2008). All use one school location as a depot and the student house location as the customer location for generating bus routes. They work with study areas at Isparta, Turkey, Hyderabad, and India respectively. Others who have conducted similar studies include Li and Fu (Li & Fu, 2002) and Bektas and friend (Bektas, 2007). Li and Fu implement a heuristic algorithm for an existing data related to a kindergarten in Hongkong, while Bektas uses integer programming for an elementary school located in central Ankara, Turkey. The results of their studies were impressive; they both promoted transportation initiatives saving 29% and 26% respectively for their generated new routes compared with the current implementation and associated expenditures.

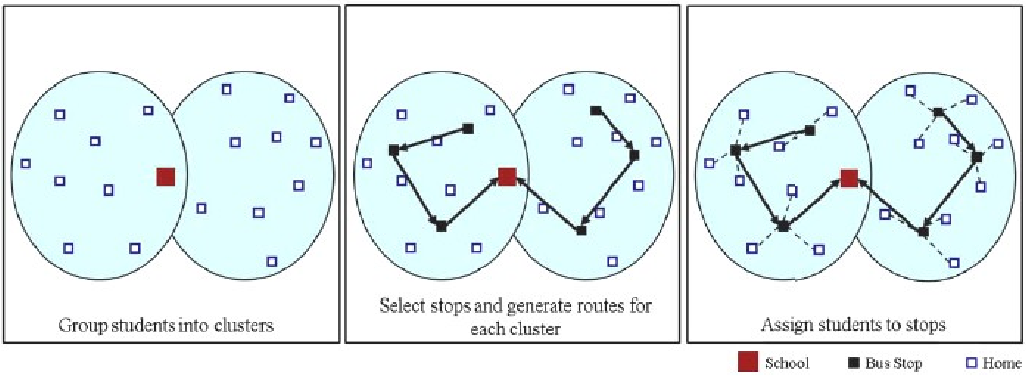
Based on the location or environment of the data, bus routing can be divided into urban and rural areas. In the urban area, the many-to-one study from Bektas and friend and Li and Fu can be a robust example. In the rural area, Armin Fu¨ Genschuh examine student locations in five counties in Germany.(Fu¨Genschuh.2009). While the destination involves multiple schools, instead of sending the bus back to the depot after having served a trip, he pushes to re-use the bus to serve other trips whenever possible. He integrates optimization of school start times with the optimization in school bus transportation. He believes that with this integration a single county can save up to 1 million Euro annually. Another work in rural area focusing in the advancing algorithm involved using rural school data in Savigny and Forel, Swiss (Spada et al 2004). This two rural area environment route proposes buses for multiple schools.

Park, J. and Kim, B. (Park & Kim, 2009) found that there are two strategies in heuristic solution approaches for bus stop selection. First is the location-allocation-routing (LAR) strategy, and the second is the allocation-routing-location (ARL) strategy. Figures 2.5 and 2.6 illustrate the concepts of these two strategies. In the LAR strategy, a set of bus stops for a school is determined and students are assigned to these stops. The generation of the routes is based on these selected stops. This approach tends to generate excessive routes, because the bus stops and the assignment of the students are determined without taking into consideration their effect on generating routes. The ARL strategy first attempts to make a cluster of students referring to the vehicle capacity constraint. Secondly, the task involves selecting bus stops and generating routes to these bus stops. Finally, the students in a cluster (route) are assigned to the bus stops. The assignment process must be satisfied and all the requirements given in the problem must be taken into account. Some requirements related with the bus stops are the number of students that can be served well at a given bus stop, the minimum distance between bus stops, and the maximum walking distance to catch the bus stop.



(source: Park & Kim 2009)

Figure 2.5. Concept of LAR Strategy



(source: Park & Kim 2009)

Figure 2.6. Concept of ALR Strategy

In the both concepts, it is assumed that students can walk to the bus stop. This assumption is used in almost all of bus school routing research. There is also the assumption that the bus must be picking up from and taking the student to their house, given that the bus is servicing special-education students. Examples of this kind of research and concept can be found in the work of Russel (Russel et al, 1986). They refer to the Tulsa Public School System in Tulsa, Oklahoma. The system has approximately 850 special-education students whose special needs are served by 66 different schools in the Tulsa area.

For gaining a better understanding and macro view of school bus study, we can view a good survey of School Bus Routing over almost 4 decades from Park and Kim (Park & Kim, 2009). They reviewed and classified sub-problem, characteristic of object, and solution method. For the problem characteristic, they had classified:

a) Based on number of schools, there are:
    i) Single school
    ii) Multiple schools
b) Based on surroundings of service, there are:
    i) Urban
    ii) Rural
c) Based on problem scope, there are:
    i) Morning
    ii) Afternoon
    iii) Both
d) Based on mixed loads, there are:
    i) Mixed loads are allowed
    ii) No mixed loads are allowed (single load problem)
e) Based on Special-education students, there are:
    i) Special-education students are considered
    ii) Only general students are considered
f) Based on fleet mix, there are:
    i) Homogeneous fleet
    ii) Heterogeneous fleet
g) Based on objectives, there are:
    i) Number of buses used
    ii) Total bus travel Distance or time

iii) Total student riding distance or time

iv) Student walking distance

v) Load balancing

vi) Maximum route length

vii) Child's time loss

h) Based on constraints, there are:

i) Vehicle capacity

ii) Maximum riding time

iii) School time window

iv) Maximum walking time or distance

v) Earliest pick-up time

vi) Minimum student number to create a route

Another researcher compares existing methodology with their case study data. Lazar and others (Lazar et al 2001) compare three different methodologies: time saving heuristic, Router software, and sweep algorithm for the Riverdale, Morris County, New Jersey. The have concluded that for one depot the router gives the best results while using multi depot, and the time saving heuristic provides the best result. They also suggest using GIS as a superior analytical tool and analyst and for varying the parameters.

Other researchers have focused on analyzing existing routes or transportation systems. Melissa Reese analyzed the rapid bus transport system in Curitia, Brazil (Reese, 2007). She used a 3D presentation to display her analysis. The coverage of facilities is another aspect that can be analyzed. The work of Schilling and others (Schilling et al, 1993) is an excellent review of this research. Most analyses use a buffer to represent and formulate the coverage area. The buffer uses distance as its primary parameter. Numerous research, especially in accessibility of transport systems, have calculated this distance. Most researchers take 400 meters as the recommended maximum distance from a place to transport lines (Dittmar & Ohland, 2004; NJTransit, 1994; Paget et al, 1992).

2.6 Problem Solution Definition

This section will determine the problem area of this research to describe its position in the VRP and school bus routing. This research is using the existing VRP algorithm and applies it to real life scenarios or problems. In the VRP area, this research has these characteristics:

a) This research is a CVRP because the school bus will have a fixed fleet of delivery vehicles with identical capacity.
b) This research is a MDVRP because the bus depot will be more than one.
c) This research is a VRPPD because the bus school will pick up the student and deliver him or her to school.
d) This research is a VRPTW because the bus has a limited time when delivering the students to schools.

In the classification from Park and Kim (2009), this research has a position in these areas:

a) Based on number of schools, this is a multiple schools research
b) Based on surroundings of service, this research is in urban area
c) Based on problem scope, this research is in the morning time
d) Based on mixed loads, this research does not allow a mixed load
e) Based on Special-education students, this research only considers general students
f) Based on fleet mix, this is a homogenous fleet

g) Based on objectives, this research has these objectives:

    i) Number of buses used

    ii) Total bus travel distance or time

    iii) Student walking distance

    iv) Load balancing

h) based on constraints, this research has following constraints:

    i) Vehicle capacity

    ii) Maximum riding time

    iii) School time window

    iv) Maximum walking time or distance

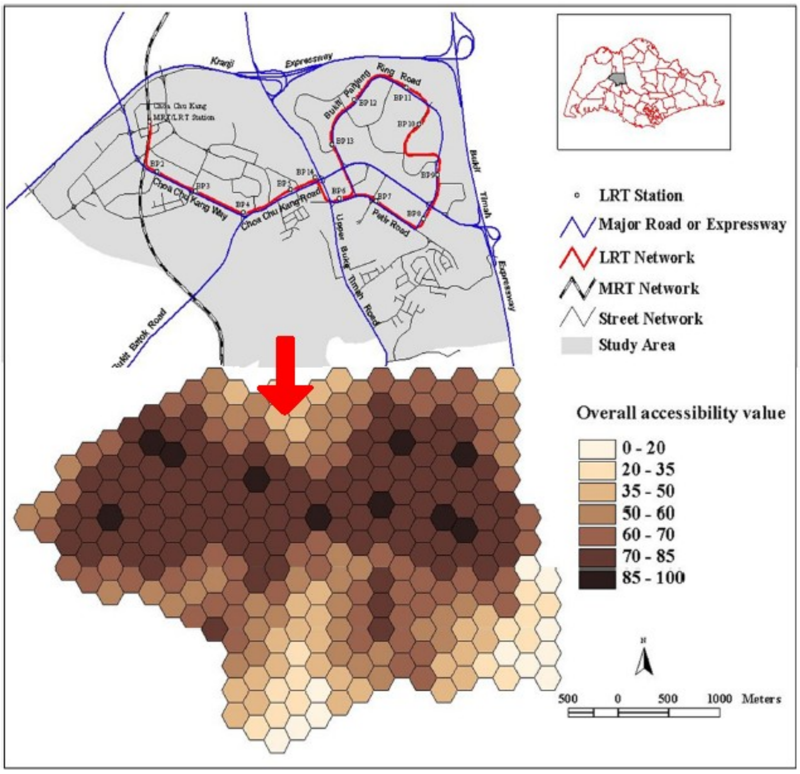
Above characteristics will be examined using the unique capabilities of the GIS, especially the network analyst engine.

2.7 Accessibility Concept

Accessibility is derived from word accessible, which according to the Oxford Dictionary means able to reach or be entered. Accessibility is related to opportunity, for example, the opportunity for an individual to participate in an activity. In the transportation area, accessibility represents the ease of reaching activities from a given location by using a particular transportation system (Zhu and Liu 2004). An accessibility measure is usually formulated in a set of destinations and a set of origins. The destination represents the activity places and the origin represents the potential users. At the activity sites there are facilities that will be used by the potential users. Individual destinations may be weighted according to their attractiveness. Their appeal can be generated by a combination of factors which might influence or pique the interest of the customers. For example, in the shopping centre, the weighted value can be measured by the value of floor space and parking space. In the case of a school, the weighted value can be measured by the maximum number of pupils and the range and quality of school services provided.

A wide range of application and research has used the accessibility factor. In measuring the level of service provision in the rural areas of Scotland, Halden used accessibility as one of the criteria (Halden et al 2002). The travel time and the travel pattern to the urban centers, shopping opportunities, and to regional health care facilities are examined and analyzed. Xuan Zhu et al proposed using GIS for discovering the accessibility of housing into public transport such as MRT and bus routes, shopping centres, health care services, etc (Xuan Zhu et al 2006). Figure 2.7 show the input and output of this accessibility research.

Accessibility also has a portion in the public transport and its stopping points. Jennifer Rogalsy calculated the quality of service of existing public transport for the working poor in Knoxville, Tennessee, USA (Rogalsky 2009). The work of Talat Munshi and Mark Brussel focuses on the accessibility of public transport in relation to different income rates and work activities in Ahmedabad city in India (Munshi & Brussel, 2006). Sankar et al use GIS to optimize accessibility and to determine the best placement of bus stops. Handy and Niemeier note that the gravity-based measure of accessibility is the most widely used measure in planning studies (Handy and Niemeier 1997). Thomas J Kimpel et al use this gravity-based measure for calculating overlapping service at existing bus stops (Kimpel et al, 2007). The more stops along the route increases accessibility because it boosts more walking distance areas and larger numbers of customers. On the other hand, more stops and greater access slow transit travel speeds and can reduce the area of service in certain time instances Greater access can actually result in greater service interruption and longer travel times. Allan T Murray and Xiaolan Wu have conducted research about accessibility to examine the trade-off between both sides of this dilemma (Murray & Wu, 2003).



(source: Xuan Zhu et al, 2006)

Figure 2.7 Example of the accessibility visualization

Impact of the new transport system in the accessibility level of the surrounding working population in Singapore has been revealed in the research of Xuan Zhu and Suxia Liu. They make a before and after accessibility map to show the effect of adding a new MRT route in the existing transportation system.

The output of this research is a proposed optimal route. This research will try to to analyse accessibility by making a before and after accessibility map similar to that presented by Xuan Zhu and Suxia Liu. The accessibility will be calculated by distance to the route and weighted by the available seats of transportation media.

2.8 Network Analyst Technology

The graph used in the vehicle routing field has a similarity with the graph used in the GIS application. Peter Keenan notes GIS approaches allow the modeling of an extended range of routing problems (Keenan, 2008). The synthesis of vehicle routing and GIS techniques in a spatial decision support system can significantly enhance the modeling of these problems.

ArcGIS is one of the common global GIS applications used. It has many functions in many areas of GIS, and among them is vehicle routing. The Vehicle Routing Problem function is one of the functions bundled in network analysis extension. It has a close relationship with the other analysis functions in the bundling extension.

In GIS there are two kinds of networks; they are transportation networks and utility networks. Transportation networks are undirected networks, meaning that although an edge on a network may have a direction assigned to it, the agent (the person or resource being transported) is free to decide the direction, speed, and destination of traversal. For example, a person in a car traveling on a street can choose which street to turn onto, when to stop, and which direction to drive. Restrictions imposed on a network, such as one-way streets or "no U-turn allowed", are guidelines for the agent to follow. This is in stark contrast to the utility network. A utility network is directed. This means the agent (for example, water, sewage, or electricity) flows along the network based upon certain rules built into the network. The path that the water will take is predetermined. It can be changed, but not by the agent. The engineer controlling the network can change the rules of the network and alter direction by opening some valves and closing others.

2.8.1 Network Dataset

The transportation network needs a special model of data. It is called network dataset (ESRI, 2008). A network dataset is created from the feature source or sources that participate in the network. It incorporates an advanced connectivity model that can represent complex scenarios such as multimodal transportation networks. It also possesses a rich network attribute model that helps model impedances, restrictions, and hierarchy for the network. The network dataset is built from simple features (lines and points) and turns.

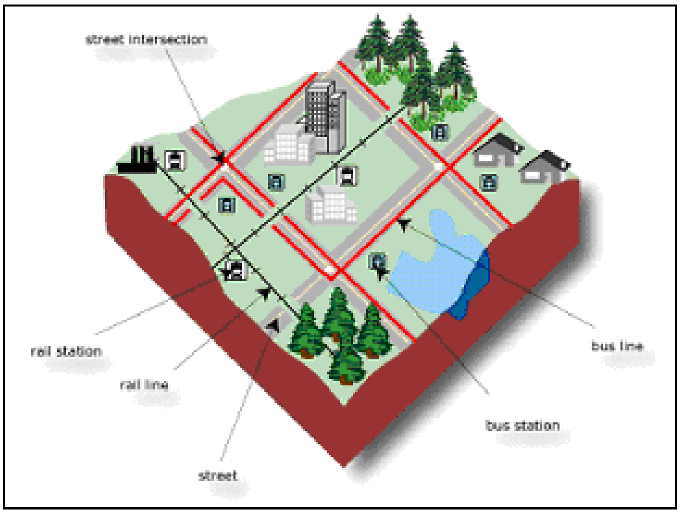
Network datasets are the presentation of network elements in Esri standard. Network elements are built from the characteristics of the road network. Street features is the common use for the main sources. The geometry of the street features is used to establish connectivity. Network elements also have attributes that are responsible for controlling navigation over the network. There are three parts in the network elements which include edges, junctions, and turns. Edges are elements responsible in connection and connect to other elements. Edges also become the links over which resources flow. Junctions connect edges and they also facilitate navigation from one edge to another. Connectivity in a network deals with connecting edges and junctions to each other. Turn elements are responsible for the movement between two or more edges. Turns are optional. This element stores information about the turning movement; for instance, a right turn is restricted from one particular edge to another in a particular junction.

There are three types of network sources that participate in the creation of a network dataset: edge feature sources, junction feature sources, and turn feature sources. Line feature classes participate as edge feature sources. Point feature classes participate as junction feature sources. Turn feature classes participate as turn feature sources in a network. A turn feature source explicitly models a subset of possible transitions between edge elements during navigation.

Each feature class that participates in a network as a source generates elements based on its assigned role. For example, a line feature class is used as a source for edge elements, and a point feature class is used to generate junction elements. Turn elements are created from a turn feature class. The generated junction, edge, and turn elements form the underlying graph which is the network.

Consider the example of a simple transportation network and the sources that participate in its creation. This network has a streets feature class that can act as an edge source, a street intersections feature class that acts as a junction source, additional line feature classes that act as edges (rail lines, bus routes), and point feature classes that act as junctions (rail stations and bus stations).

The transportation network has some attributes. Network attributes are properties of the network elements that control traversability over the network. Examples of attributes include the time to travel a given length of road, which streets are restricted for which vehicles, the speeds along a given road, and which streets are one-way.



(source: http://webhelp.esri.com)

Figure 2.8. Example of a simple transportation network

Network attributes have a usage type. The usage type specifies how the attribute will be used during analysis, which is identified either as a cost, descriptor, restriction, or hierarchy. Certain attributes are used to measure and to model the cost factor in VRP field. The examples for these kinds of attributes are travel time and the demand value. These attributes are apportioned along an edge. They are divided proportionately along the length of the edge. For example, if travel time is modeled as a cost attribute, then traversing half an edge will take half the time as traversing the whole edge. This means that if the travel time to traverse the edge is 3 minutes, it takes 1.5 minutes to traverse half the edge. If we are looking for a 1.5-minute route along this edge, the route feature will be created from the first half of the edge feature.

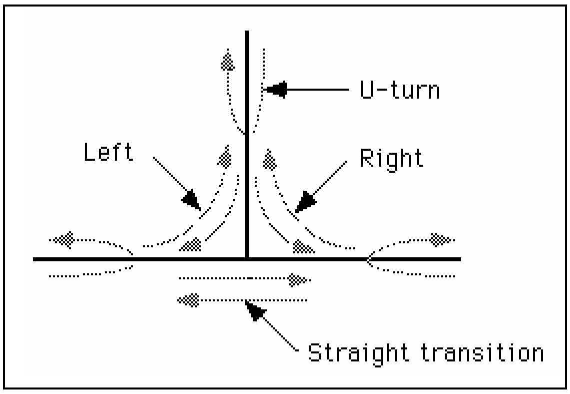
Network analysis often involves the minimization of a cost (also known as impedance) during the calculation of a path (also known as finding the best route). Common examples include finding the fastest route (minimizing travel time) or the shortest route (minimizing distance). Travel time (drive time, pedestrian time) and distance (meters) are also cost attributes of the network dataset.

Descriptors are attributes representing the characteristics of the network elements. The nature of the descriptors is that they are not apportioned. The value of a descriptor does not depend on the length of the edge element. In a street network, some examples of descriptors are speed limit, street width, and number of lanes. Although it is not a cost attribute, it can cooperate with the distance value to create cost attribute. One of the common examples is drive time, which is derived from length of the segment road in relation to driving speed.

Particular elements can have Restriction. The most common is the restriction in the street, it cannot be traversed. One-way streets can also be modeled with a restriction attribute. The restrictions can apply to one direction, so they can only be traversed from one end to another and not in the reverse direction. A restriction attribute is defined using a Boolean data type. In another example, certain sources where pedestrians are not allowed could be restricted using the attribute No Pedestrians. In this case, the restriction can be used as a parameter during best route analysis to ensure the pedestrian does not use streets that are restricted.

Hierarchy is the concept of ordering or grading the network elements. The commonly used this concept is the street network. A street network have an interstate class, highway class, local class, and so on. The hierarchy can also be used to model the network with preferences to take (or not to take) a particular street class, such as interstate, in finding a shortest path from one point to another.

The different turning characteristic can be modeled in the Turns. Turns can be made at any junction where edges connect, even at a junction with a single edge, since it is possible to make one U-turn. In a quotation, there are n^2 possible turns at every network junction, where n is the number of edges connected at that junction.



(source: http://webhelp.esri.com)

Figure 2.9. Combination of Turn

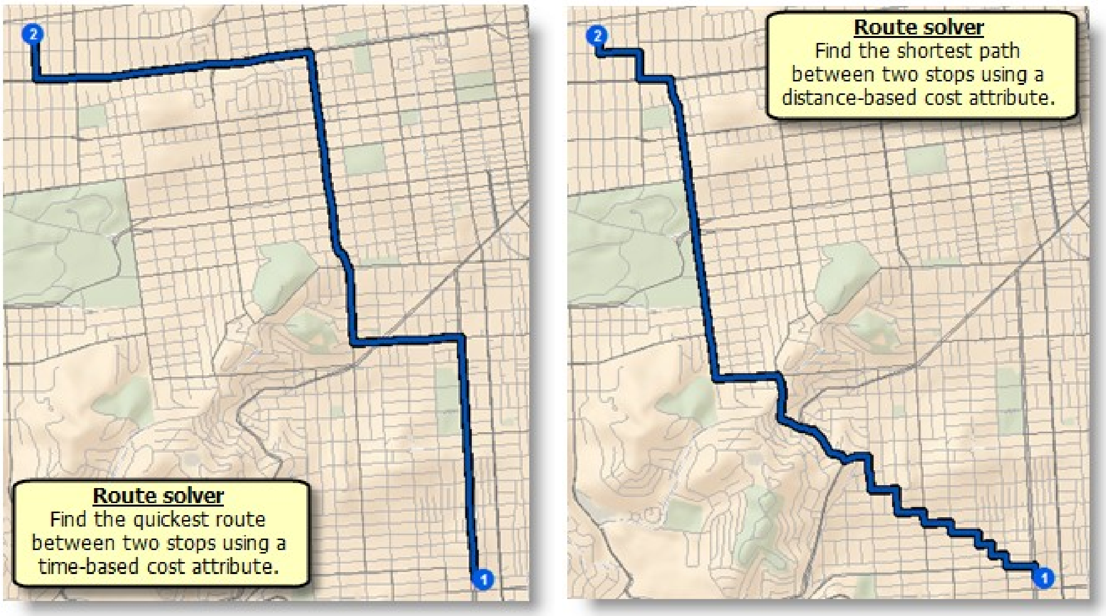
2.8.2 Functions in Network Analyst

Network Analyst functions in ArcGis 9.3 allows us to solve more common network problems. That version completed the previous set of functions with the VRP function. The previous function, people can find the best route across a city, finding the closest emergency vehicle or facility, identifying a service area around a location, and creating OD cost matrix. Now people can also deals with servicing a set of orders with a fleet of vehicles (ESRI, 2008).

a. Finding the best route

This first function of ArcGIS Network Analyst one can use it to find the best way to travel from one location to another. It also can apply to several locations in a row. The locations are added in 3 ways: interactively placing points on the screen, entering an address, and by importing an existing feature class or feature layer. The order of the locations to be visited is determined by user and the shortest route. ArcGIS Network Analyst can also suggest the best sequence to visit all locations.

People usually try to take the best route, whether finding a simple route between two locations or one that visits several locations. However, best route can mean different things in different situations. The best route can be the quickest, shortest, or most scenic route. It depends on the chosen impedance. The impedance used is chosen by user. For example, let's say the impedance is time. The best route is the quickest route. When impedance is cost, the best route is the one with minimum cost. In the example below, the first case uses time as an impedance. The quickest path is shown in blue, and has a total length of 4.5 miles which takes 8 minutes to traverse. In the second case, distance is chosen as the impedance. Consequently, the length of the shortest path is 4.4 miles, which takes 9 minutes to traverse.

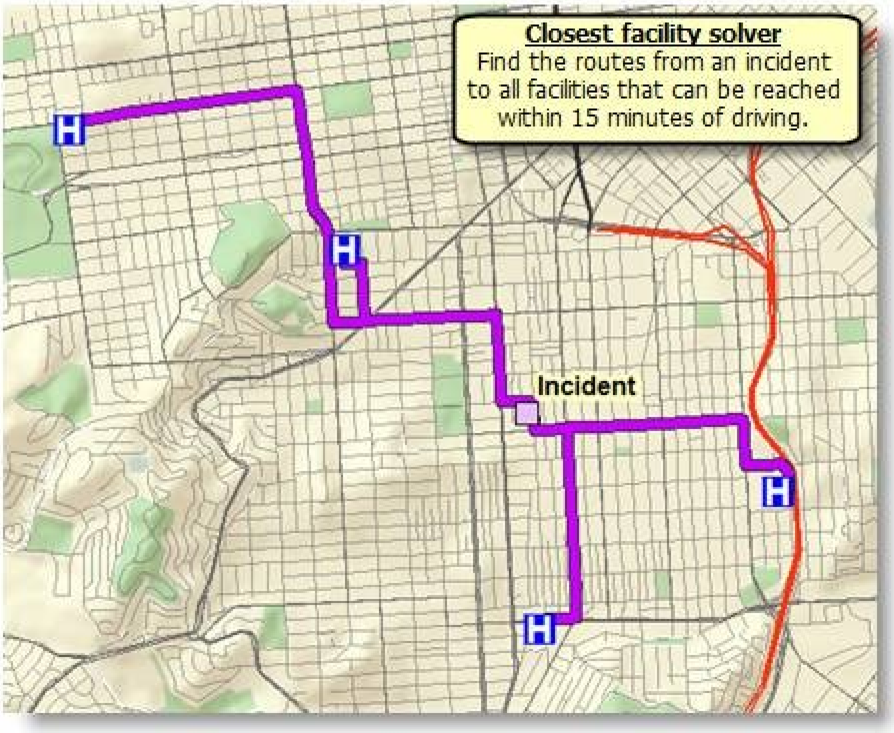


(source: http://webhelp.esri.com)

Figure 2.10. Two different base attribute routing: a. time based, b. distance based

b. Finding the closest facility

The second function is finding the closest facility. This function can used for finding the closest police cars to a crime scene, finding the closest hospital to an accident, or finding the closest ATM machine from a given location. Once we've found the closest facilities, we can display the best route to or from them, display directions to each facility, and return the travel cost for each route. We can also specify an impedance cutoff, so the Network Analyst can stop the route searching after a particular value is met. For example, we can set up a closest facility problem to search for ATM machine within 5 minutes' drive time from our location. Any ATM machines that need more than 5 minutes to reach will not be included in the results. Another example is searching hospital that can reach within 15 minutes of driving as shown in following figure.



(source: http://webhelp.esri.com)

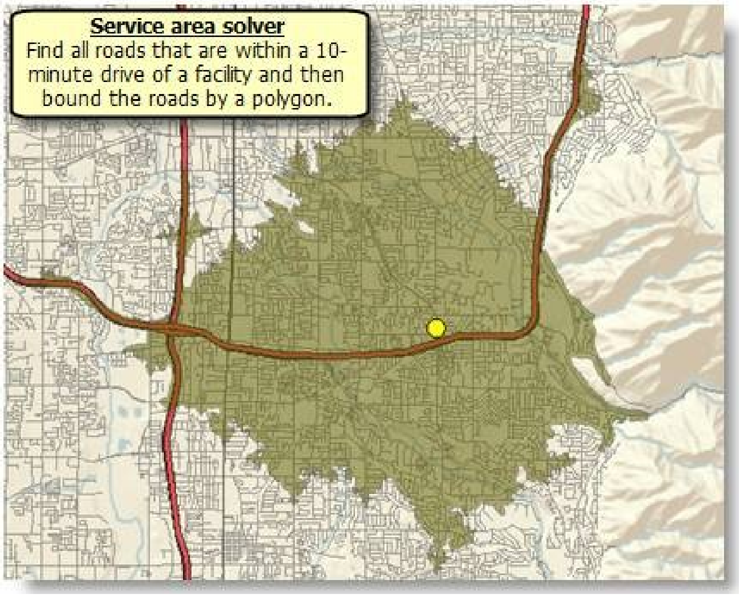
Figure 2.11. Examples output of Closest facility solver

The hospitals are referred to as facilities, and the accident is referred to as an incident. ArcGIS Network Analyst allows performing multiple closest facility analyses simultaneously. This means we can have multiple incidents and find the closest facility or facilities to each incident.

This function is used by McAbee et al in their research (McAbee et al 2006). They develop a model that provides unique results across geographic space to improve the recovery of the blackside dace (Phoxinus cumberlandensis). This function is used for calculating the shortest stream path between two populations and also the distance. It needs to determine one population as a Facility and the second population as an Incident. Solving for this route provides the shortest stream path between two populations.

c. Finding service areas

The next function is like reversing the use of closest facility function. This time the input parameter is the time and the output is every possible location that can be reached within that time or less. It means we can get service areas around any location on a network. The output is in the form of a region and called a network service area region. A 10-minute service area for a facility, for example, will have a collection of streets segments that can be reached within ten minutes from that facility.



(source: http://webhelp.esri.com)

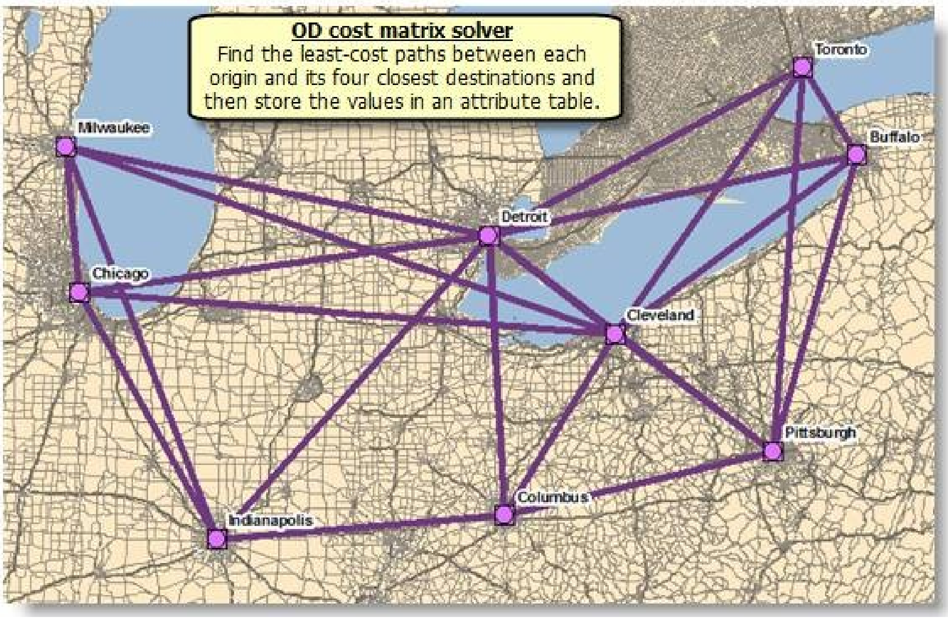
Figure 2.12. Example output of Service Area Solver

This function has been used in the research of John Cullinan et al (Cullinan et al 2008) in estimating catchments area population indicators. The spatial coordinates of every building in the Republic of Ireland were used, and coordinated with estimated service areas. The results were used to calculate the number of residential buildings within a catchment area. The number of residences within the catchments were then analyzed. They act as an indicator or proxy for its population. The finding services areas" can be useful when the catchments area does not match closely with any combination of geographical or administrative areas for which data exists.

d. Creating an OD cost matrix

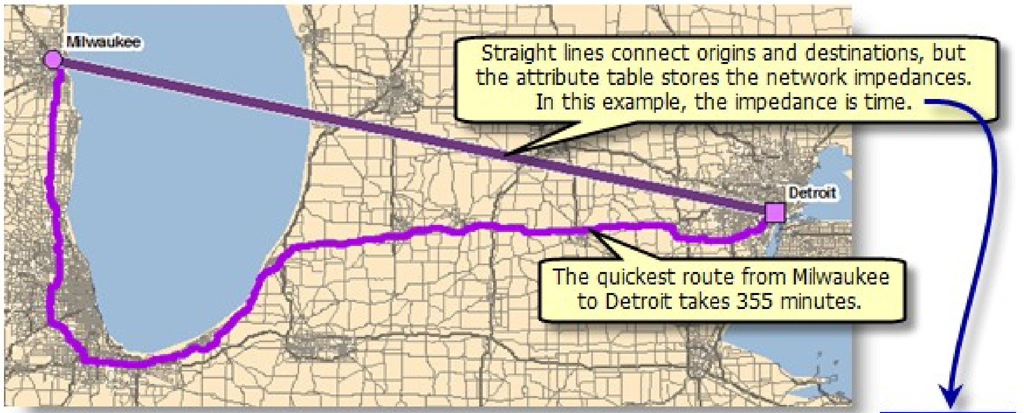
An OD cost matrix is a table that contains the network impedance from each origin to each destination. This table makes the characteristics of the network among each origin to each closest destination easy to figure out. The function in network analyst can create an origin–destination (OD) cost matrix from multiple origins to multiple destinations. It also provides the ranks of the destinations that each origin connects to. The rank is in ascending order based on the minimum network impedance required to travel from that origin to each destination.

The best network path is discovered for each origin-destination pair, and the cost is stored in the attribute table of the output lines, as shown in the straight lines below. The graphic below shows the results of an OD cost matrix analysis that was set to find the cost to reach the four closest destinations from each origin.



(source: http://webhelp.esri.com)

Figure 2.13. Example of OD Cost matrix using



	ObjectID	Shape	Name	OriginID	DestinationID	DestinationRank	Total_Time
	133	Polyline	Buffalo - Detroit	5	2	5	252
	134	Polyline	Milwaukee - Milwaukee	6	6	1	0
	135	Polyline	Milwaukee - Chicago	6	1	2	98
	136	Polyline	Milwaukee - Indianapolis	6	7	3	265
	137	Polyline	Milwaukee - Detroit	6	2	4	355
	138	Polyline	Milwaukee - Cleveland	6	3	5	419
	139	Polyline	Indianapolis - Indianapolis	7	7	1	0
	[illegible]	[illegible]	Indianapo[illegible] Columbus	[illegible]	9	2	[illegible]

(source: http://webhelp.esri.com)

Figure 2.14. Example of storage data in OD Cost matrix

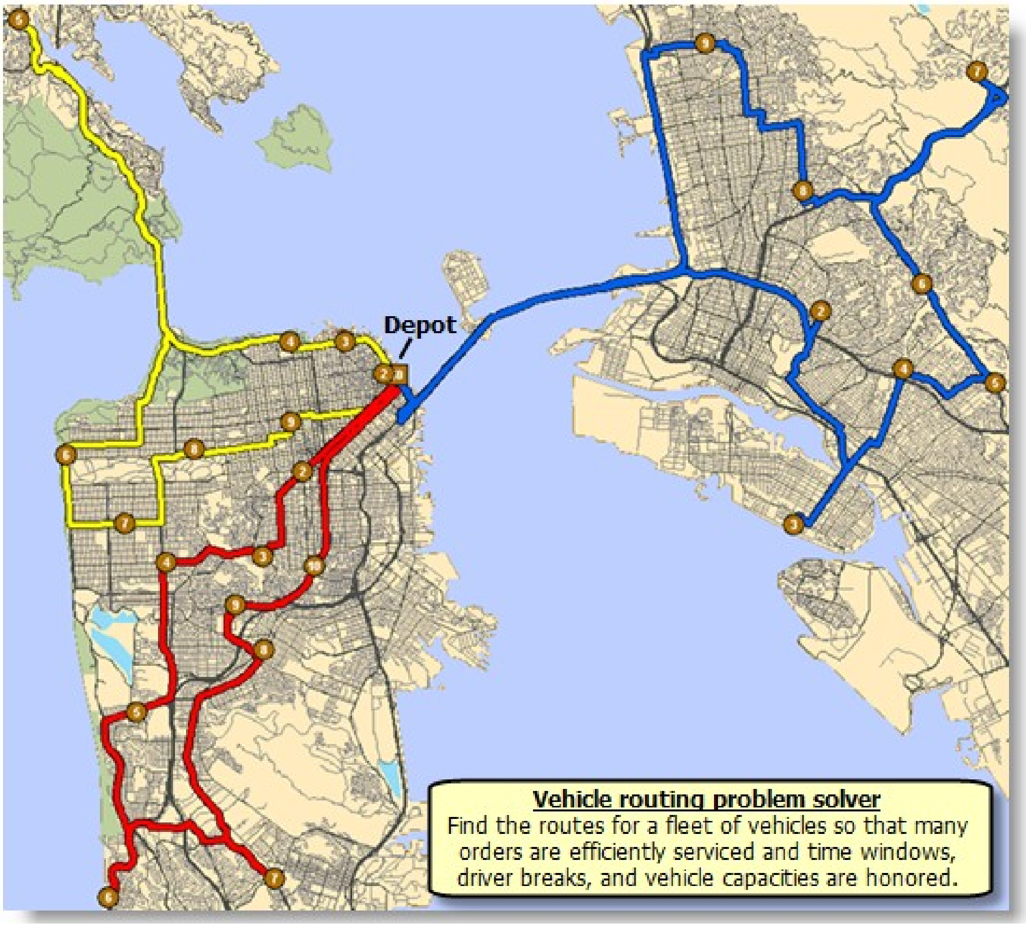
The straight lines can be symbolized in various ways, such as by color, representing which point they originate from, or by thickness, representing the travel time of each path. The closest facility and OD cost matrix solvers perform very similar analyses; the main difference, however, is in the output and the computation speed. The OD cost matrix solver is designed for quickly solving large MxN problems and as a result does not internally contain information that can be used to generate true shapes of routes and driving directions. If one requires driving directions or true shapes of routes, ideally the closest facility solver should be used, or otherwise the OD cost matrix solver can be used to reduce the computation time.

A research study by Cavalli et al had used these tools for finding a successful strategy related to a dwarf pine reclamation project in a an area located in Altopiano dei Sette Comuni in the northeastern part of Italy. In this work a strategy to contain the spreading of this species along the forest road network is analyzed. The analysis considers the operative costs for both the felling and transport of dwarf pine stems to intermediate landings. The logistic problem has been widely analyzed by an OD Cost matrix and assuming different scenarios considering different levels of upgrade to the forest road network (Cavalli et al 2010).

e. Solving a vehicle routing problem

A dispatcher managing a fleet of vehicles is often required to make decisions about vehicle routing. One such decision involves how to best assign a group of customers to a fleet of vehicles and to sequence and schedule their visits. The objectives in solving such vehicle routing problems (VRP) are to provide a high level of customer service by honoring any time windows, while keeping the overall operating and investment costs for each route as low as possible. The constraints are to complete the routes with available resources and within the time limits imposed by driver work shifts, driving speeds, and customer commitments.

ArcGIS Network Analyst provides a vehicle routing problem solver that can be used to determine solutions for such complex fleet management tasks. Consider an example of delivering goods to grocery stores from a central warehouse location. In this imagined scenario, a fleet of three trucks is available at the warehouse. The warehouse operates only within a certain time window—from 8 A.M. to 5 P.M.—during which all trucks must return back to the warehouse. Each truck has a capacity of 15,000 pounds, which limits the amount of goods it can carry. Each store has a demand for a specific amount of goods (in pounds) that needs to be delivered, and each store has time windows that restrict when deliveries should be made. Furthermore, the driver can work only eight hours per day, requires a break for lunch, and is paid for the amount spent on driving and servicing the stores. The goal is to come up with an itinerary for each driver (or route) such that the deliveries can be made while honoring all the service requirements and minimizing the total time spent on a particular route by the driver. The figure below shows three routes obtained by solving the above vehicle routing problem. This important function will be used in this research which will ultimately facilitate finding the optimal school bus route.



(source: http://webhelp.esri.com)

Figure 2.15. Example output of Vehicle Routing Problem solver

2.8.3 Algorithms used by Network Analyst

The routing solvers within Network Analyst—namely the Route, Closest Facility, and OD Cost Matrix solvers—are based on the well-known Dijkstra's algorithm for finding shortest paths (ESRI 2008). Each of these three solvers implements two types of path-finding algorithms. The first type is the exact shortest path, and the second is a hierarchical path solver for faster performance. The classic Dijkstra's algorithm solves a shortest-path problem on an undirected, nonnegative weighted graph. To use it within the context of real-world transportation data, this algorithm is modified to respect user settings such as one-way restrictions, turn restrictions, junction impedance, barriers, and side-of-street constraints, while minimizing a user-specified cost attribute. The performance of Dijkstra's algorithm is further improved by using better data structures such as d-heaps. In addition, the algorithm needs to be able to model the locations anywhere along an edge, not just on junctions.

The classic Dijkstra's algorithm solves the single-source, shortest-path problem on a weighted graph. To find a shortest path from a starting location ***s*** to a destination location ***d***, Dijkstra's algorithm maintains a set of junctions, ***S***, whose final shortest path from ***s*** has already been computed. The algorithm repeatedly finds a junction in the set of junctions that has the minimum shortest-path estimate, adds it to the set of junctions ***S***, and updates the shortest-path estimates of all neighbors of this junction that are not in ***S***. The algorithm continues until the destination junction is added to ***S***.

These network analyst functions use Dijkstra's in different views:

a) Route

This uses the well-known Dijkstra's algorithm, which is described above.

b) Closest facility

This uses a multiple-origin, multiple-destination algorithm based on Dijkstra's algorithm. It has options to only compute the shortest paths if they are within a specified cutoff or to solve for a fixed number of closest facilities.

c) OD cost matrix

This uses a multiple-origin, multiple-destination algorithm based on Dijkstra's algorithm. It has options to only compute the shortest paths if they are within a specified cutoff or to solve for a fixed number of closest destinations. The OD

Cost Matrix solver is similar to the Closest Facility solver but differs in that it does not compute the shape of the resulting shortest path for less overhead and faster performance.

d) Hierarchical routing

Finding the exact shortest path on a nationwide network dataset is time-consuming due to the large number of edges that need to be searched. To improve performance, network datasets can model the natural hierarchy in a transportation system where driving on an interstate highway is preferable to driving on local roads. Up to three levels of hierarchy can be supported by the network dataset. Once a hierarchical network has been created, a modification of the bidirectional Dijkstra is used to compute a route between an origin and a destination.

The overall objective presented here is to minimize the impedance while favoring the higher-order hierarchies present in the network. This is accomplished by simultaneously searching from both origin and destination locations as well as connection or entry points into higher-level roads, then searching the higher-level roads until segments from both origin and destination meet. As the search is restricted to the upper hierarchy, a smaller number of edges are searched, resulting in faster performance. For this heuristic to be successful, the top-level hierarchy must be connected, as it will not descend to a lower level if a dead end is reached.

e) Traveling salesman problem option for the Route solver

The Route solver has the option to generate the optimal sequence of visiting the stop locations. This is the traveling salesman problem (TSP). The TSP is a combinatorial problem, meaning there is no straightforward way to find the best sequence. Heuristics are used to find good solutions to these types of problems in a short amount of time. The TSP implementation within ArcGIS Network Analyst also handles time windows on the stops; that is, it finds the optimal sequence to visit the stops with a minimum amount of lateness.

The traveling salesman solver starts by generating an origin-destination cost matrix between all the stops to be sequenced and uses a tabu search-based algorithm to find the best sequence of visiting the stops. Tabu search is a metaheuristic algorithm for solving combinatorial problems. It falls in the realm of local search algorithms. The exact implementation of the tabu search is proprietary, but it has been researched and developed extensively in-house at ESRI to quickly yield good results.

f) Vehicle routing problem with time windows

The vehicle routing problem (VRP) is a superset of the traveling salesman problem. In a TSP, one set of stops is sequenced in an optimal fashion. In a VRP, a set of orders needs to be assigned to a set of routes or vehicles such that the overall path cost is minimized. It also needs to honor real-world constraints including vehicle capacities, delivery time windows, and driver specialties. The VRP produces a solution that honors those constraints while minimizing an objective function composed of operating costs and user preferences, such as the importance of meeting time windows.

The VRP solver starts by generating an origin-destination matrix of shortest-path costs between all order and depot locations along the network. Using this cost matrix, it constructs an initial solution by inserting the orders one at a time onto the most appropriate route. The initial solution is then improved upon by re-sequencing the orders on each route, as well as moving orders from one route to another, and exchanging orders between routes.

g) **Service area**

The Service Area solver is also based on Dijkstra's algorithm to traverse the network. Its goal is to return a subset of connected edge features such that they are within the specified network distance or cost cutoff; in addition, it can return the lines categorized by a set of break values an edge may fall within. The service area solver can generate lines, polygons surrounding these lines, or both. The polygons are generated by putting the geometry of the lines traversed by the Service Area solver into a triangulated irregular network (TIN) data structure. The network distance along the lines serves as the height of the locations inside the TIN. Locations not traversed by the service area are put in with a much larger height value. A polygon generation routine is used with this TIN to carve out regions encompassing areas in between the specified break values.

2.9 Summary

This chapter presents key information and background knowledge that one needs to know before beginning research in school bus routing. Sections in this chapter are arranged following the order of the used terms in the title of this research. In the first section, the description about school bus, the types and each type specification is introduced. This section is then followed with a discussion of characteristics related to the identification of the needy family. Needy family specification in Surabaya is explained with a sample case. The next section is focused on finding the route. It begins with an explanation of VRP, continues with several research studies in VRP area, and ends with a detailed focus on the several different types of VRP. Several school bus research examples are also explained and classified into different characteristics. The position of this research in the VRP area and school bus area then explained. In order to prepare the analysis process, a survey about accessibility research has been conducted and some related ones have been selected and introduced in the next section. Finally, this chapter concludes with an explanation about the engine that will be used in this research; the network analysis engine. The explanation is focused in the concept, data, and the functions.