CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The purpose of this chapter is to understand the context area of this research. Firstly, this chapter will explain features of the conceptual framework as they relate to finding and analyzing the school bus route. Secondly, this chapter explains about the parameter used in this research process. Next, this chapter details the concept more technically in a research methodology. The last section will introduce the geospatial technology that will be used in this research.

3.2 Conceptual Framework of School Bus Route Finding and Route Analysis

This section will explain the concept of school bus route finding and route analysis with a simple example. The key to the process of finding a route for this free school bus is how to find one optimal route with a coverage area that will have the most number of needy children. Then continue with analyzing the chosen route. In a concept, the process is a continuation of smaller steps. These steps begin with preparing and adding the data to the developed model. The initial data needed in the model are needy layer data, street network, school location, and depot location. The model then will refining the needy layer, considering the needy around school is not need to use the school bus and directly adjusted to attend to the concerned school. Next process is finding the possible route and calculating the number of needy that discovering which is being the best route, with the most number of coverage needy. This most optimal route then analyzed. These steps are shown in Figure 3.1 below. Next paragraphs will details each step.

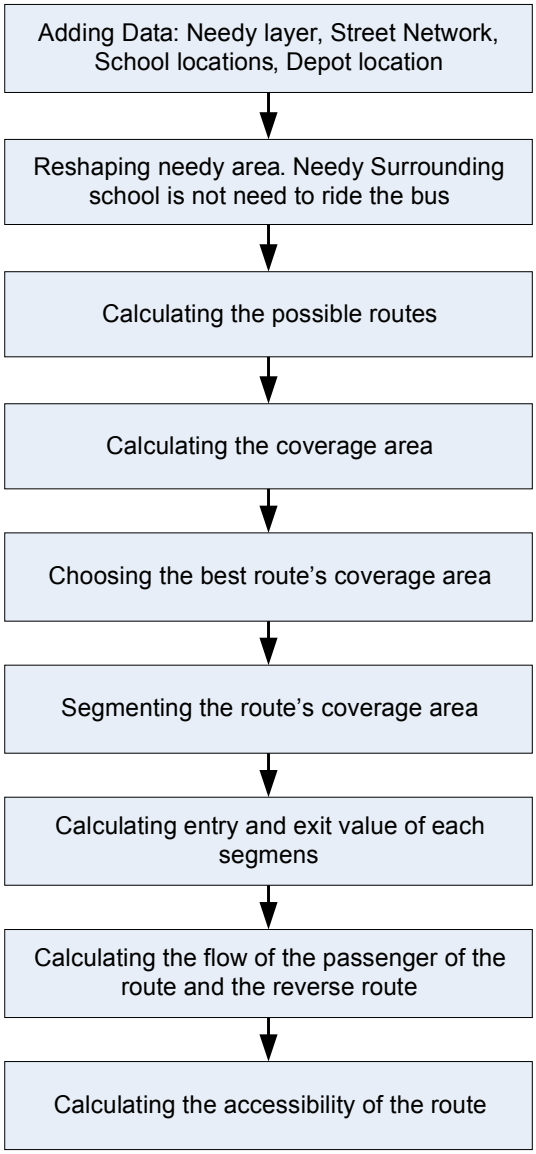


Figure 3.1. Conceptual framework diagram

Coverage area is the area surrounding the route. Needy children in this area will need to be calculated to determine the total number of needy. It will need a needy layer, or layer that represents needy number in a specific area distribution. Figure 3.2 is the example of this needy layer.

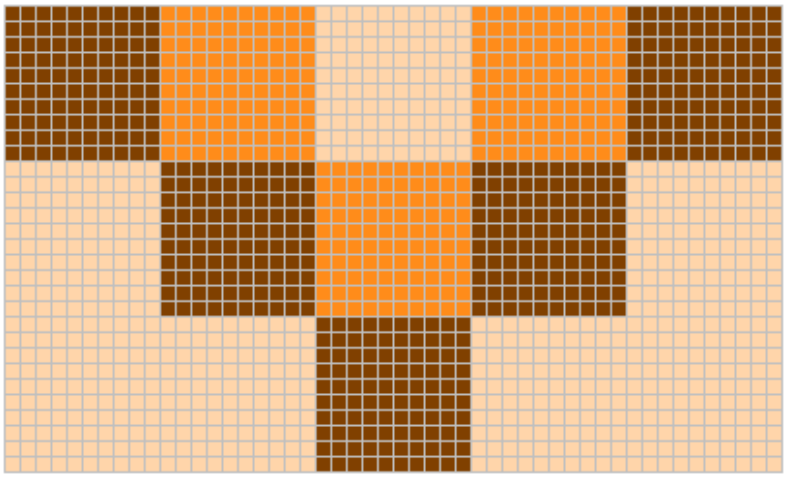


Figure 3.2. Needy area layer (example).

In order to keep the process easy to follow, this explanation uses a simple example map layer. In the example needy area layer, there are 5x3 square district areas. Each district has a number of needy. The number of needy is represented in little squares in the district square. There are 3 different colors in the little squares. For example, let's assume the darkest color represents 3 needy students, the lighter color represents 2 needy students, and the lightest color represents 1 needy student.

The other layers we need to provide include the street layer, school layer, and bus depot layer. Figure 3.3 is the example of these 3 layers. Street layer is represented in red line. School layer is represented in green square (green point). Bus depot layer is represented in blue square (blue point). The school layer has a specific capacity of student. The street layer is segmented with the end of each segment is the junctions. Each segment has its own drive time value.

Figure 3.3 Street, School, and Depot layer (example)

The school bus is used by the needy student to take them to the school. However, students who live around their school do not have to ride the bus. This situation will decrease the number of the buses which will be provided. So, we need to calculate how many needy are surrounding the schools and decreasing the number of needy in those areas. A specific distance needs to be set and a specific percentage of needy whose school is nearby needs to be determined. The output of this process will result in two changes with the first being the reduction in the capacity of the school because some seats have been filled in by the surrounding needy students. Secondly, because the number of needy in the needy area around the schools is diminished, the needy area will be reshaped. Figure 3.4 is the example output of this process.

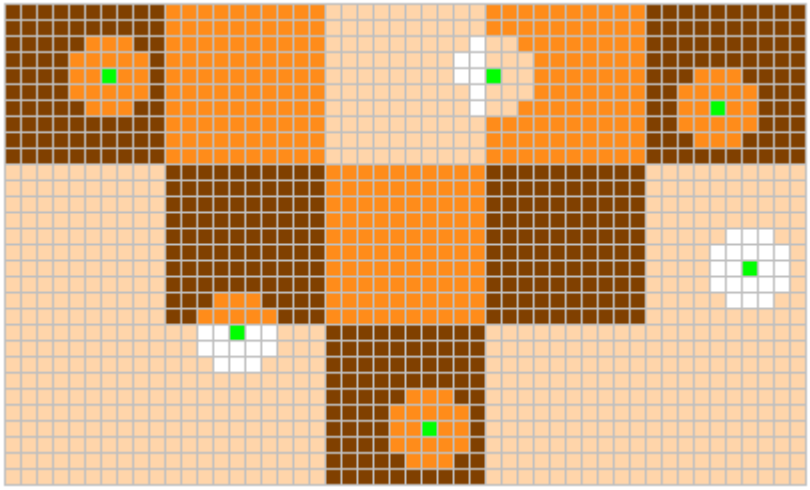


Figure 3.4 School layer and the reshaped needy area

The next process is finding the route. But first, take a look the all layers in one view as shown in Figure 3.5. The needy layer is placed in the lowermost. The street layer in the second layer and followed by school layer and bus depot layer. In GIS, it is called overlay.

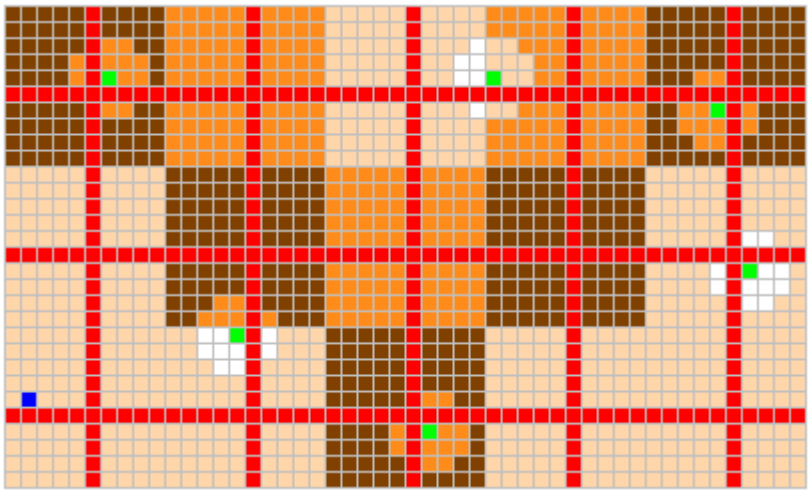


Figure 3.5 Street, School, and Depot layer overlay on the new needy layer

In the route finding process, start and end place of the school bus have to be determined. In this example, the one and only bus depot which is symbolized in blue point acts as both start depot and end depot. The school bus is starting the journey at start depot, visiting all school locations, and ending the journey at end depot. In this example case, the route will looks like a loop. By providing a certain algorithm, with an input including street network, school locations, and bus depot location, the process will generate an output in the form of a new polyline type layer. The routing process is done in several times in different settings. In this example, there are two different example of routing output. Along with the polyline type layer, the output was also provided with drive consumed time. Consumed time is how long the bus takes from beginning to end of the journey. The value is taken from summing all the drive time in all passed road segments and adding with the visiting time of the bus in the school location and in every bus stop. Figure 3.6 shows these two example routing process output.

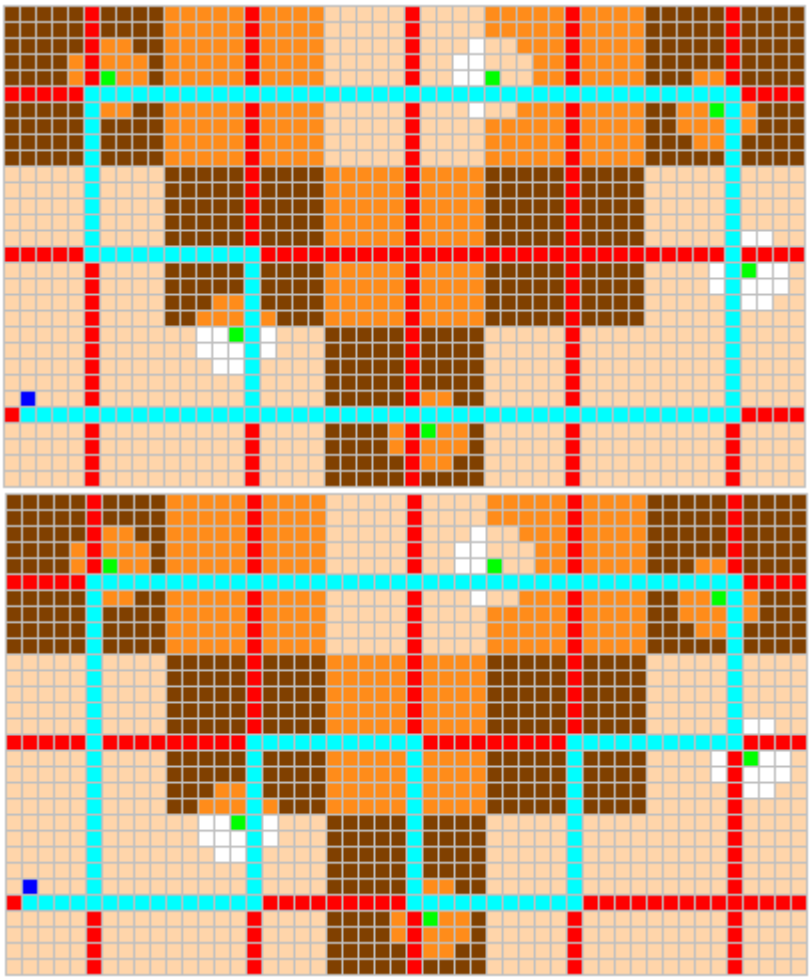


Figure 3.6 Two different output of routing process (example)

In the above figure, route is symbolized in blue line. These two routes have about the same length. However, having same length does not always mean having the same consumed time. Since this research is for a free school bus, with the main objective being to service as much needy students as possible, the length of the route is omitted. The consumed time is also not the primary factor in choosing the route. As long as the time range is in an acceptable span, the candidate route will not be eliminated. The next process is calculating the number of needy covered by the route. This will require a certain distance value which represents the maximum acceptable walking distance a needy student must go when traveling from their home to the closest route. The process needed to make a surrounding area along the route is called Buffer to those familiar with Geographic Information System (GIS). After the buffering process, there will be an area in the form of a polygon layer type. This polygon then will be used

to get the region's needy area placed inside it. In GIS this is called Clipping. The clip process is like cookie dough and a mold. In this case, the cookie dough is the needy area and the mold is the polygon surrounding the route. Let's refer to it as the output coverage needy layer. Figure 3.7 shows the result of the coverage needy layer of the two example routes.

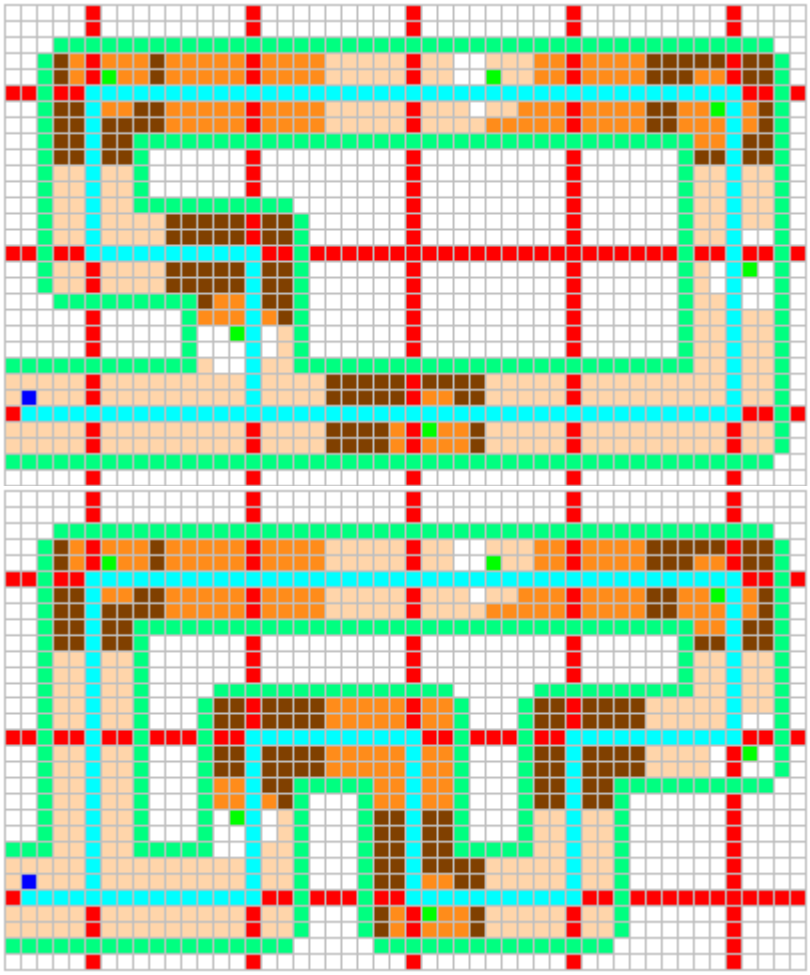


Figure 3.7 Two different coverage needy layer (example)

By counting the needy representation point (little squares), the total needy number in the coverage needy layer can be counted. The coverage of needy in the first route has the dark brown needy points, as much as 114, lighter brown point 99, and the lightest brown 254. By multiplying the value of one point as 3 for the dark brown, 2 for the lighter brown, and 1 for the lightest shade, the total number of needy in the coverage needy area is 794. On the other hand, the second coverage needy layer has 135 dark brown points, 140 lighter brown points, and 192 of the lightest spots. So, we can

reasonably conclude that the number of needy is 877. Furthermore, this figure demonstrates that the second route covers more number of needy student areas than the first route. Importantly if the second route has more consumed time than the first route, this research will still pick the second route for the solution route. After getting arriving at this solution, the process can now continue to the next stage which is the analysis stage.

Once the optimal route is determined, this research continues to explore the particular characteristics and benefits it offers to residents. The first area of exploration is determining the characteristics of the passenger load along the journey of the bus. The coverage needy layer from the previous section is used for this calculation. This layer needs to be segmented in a certain distance. Each segment will be calculated referring to the number of potential entry and exit of the passengers. The number of entry passengers is calculated from the total number of needy in the segment. The number of exit passengers is determined in reference to student capacity in the school of the targeted segment. If there is no school in the segment, this exit number is set to 0. In order simplify this example calculation and be consistent, all schools have been set to have 150 seats. With these entry and exit numbers, the load of passenger will be predicted. Figure 3.8 shows the chosen route with the entry and exit number.

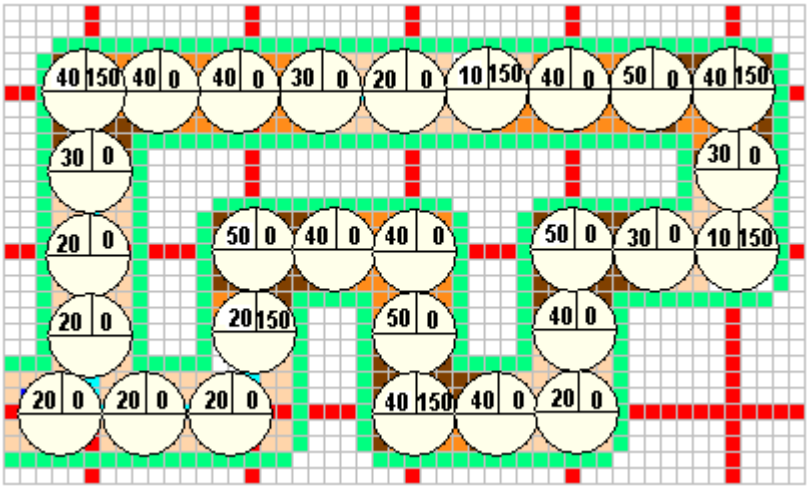


Figure 3.8 The entry and exit number in each segment (example)

Say the figure above is a map that has 4 directions; north is the top side, south is the bottom side, west is the left side, and east is the right side. The bus is doing its journey counter clockwise, so that it will travel east then make two weaves, go to the north, back to the west and finally end going to the south. In every visited segment the load of passengers is counted. The load of passenger value is added with the entry number and subtracted with the exit number. If the load of passenger is less than the exit number, the load will set to 0 and the segment is marked for use in the reverse route. Figure 3.9 above shows the calculation of each segment in the example route output.

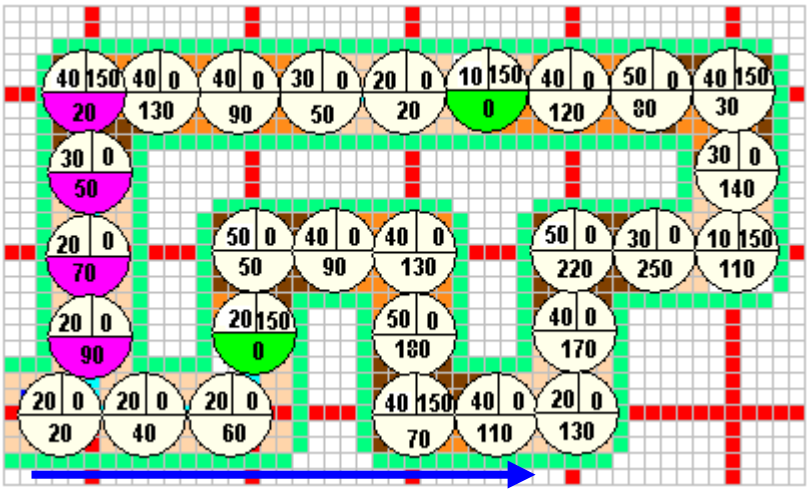


Figure 3.9. The calculation of passenger load (example)

The entry point is rounded for simplification purposes. The first segment has 20 as an entry value and 0 as an exit value, and this data also applies to the second and third segments. In the forth segment there is an exit value of 150. The passenger load in the second and third segment is 40 and 60 respectively. In the fourth segment the total passenger load is 0 because there is a school located there. The school capacity, also known as the exit value, is 70 seats more than the passenger value, therefore in this fourth segment the passenger load is set back to 0, and the segment is marked. In the above figure, this segment is marked with green color and the unused seats value is kept for the calculation in reserve route. The calculation in the next segment is formulated in a similar way. There are no unused seats in the next 4 school locations. In those locations, the passenger exceeds the school capacity. Some unused seat space is

generated in the fifth school, and the segment is colored with green again. After the last school, the sixth school, the segment is purple-colored indicating that passengers cannot make use of this route for the purpose of traveling to school. For passengers in these segments the reserve route is needed to conduct transportation services.

The reserve route needs an initial value of entry and exit passenger. The entry value is taken from the undelivered needy student, which in the previous figure is from the segments marked in purple. The exit value is taken from the unused seat, which in the previous figure is from the segments marked in green. Figure 3.10 shows the initial value for the reserve route.

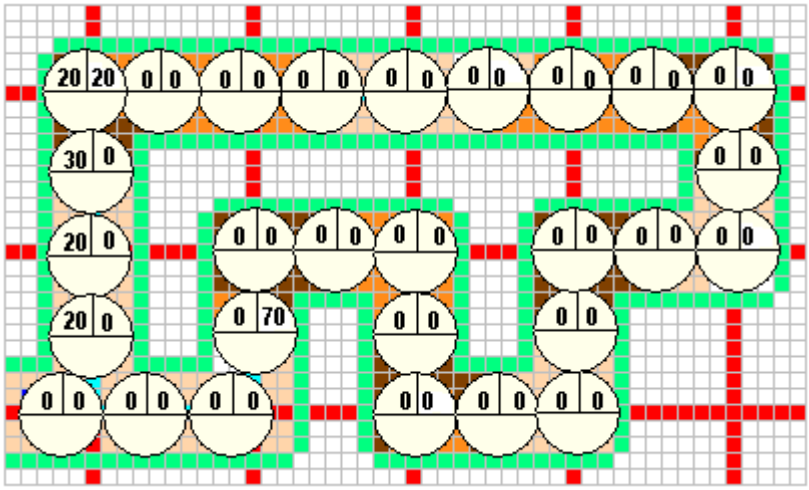


Figure 3.10 Initial entry and exit value for reserve road (example)

There are 2 school locations in this reserve route with capacity 20 and 70. The reserve route is running clockwise with the similar calculation method. So the potential passenger load is located at the beginning of the journey. There are 20, 20, 20, and 30 values in the first to fourth segments respectively. In the fourth segment, there is an exit value but less then the passenger load which as much as 20. The rest passenger load, as much as 70, is delivered until the last school which still has a capacity. Figure 3.11 shows this calculation. The maximum number of passengers in the main route is 250, while in the reverse route is 70. These numbers can used for calculating how many

numbers of buses need to be provided. If the bus capacity is 50 seats, the number of buses has to provide for the main route (5 buses and the reverse is 2 buses).

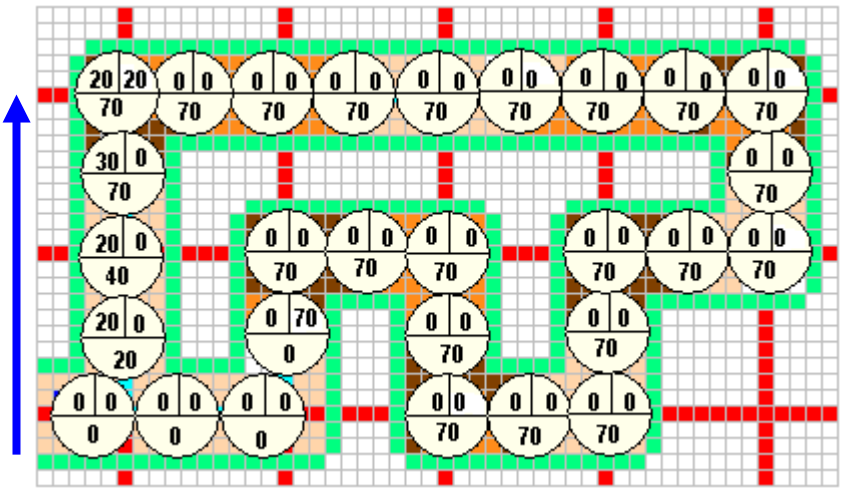


Figure 3.11. The calculation of passenger load in reserve route (example)

The number of buses will then be used in the second analysis; the accessibility analyst. The purpose of accessibility analyst is to discover improvements to the level of accessibility in the study area after a set of new school buses has been added to the existing transportation system. Therefore, this part will need an existing transport characteristic map. For an example, this section uses an existing transport map like that shown in figure 3.12 below.

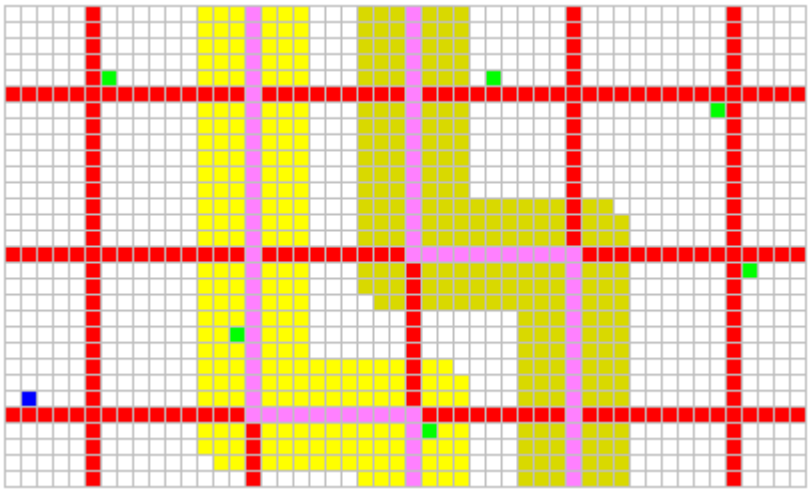


Figure 3.12 The existing transport system and it's accessibility layer (example)

The accessibility of existing transport system has to be determined. The accessibility layer is made from a buffer process similar to the process used in finding

the coverage area. The difference is that in this surrounding area, the value of the buffer is set based on the number of available seats in the transport media. In this example, the existing route on the east side has more available seats then the west, there the accessibility level of the right route is higher and represented in darker color. The transportation system is then added with a set of school buses.. This additional process is then followed with a recalculation of the accessibility level of all routes. The existing accessibility layer is added with accessibility of the bus route. Since this is an adding process, a street which has many transportation routes along it will have a high accessibility level. Figure 3.13 shows an example of new accessibility map. In this example map there are four different accessibility levels. The darker color shows the area shared street of the east existing route and the bus route.

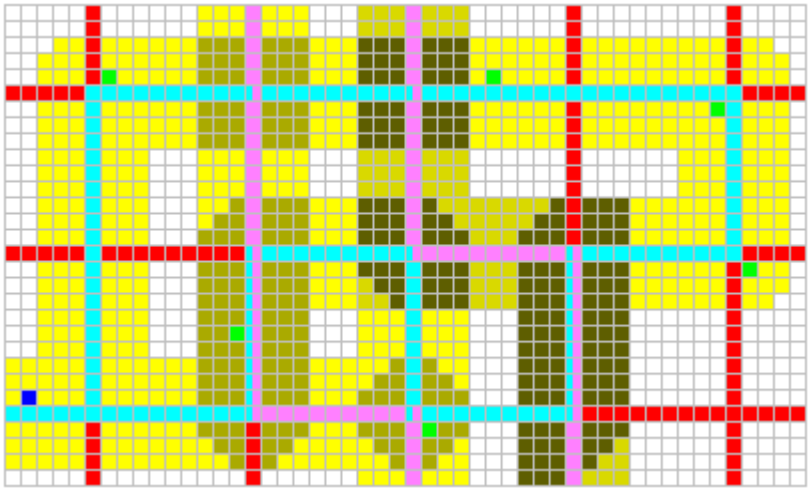


Figure 3.13 Accessibility after buses added (example)

3.3 **Variables and Notations**

In the developed model, there will be 3 types of variable. The first type is an Independent variable. Independent variable is some variable that changed several times when conducting experiment within the model. Second are Dependent variables which are observed in the experiment. The last is a Controlled variable. This is a variable which keep the same value while there are several changing of the independent variable.

Independent variables in that will be set in this model are :

a) Number of bus routes

   Number of bus routes will be setting up differently, starting from just one route and increasing until the number of routes can covering all schools location.

b) Start and End depot

   In each number of routes, there will be some experiment with different start and end depot. The result will determined with some method to get the best setting of Start and End depots.

c) The group of school

   In some experiment, the engine will supply with a group of school that have to be placed in a certain route. There will be different combination of the school group so it can make a comparison an choosing the best grouping school.

The different value of above Independent variable will generate a different value also in the dependent variable. The dependent variables in this model are:

a) The generated route

   The route is the output of the routing engine. With different number of routes, different start and end depots, and different of orders group (school group), the generated route will be different.

b) The reliability of the route

   Beside the number of the covered needy, the route also has to be reliable. The number of the school capacity along the route must be in the same level with the number of the needy covered along the route.

The different setting of above independents variables and the generated output is restricted with the value of controlled variable. The controlled variable in this model is the time limit. Whatever the number of routes, the different combination of start and end depots and the school groups, the bus route to be generated is limited in a certain

time limit. The bus must be operated within this time limit while driving from its start depot and its end depot.

In finding the optimal route, there is several calculation need to conduct. The calculation is conduct after the vehicle routing finished by the ArcGis VRP engine. First it needs to see the total needy in the study area.

Say D is the number of district in the study area, and N is number of needy in such area, the formulation of total needy is:

$$\text{Total needy} = \sum_{d=1}^{D} N_d$$

If C is the school capacity then the total number of all school capacity becomes:

$$\text{Total school capacity} = \sum_{s=1}^{S} C_s$$

But the actual needy that need to be picked up is needy that not live around the schools. If R is the number of surrounding needy in certain school, it makes the actual needy to be picked and the actual school capacity (for needy live far from school) is:

$$\text{Total needy to be picked} = \sum_{d=1}^{D} N_d - \sum_{s=1}^{S} R_s \quad \begin{array}{l} \text{if } C>R \\ \text{if } C>=R \\ \quad R \text{ replaced with } C \end{array}$$

$$\text{New Total school capacity} = \sum_{s=1}^{S} C_s - R_s \quad \text{if } C>R$$

$$= \quad 0 \quad \text{if } C>=R$$

The network data, school location, and depot location is used for the VRP engine. This research uses the VRP algorithm in ArcGis Network Analyst extensions that have been proved that the generated route is optimal in term of time and distance

consume. The engine is using the independent variables which have been explained before. The engine will produce a route that will be a dependent variable for the developed model. The model will have several outputs that indicate the reliability of the route. The first variable, number bus routes, is valid if the generated route visit all schools. If E as number of route, the formula becomes:

$$\text{Number of Route} = E \text{ is valid if } \sum_{e=1}^{E} V_e = S$$

Another constraint is in each generated route must be have school capacity more than covered needy of the route. In order to count the needy covered along the route, the route will be buffered. Says b as the buffer distance, and Db means the buffered area in the district, the constraint of the generated route formulated as:

$$\text{Route e is valid if } \sum_{s=1}^{S_e} C_{e,s} > \sum_{d=1}^{D_b} N_{e,d}$$

Then, if all constraint has been passed, the process will continue to the assessment process. Different combination of independent variable will use to generated another routes, and continue again with the same assessment process. All assessment result will be compared in order to get the most optimal route. The assessment process will check: route balance, number of covered needy and time consumed. The constraint above is already checking the load balance, while time consumed is already done by the VRP engine along with the generated route. For the total covered needy, still with the same variable above, the formula is:

$$\text{covered needy} = \sum_{e=1}^{E} \sum_{d=1}^{D_b} N_{e,d}$$

After the optimal route is chosen, the process will continue with analyzing the route. It have to analyze the passenger load in order to know how is the characteristic of

the route and how many buses have to be assigned in such route. Each route need to be segmented, and each segment need to calculate how many student getting into and or going out from the bus. Total passenger load in certain segment q of certain route e is formulated below:

$$\text{Passenger load of route e segment q} = \sum_{q=1}^{Q_e} (\text{Entry}(q) - \text{Exit}(q))$$

$$\text{Entry}(q) = N_{e,q} - \left(\sum_{q=1}^{Q_e} N_{e,q-1} - \sum_{s=1}^{S_e} C_{e,s} \right)$$

$$\text{Exit}(q) = C_e \quad \text{if there is a school in segment q}$$

$$= 0 \quad \text{if there is NO school in segment q}$$

Maximum passenger load in the segments of route e is used to calculate how many busses have to provide in that route. The route, equipped with the number of the bus, then added in the existing transport system to get the new transportation accessibility level.

3.4 Methodology Schema

The conceptual framework and the variables used to make the model methodology. This research has a big picture which resembles the schema shown in figure 3.15. This schema shows all data that was used as well as all processes that will be followed and implemented in this research. The schema is divided into three stages which include the initial stage, the routing stage, and the analysis stage.

In the first row of the initial stage, there is a listing of data that will explained in chapter 4, namely needy citizen database, region map, school map, depot location, street map, and existing transportation system map. Some processes need to be conducted before this data can move into the routing stage. The needy database was extracted to

get needy student data and then joined with the region map (block map) to make the needy map. This process will be explained in chapter 4. Next process for this needy map is to refine its characteristics and parameters. The refined needy layer will be then used for other processes before the routing process, specifically the extract visiting point process and calculate street load process. The refined needy process, extract visiting point process, and calculate street load process will be detailed in the next section. For school map, there is an extraction process to get the public school data. This public school will join the depot location to be used in the network dataset as a depots layer class. The street map data, after the refinement process to make it fit the requirement of the network analyst, will be used for making a network dataset. This process will be explained in more detail in the next section too. At last, the existing transport system will be used in the third stage, the analyst stage, for creating a 'before and after' view of accessibility level.

The routing stage is the stage where VRP process is followed. The VRP analyst needs Orders layer class, Routes layer class, and Depots layer class. Orders layer get the data from visiting the point map, street load map, and the school map. The school map will be used without a distribution and with a distribution for the undirected VRP and directed VRP respectively.

The routing process will doing several times and will generate several output. The next process is to determine the most optimal solution. The criteria for looking for the best solution will be explained in the next section, and the output will be detailed in the next chapter. The chosen route then moves into analyst processes. There are 2 analyst processes: load analyst and accessibility analyst. The load analyst will use 3D perspective. The accessibility analyst will draw a comparison between the accessibility level in the existing transport system and the accessibility level following the government's addition of the school bus into the transport system.

In this schema, there are differences in color of the entities. Entities with green color are the independent variables while the yellow one are the dependent variables. Next chapter will explain about how data, tabular and spatial, in this research have been developed. Continue with how the each process are and how the transformation of the data in each process.

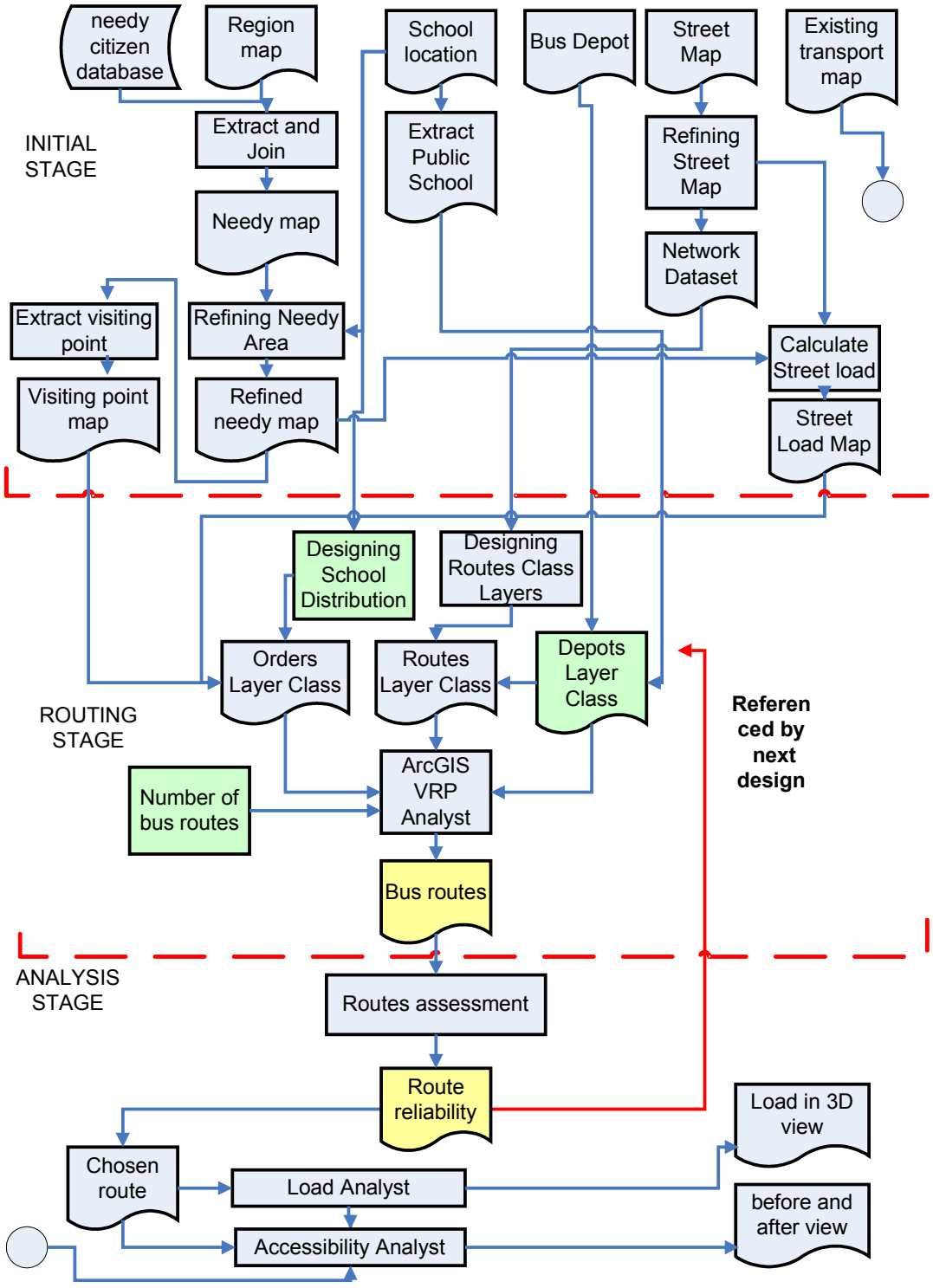


Figure 3.14. Methodology schema

3.5 The Geospatial Technology Used

This section will explore the technology that will be used in this research. The spatial technology for the routing process, the network analyst, has been covered in depth in the previous chapter. This section will now introduce other spatial technology that is used before and after the use of the network analyst. All processes in this research are based on a model. Therefore, first there will be an explanation of the Model, and then a continuation of spatial functions, and finally the end will provide an explanation about Arcscene that is used to present data in 3D perspective.

Models are how people automate their work. When someone creates a model, he is preserving a set of tasks, or a workflow, that he can execute multiple times. There are an infinite number of workflows can be automated using models. People create models using ModelBuilder to chain together tools, using the output of one tool as the input to another tool. The model which has been created is added to ArcToolbox as a model tool, which can execute using its dialog or the Command Line window. People can also execute the model within ModelBuilder.

At the highest level, models contain only three things; elements, connectors, and text labels. Elements are the data and tools to work with. Connectors are the lines that connect data to tools. Text labels can be associated with the entire model, individual elements, or individual connectors. There are two types of model elements, tools and variables. Tool elements are represented with rectangles and are created when we add a tool from ArcToolbox. The color of the tool has meaning, as described in the table below. Table 3.1 shows two different types of tools.

Variables are represented with ovals. We can think of variables as containers that hold values that can be changed. In the context of a model, a variable can be created and its value used in place of a tool's parameter value. There are two types of variables: data and values. Data variables reference data on disk or in an in-memory layer (such as

a layer the ArcMap table of contents). Values are everything else, such as numbers, strings, spatial references, and geographic extents. Table 3.2 shows two different types of variables.

Table 3.1 Color and meaning of the Tool

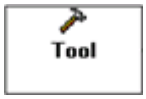
Color	Meaning
	Not all parameters have been supplied for the tool to run. Tool parameters will be supplied by the user of the model when the model is executed.
	All parameters have been supplied. The user does not have to supply parameters when the model is executed.

Table 3.2 Color and meaning of the Variable

Color	Meaning
	Project data is data that we add to the model. Typically, it is the result of specifying a dataset to a tool's input parameter.
	Derived data is new data created by a tool in the model.
	An empty variable has no value.

A tool equipped with its data is called a process. Processes can be in different states: ready-to-run, has-been-run, and not-ready-to-run. A process consists of a tool and all variables connected to it. Connector lines indicate the sequence of processing. Figure 3.24 shows what called process. There will often be several processes in a model, and they can be chained together so that the derived data from one process becomes the input data for another process, as shown in the following diagram.

Starting with this, next paragraph will explain about spatial function that will need to know to build the model in this research. The first function is Select function. Select is a function which extracts selected features from an input coverage and stores them in the output coverage. Features are selected for extraction based on logical

expressions or by applying the criteria contained in a selection file. Any item, including redefined items in the specified feature attribute table of the Input coverage, can be used.

Next is the Calculate function. This function is for calculating the values of a field for a feature class, feature layer, or raster catalog. The input table will be modified; a copy should be made to preserve the original information. Calculate Field computes and assigns a value to the specified field of the Input table. Expressions can be created in a standard Visual Basic (VB) format or in a standard Python format. The formatting style of the string used for the expression should be appropriate to the environment (type). The calculation can only be applied to one field per operation. When calculating new values for a field, existing values will be overwritten. Retain a copy of the input table before using Calculate Field in case an error is made. When calculating joined data, we cannot calculate the joined columns directly. However, we can directly calculate the columns of the origin table. To calculate the joined data, we must first add the joined tables or layers to ArcMap. We can then perform calculations on this data separately. These changes will be reflected in the joined columns.

Buffer function is the next function. We can use the Buffer tool to identify or define an area within a specified distance around a feature. For example, we may create a buffer to define an area around a river to identify land that can't be developed, or we may want to create a buffer to select features within a specified distance of a feature. The Buffer tool creates a new coverage of buffer polygons around the input coverage features. Input features can be polygons, lines, points, or nodes. The width of the buffer can be specified as a fixed distance from an attribute table field (item) or from a distance table.

Buffer function has a close relation with Clip function. Clip creates a new coverage by overlaying two sets of features. The polygons of the Clip Coverage define

the clipping region. Clip uses the clipping region as a cookie cutter; only those input coverage features that are within the clipping region are stored in the output coverage. Input coverage features can be polygons, lines, or points. Clip Coverage features must be polygons. Output coverage features are of the same class as the input coverage features. They are clipped to the outer boundary of the Clip Coverage, and topology is rebuilt for the output coverage.

Next is the most complex function, the identity function. Identity function is used to create a new coverage by overlaying two sets of features. The output coverage contains all the input features and only those portions of identity coverage features that overlap the input coverage. Input coverage features can be polygons, lines, or points. Identity Coverage features must be polygons. Output coverage features resulting from the overlay are of the same class as the input coverage features. Topology is built for the output coverage. Feature attribute tables are updated. The feature attribute table for the output coverage contains items from both the input and identity coverage feature attribute tables. Items are merged using the old internal number of each feature. The following two tables list the items contained in the output coverage feature attribute table.

Next there is Dissolve function. Dissolve is useful when we want to aggregate features based on a specified attribute or set of attributes. For example, we could take a feature class containing sales data collected on a county-by-county basis and use Dissolve to create a layer containing contiguous sales regions based on the name of the salesperson in each county. Dissolve creates the sales regions by removing the boundaries between counties represented by the same salesperson. Features with the same value combinations for the specified fields will be aggregated (dissolved) into a single feature. The Dissolve fields are written to the Output Feature Class table. As part of the Dissolve process, the aggregated features can also include summaries of any of

the attributes present in the input features. For instance, the revenue generated in the counties making up each sales region could be summed to give the total revenue for each sales region.

The process will also need Update function. Update function creates a new coverage by overlaying two sets of features. The features of the update coverage define the updating extent. Update uses the updating extent in a cut and paste operation; update coverage features replace the area they overlap in the input coverage. The result is stored in the output coverage. Both the input and update coverage must have polygon topology. Topology is rebuilt for the output coverage. Attributes are updated. Items in the polygon feature class are merged using the old internal number of each polygon.

In order to make a new layer with operating AND and OR operand, we need Intersect and Union function. Union creates a new coverage by overlaying two polygon coverages. The Output Coverage contains the combined polygons and attributes of both coverages. Only polygon coverages can be combined using Union. Arcs of the Input Coverage polygons are split at their intersection with polygons of the Union Coverage. The resulting arcs are used to build polygons using a process similar to the Build tool with the POLY option. The feature attribute table for the Output Coverage contains items from both the input and Union Coverage attribute tables. Items are merged into the output polygon feature class using the old internal number of each polygon. The following two tables list the items that are saved in the polygon feature class for the Output Coverage.

Intersect creates a new coverage by overlaying the features from the input coverage and intersect polygon coverage. The output coverage contains the input features or portions of the input features that overlap features in the intersect coverage. The output features have the attribute from the original feature from the input coverage and the feature in the intersect coverage, which they intersect. Intersect is one of several

Overlay tools available. The tool most similar to Intersect is Clip, which does not transfer any attribute from the overlay feature class to the output. Input Coverage features can be polygons, lines, or points. The intersect coverage must have polygon topology. Output coverage features resulting from the overlay are of the same type as the input coverage features. They are split when they intersect with the polygons of the intersect coverage. Topology is built for the output coverage. Attribute tables are updated. The attribute table for the output coverage contains items from both the input and intersects coverage attribute tables. Items are merged using the old internal number of each feature.

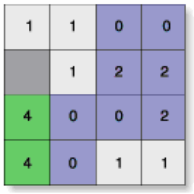
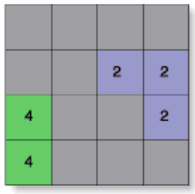
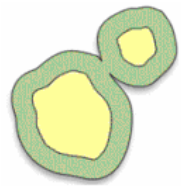
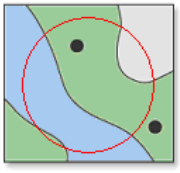
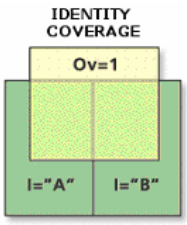
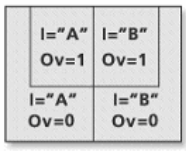
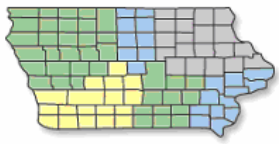
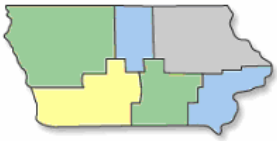
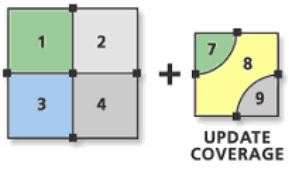
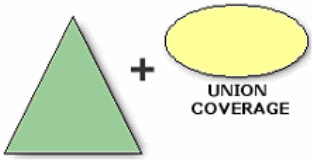
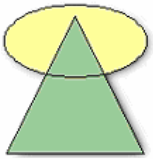
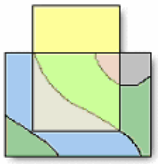
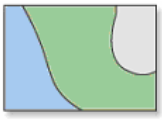
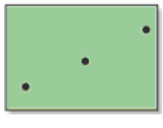
The model will also need to manipulate the tabular data. Most database design guidelines promote organizing our database into multiple tables—each focused on a specific topic—instead of one large table containing all the necessary fields. Having multiple tables prevents duplicating information in the database because we store the information only once in one table. When we need information that isn't in the current table, we can link the two tables together. For example, we might obtain data from other departments in our organization, purchase commercially available data, or download data from the Internet. If this information is stored in a table, such as a dBASE, INFO, or geodatabase table, we can associate it with our geographic features and display the data on our map. ArcMap provides two methods to associate data stored in tables with geographic features: joins and relates. When we join two tables, we append the attributes from one onto the other based on a field common to both. Relating tables defines a relationship between two tables—also based on a common field—but doesn't append the attributes of one to the other; instead, we can access the related data when necessary.

Some function is better using point than using polygon as the input. For that reason it also need feature to point function. This function is to create a point feature

class based on an input polygon, line, or multipoint feature class. The attributes of the input features are maintained in the output points feature class.

All above function example described in Table 3.3. The model is also need to show the analyst process output in its best form. The Arcscene technology fit with this requirement. ArcScene is one of two applications provided by the 3D Analyst extension and allows us to effectively manage our 3D GIS data, perform 3D analysis, create 3D features, and display layers with 3D viewing properties. We can create 3D features from existing two-dimensional (2D) GIS data, or we can digitize new 3D vector features and graphics in ArcMap using a surface to provide the z-values. ArcScene allows us to make realistic scenes in which we can navigate and interact with our GIS data. ArcScene allows us to overlay many layers of data in a 3D environment. Features are placed in 3D by reading height information from feature geometry, feature attributes, layer properties, or a defined 3D surface, and every layer in the 3D view can be handled differently. Data with different spatial references will be projected to a common projection, or data can be displayed using relative coordinates only. ArcScene is also fully integrated with the geoprocessing environment, providing access to many analytical tools and functions.

Table 3.3. Example of geospatial functions

Function	Input	Output
Select		
Buffer		
Clip		
Identity		
Dissolve		
Update		
Union		
Intersect		
Feature to point		

3.5. Summary

This chapter has detailed and summarized the conceptual framework, variables, methodology schema, and the technology used in this research. In the first part, the reader is provided with data which is explained conceptually, then is shown how to process it, and finally how to choose the optimal route. After the optimal route was chosen, this chapter continued with the concept of the analyst process. There are two analyst processes:load analyst and accessibility analyst. This conceptual explanation is supported with an abundance of example figures to facilitate better understanding. The second part is concerned with the variables used in this research, namely the Number of bus routes, Start and End depot, and the group of school as the independent variables and the generated route and its reliability as the dependent variables. The conceptual model and the variables then moved more details in a methodology schema. The last part provides an explanation of geospatial techniques that will be used in the preparation process through the analysis process. These geospatial techniques need to be introduced in order to make the reader easily understand the next chapter which covers the research methodology.