Lesson 1.3: GIS Importing Document

1.3.1 Software and dataset:

Download the GIS Network data into the NeXTA.

Link: https://github.com/xzhou99/learning-transportation/blob/master/Lessons/Lesson%201/Lesson%201.3/GIS_Import_Export_Tool/GIS-Import_Export_Tool.zip

Unzip the file to a known location on your computer, and open the NeXTA_for_GIS.exe file.

1.3.2 Learning Objective:

- Importing GIS Network data into the NeXTA.

The following sections describe techniques and processes for importing network data using NeXTA_for_GIS. Specifically, these tools provide the capabilities of preparing and importing shapefiles from VISUM, Cube, TransCAD and other shapefiles. To demonstrate this functionality, and assist the user throughout the conversion process, one example is provided for converting network data. Please go to the project folder and open "1.Cube_sample_GIS_files" folder.

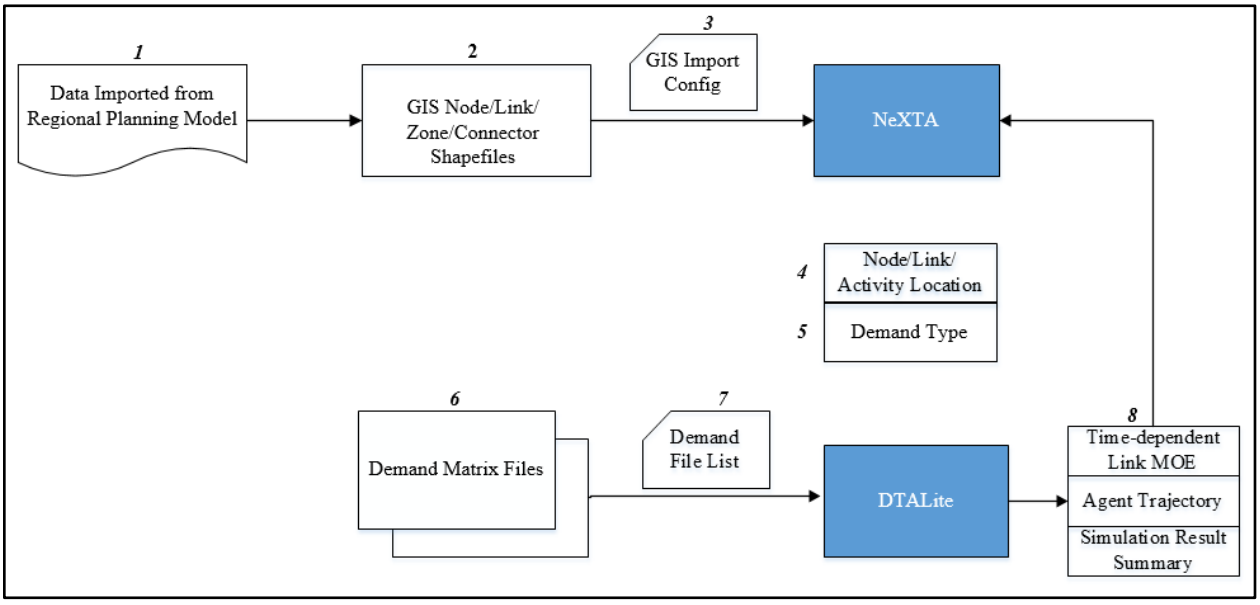


Fig 1: Data Flow for Importing GIS Network to NeXTA_for_GIS

1.3.3 Introduction:

NeXTA_for_GIS network conversion tool specifically imports network data from shapefiles, which are geospatial vector data files used with Geographic Information System (GIS) software and commonly supported by many transportation modeling software packages. Shapefiles contain spatial information about points, lines, and polygons in the transportation network, with separate files for different shapes. Road networks are represented as a graph of links (lines) and nodes (points), where links may represent road segments or transit lines and nodes may represent intersections or connections between individual links, and trips are often made between zones (polygons) on the network.

Network attribute data is stored in database (DBF) files for each shapefile, which must be read by NeXTA during the conversion process. Network data formatting is very flexible between network modeling software packages, with many applications allowing users to define custom formats and data fields for use within a software package. To support these applications, NeXTA_for_GIS uses a configuration CSV file for importing GIS settings, to identify and connect the fields in the input DBF files to their corresponding fields in the NeXTA data format (import_GIS_settings.csv).

Data preparation for building a regional traffic network data hub

The following sections describe processes for preparing network data using NeXTA's network conversion utilities. A network data set for the NeXTA package includes four major CSV files: input_node.csv, input_link.csv, input_zone.csv, input_activity_location.csv, as well as a number of definitional files such as input_node_control_type.csv and input_link_type.csv. There are multiple ways of constructing a network data set for the NeXTA_for_GIS package:

Step 1: Preparing required shapefiles

1. Obtain **shapefiles from regional planning packages** such as VISUM, TransCAD, Cube, which typically consist of node, link and traffic analysis zone layers, with or without centriod and connector layers. Zone shapefile is optional. Sample GIS files can be accessed in GIS-Import_Export_Tool folder from downloaded package.

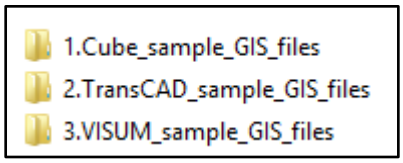


Fig 2: Sample GIS Files

Shape files of South Jordan network is shown in Figure 3.

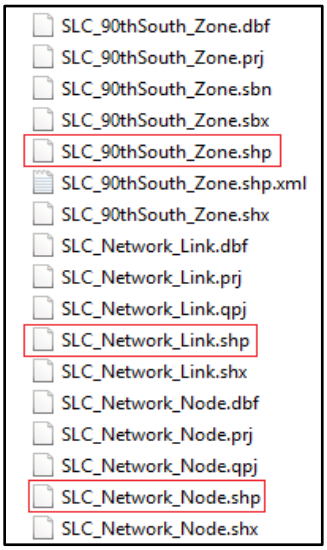


Fig 3: Shape Files of South Jordan Network

2. Or obtain shape files from **GIS shape files with a link layer only**, such as an OpenStreetMap layer, a Traffic Message Channel (TMC) coded network file. First, those GIS sources might not have a separate node layer, so users need to use NeXTA_for_GIS to automatically create the node layer based on the end feature points of a link curve. In order to support the traffic simulation, users also need to create zone layers and define the additional activity layer to link the node layer to the traffic analysis zones.

Step 2: Extract GIS information from shapefiles and importing to configuration files

For importing shapefiles from regional planning packages, users need to prepare input_link_type.csv, import_GIS_settings.csv (required) and input_node_control_type.csv.

1. input_link_type.csv file

The input_link_type table allows users to define their own specific link types, as long as the flag variables

are correctly used to identify how the different link types are connected/related (e.g., freeways connect to arterials using ramps). Only one flag may be used for each link type. Link types can also be used to determine how links are visualized in NeXTA. The data set of link types can be seen at project folder …\GIS-Import_Export_Tool\1.Cube_sample_GIS_files and then open the input_link_type.csv.

Table 1. input_link_type.csv

link_type	link_type_name	type_code	default_lane_capacity	default_speed_limit	default_number_of_lanes	travel_time_bias_factor
1	Freeway	f	1800	0	0	1
2	Principal arterial	a	1000	0	0	1
3	Principal arterial	a	1000	0	0	1
4	Major arterial	a	900	0	0	1
5	Minor arterial	a	850	0	0	1
6	Collector	a	650	0	0	1
7	Local	a	600	0	0	1
8	Frontage road	a	1000	0	0	1
9	Ramp	r	1300	0	0	1
10	Zonal connector	c	2000	0	0	1
100	Transit link	t	1000	0	0	1
200	Walking link	w	1000	0	0	1

2. input_node_control_type.csv

The input_node_control_type table defines the control type of nodes in the network in terms of control type name, unknown control, no control, yield sign, 2way stop sign, 4way stop sign, pretimed signal, actuated signal and roundabout. This file is required when using the network import tool, and the control type field is read from the node shape file. The data set of node control type can be opened at project folder …\GIS-Import_Export_Tool\1.Cube_sample_GIS_files.

Table 2. input_node_control_type.csv

control_type_name	unknown_control	no_control	yield_sign	2way_stop_sign	4way_stop_sign	pretimed_signal	actuated_signal	roundabout
control_type	0	1	2	3	4	5	6	100

3. **import_GIS_settings.csv file**

File import_GIS_settings.csv is a user-defined configuration file, and a user can use Excel to edit it. This file can help users to associate/map fields in shapefile DBF files to the AMS Data Hub schema data format, and **NeXTA_for_GIS** can use this file to read the network geometry from shapefiles and create

an Analysis Modeling Simulation (AMS) Data Hub compatible transportation network project file (which is readable by both DTALite and NeXTA).
Sample GIS importing files can be found under folder ...\GIS-Import_Export_Tool\1.Cube _sample_GIS_files.

Depending on the type of data to be imported. the configuration file consists of several **sections**. As defined in Column **A**, there are general sections of file, configuration, and sections for different network objects such as node, link, and zone.

The "**key**" settings given in Column **B** are used by DTALite/NeXTA to match and convert the shapefiles from the DBF files.

"**value**" in Column **C** is the matching name between NeXTA's "key" values, and those given in the shapefiles.

Column "**required_or_optional**" shows which values must be defined for a successful import. Column "**allowed_values**" show which definitions are related to "value" for binary values (such as oneway vs. twoway, lane vs. link, etc). **Notes** column give some information how to fill the column **value**.

Hints:

- Although this configuration contains many columns, **users only need to prepare or change the data in column "value".** The converted shapefiles can be opened in any GIS software (we recommend open-source Q-GIS) and accessed through attribute tables.
- Once a configuration file is established for a regional travel demand format, it can be **replicated as a template** for future imports from that format.
- The open-source GIS library GDAL is used by NeXTA to import the geometry and field data from GIS shape files.
- The user needs to be careful with **attribute names that have more than 10 characters** (such as ControlType in this case). For example, an exported SHP/DBF file will convert the file name to the 10-character length (Controlt~1). To see the correct attribute name definitions, the user can use any GIS software to open the corresponding shapefile and read the name from the attribute table.

The first **file section** defines the shape file names for GIS layers of node, link, and zone. In this example, the file name table looks like the following table. The reference file names should correspond to the layers exported from planning packages such as VISUM, CUBE and TransCAD. Node, link and zone are the required layers for a successful GIS network import.

Table 3. File Section

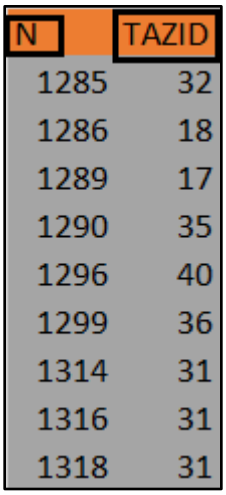
section	key	value	required_or_optional
file_name	node	SLC_Network_Node.shp	optional
file_name	link	SLC_Network_Link.shp	**required**
file_name	zone	SLC_Network_Zone.shp	optional

Remarks:

- ❏ Only the link layer is required.

The next **section** defines **node** shapefile attributes. The import values are node ID, node name, the zone (TAZ) to which a certain node belongs to, and node control type. The sample node table is as follows, while fields node_id and TAZ are **required** if we have had node layer. This part will fill out with the data provided in SLC_Network_Node.dbf as shown in table 5.

Table 4. SLC_Network_Node.dbf



N	TAZID
1285	32
1286	18
1289	17
1290	35
1296	40
1299	36
1314	31
1316	31
1318	31

Table 5. Node Section

section	**key**	**value**	**required_or_optional**
node	node_id	N	Required only when the node layer file is provided
node	name		Optional, for display only
node	TAZ	TAZID	Optional, if not provided, then users can use node_number_threshold_as_centroid to specify the first n nodes directly as centroids of zones.
node	control_type		optional, if node control type has been specified. If not, set value from file input_node_control_type.csv

Table 6: Configuration

Configuration	**key**	**value**	**allowed_values**

1	with_decimal_long_lat	yes	yes;no
2	node_number_threshold_as_centroid	0	Default 0

Item-by-item descriptions:

1. Key **with_decimal_long_lat** indicates if a decimal long/lat format is used.
   For example, longitude -111.943375 has a decimal point, while -111943375 has no decimal point. If "value" is set to no, a multiplier of 0.000001 is used to convert the sample value of -111943375 to 111.943375.
2. Key **node_number_threshold_as_centroid** specifies the value for using nodes directly as zone centriods. In VISUM or TransCAD data set, a centroid layer is provided, while Cube might assume a set of nodes below a certain number (say 3000) is used as zone centroids by default. This number typically is specified in a script for traffic assignment in Cube. For example, there are 2000 nodes with node numbers < 3000. By reading this number threshold, NeXTA can automatically (a) add 2000 traffic analysis zones, (b) assign a node to the activity location for each newly created zone. Thus, the zone and activity location layers (namely files input_activity_location.csv and input_zone.csv) are created automatically.

The **link section** defines the link shapefile attributes. In order to code the corresponding attributes correctly, the user can access the link shapefile through GIS software, and read the link attributes. The available link attributes are from and to node, name, link ID, link type, transportation modes that link is open for, direction definition, number of lanes, hourly capacity, speed limit, and number of lanes, capacity, speed limit and link type for reversible lanes, if links are defined as two-way links. Table 8 represents how user should extract GIS information from link dbf file as shown in table 7 and modifying GIS setting file based on that.

Table 7. SLC_Network_Link.dbf File

A	B	DISTANCE	STREET	ONEWAY	TAZID	LINKID	LANES	FT	SFF	SFF_TIME	CAP1HR1LN	Length
1285	5018	0.23840	Cent	2	1285	1285_5018	7	1	21.17647	0.67547	10000.00000	
1286	11125	0.46596	Cent	2	1286	1286_11125	7	1	21.17647	1.32022	10000.00000	
1289	4952	0.24269	Cent	2	1289	1289_4952	7	1	21.17647	0.68762	10000.00000	
1289	5018	0.26213	Cent	2	1289	1289_5018	7	1	21.17647	0.74270	10000.00000	
1289	11124	0.50688	Cent	2	1289	1289_11124	7	1	21.17647	1.43616	10000.00000	

Table 8: Link Section of import_GIS_setting.csv

	section	**key**	**value**	**required/optional**	**Remark**
1	link	from_node_id	A	Optional but highly recommend to fill in the cell	Typical values include usn. If data is not provided, please set identify_from_node_id_and_to_node_id_based_on_geometry to yes
2	link	to_node_id	B	Optional but highly recommend to fill in the cell	Typical values include dsn. If data is not provided, please set identify_from_node_id_and_to_node_id_based_on_geometry to yes
3	link	name	STREET	optional	For display only
4	link	link_id	LINKID	optional	For searching links by ids
5	link	link_type		optional	If data is not provided, it will set value from input_link_type.csv
6	link	mode_code			Reserved for transit, pedestrian modeling.
7	link	direction	ONEWAY		If no value is given, a one-way link is assumed by default.
8	link	length	DISTANCE	required	The unit should be miles.

9	link	number_of_lanes	LANES	optional	Users can specify use_default_number_of_lanes_from_link_type based on link type.
10	link	hourly_capacity	CAP1HR1LN	optional	Users can specify use_default_lane_capacity_from_link_type based on link type
11	link	speed_limit	SFF	optional	Users can specify use_default_speed_limit_from_link_type based on link type

Remarks

There is a separate CSV file for the link types (input_link_type.csv) that needs to be updated with the correct link types. Number of lanes, speed and capacity data are indirectly defined through link type table, users can also set the following key to **yes** in the configuration section:
use_default_speed_limit_from_link_type;use_default_lane_capacity_from_link_type;
use_default_number_of_lanes_from_link_type

The next **configuration section** describes general model attributes and import options that can accommodate different network coding conventions. These settings include generating from node and to node ids, whether the capacity is given per lane, or per link, and default direction of links. Its advanced settings also allow users to import link typ, specific speed limit and capacity values, number of lanes, and minimum length for importing links. In the following example, the configuration settings are given with sample values.

Table 9: Link Configuration

configuration	**key**	**value**	**allowed_values**
1	identify_from_node_id_and_to_node_id_based_on_geometry	no	yes;no
2	lane_capacity_vs_link_capacity	lane	lane;link

3	default_link_direction	oneway	oneway;twoway
4	use_default_speed_limit_from_link_type	no	yes;no
5	use_default_lane_capacity_from_link_type	no	yes;no
6	use_default_number_of_lanes_from_link_type	no	yes;no
7	minimum_length_for_importing_links	0.00001	

Item-by-item descriptions:

1) In the link layer, if the from_node_id and to_node_id of a link is not provided, for example, in a TransCAD data set, then the flag of **identify_from_node_and_to_node_id_based_on_geometry** can allow NeXTA to use the geometry information of a link to find the nearby nodes from the node layer to construct the from_node_id and to_node_id fields for a link. If no such a node is available or even no node layer is provided, then NeXTA_for_GIS will create new nodes and use the corresponding new node numbers for the fields of from_node_id and to_node_id. This setting significantly relaxes the data requirements.

2) Key **lane_capacity_vs_link_capacity** indicates if the capacity filed in the link layer is referred to link capacity across all lanes or simple lane capacity. If "value" = **link**, then the capacity of link will be divided by the number of lanes as DTALite/NeXTA uses **lane-based** capacity.

3) Key **default_link_direction** specifies if a link record is a oneway link or a two-way link. If a record is a two-way link, then two directional links will be created by NeXTA.

4,5,6) Keys **use_default_speed_limit_from_link_type; use_default_lane_capacity_from_link_type; use_default_number_of_lanes_from_link_type** allow users to input their link-type specific lane capacity, number of lanes and speed limit fields in file input_link_type.csv. If "value" = yes, then users do not need to specify the field names in section link for these three important link attributes.

7) Key **minimum_length_for_importing_links** specifies the minimum length threshold for importing a link. With a default value of 0.00001, NeXTA does not import a link with a length of 0. If a negative value

of -0.00001 is given, then NeXTA will keep the links with zero distance in the final imported data set.

The **zone section** defines the zone shapefile attributes (SLC_90thSouth_Zone.shp in this example). Only a zone id field is needed. If the zone layer does not present, one can set a positive value (say 3000) for key **node_number_threshold_as_centroid** in section configuration to add zones and the corresponding activity locations.

Table 10: Zone Section

section	**key**	**value**	**required_or_optional**
zone	zone_id	Id	optional

Remarks:

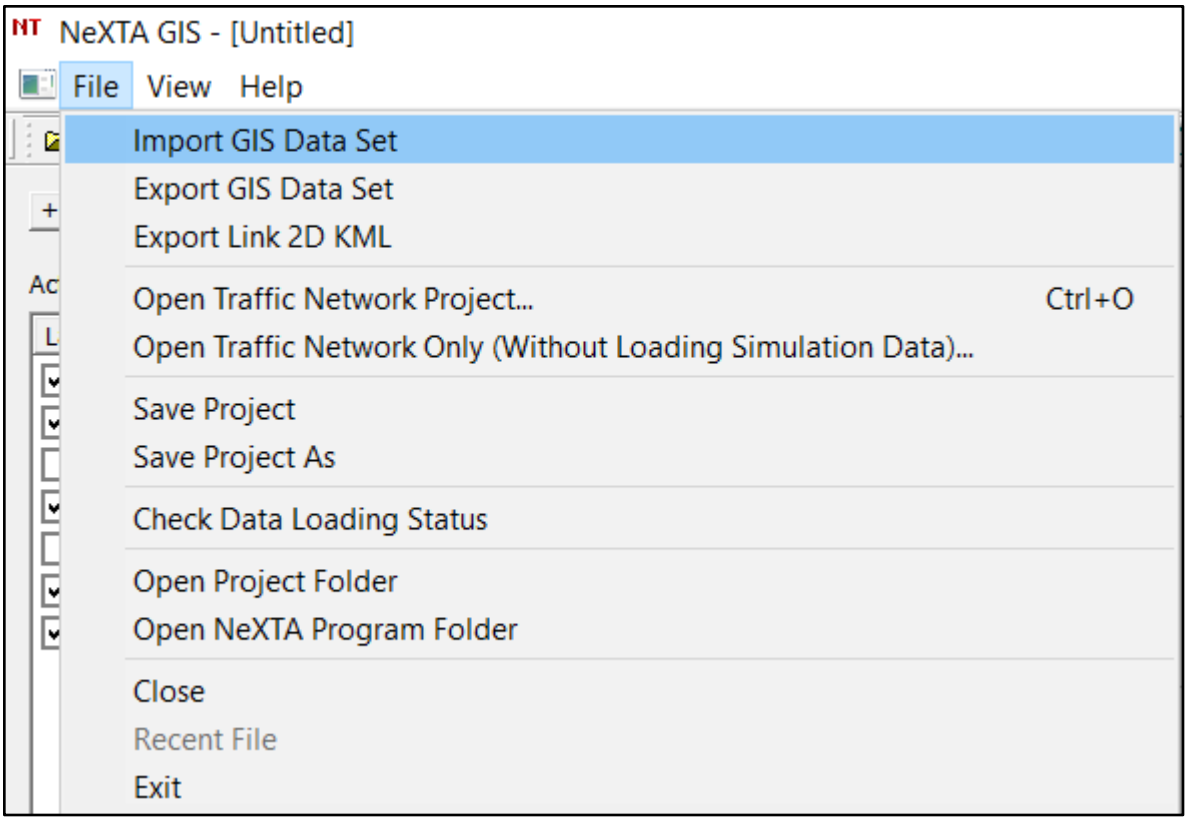
- ❑ Zone is required only if zone file is given in the section file_name.

A complete sample imput_GIS_settings.csv file can be seen under project folder …\GIS-Import_Export_Tool\1.Cube_sample_GIS_files.

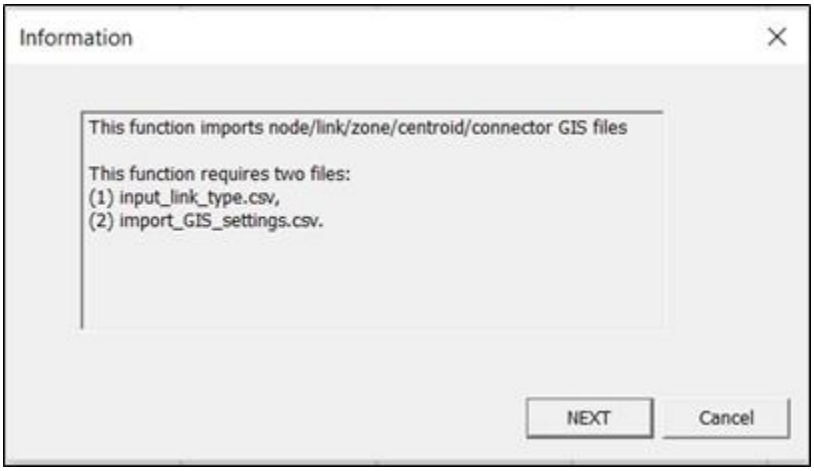
Step 3: Generating the traffic network using NeXTA_for_GIS

Step 3-1: Open NeXTA_for_GIS.exe

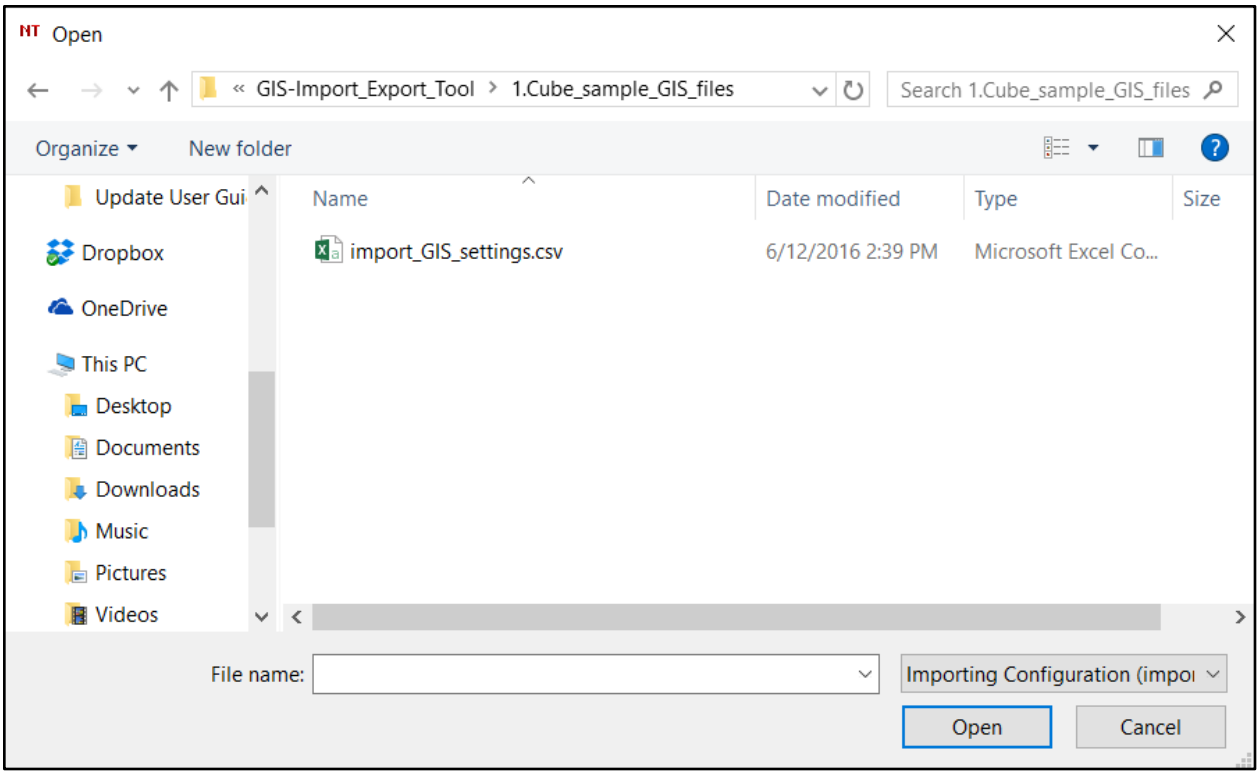
Step 3-2: Click on “File”→“ Import GIS Data Set”



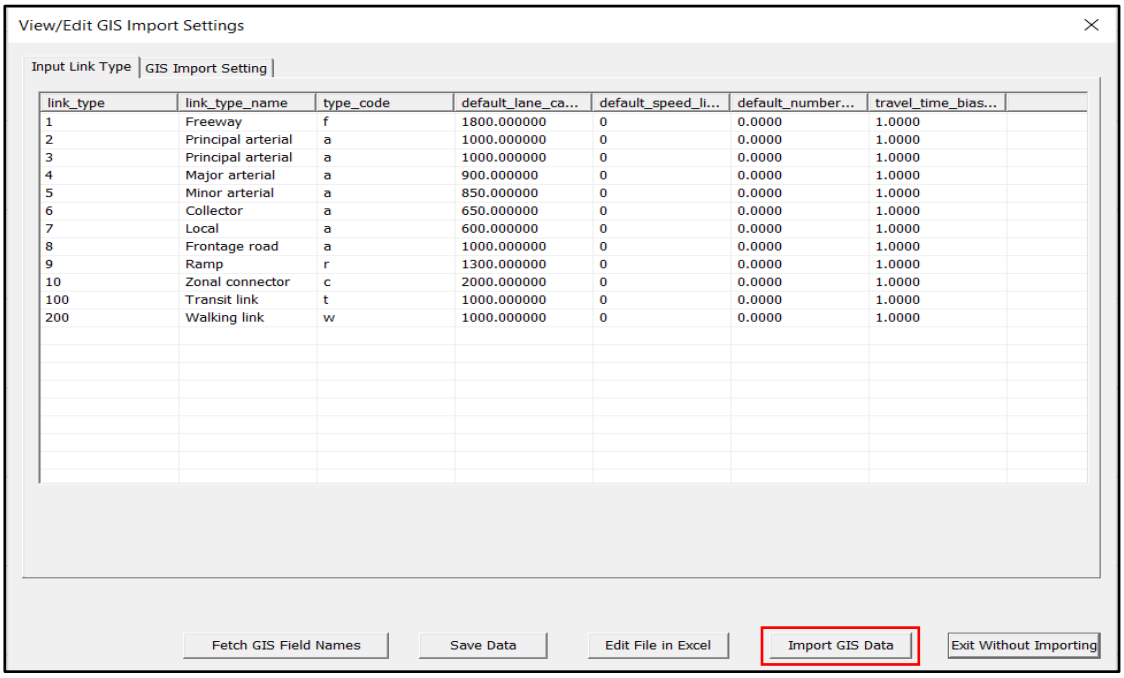
The following dialogue will be opened:



Click on NEXT to load import_GIS_setting.csv from folder GIS-Import_Export_Tool\1.Cube_sample_GIS_files.



Step 3-3: Select “Import GIS Data”

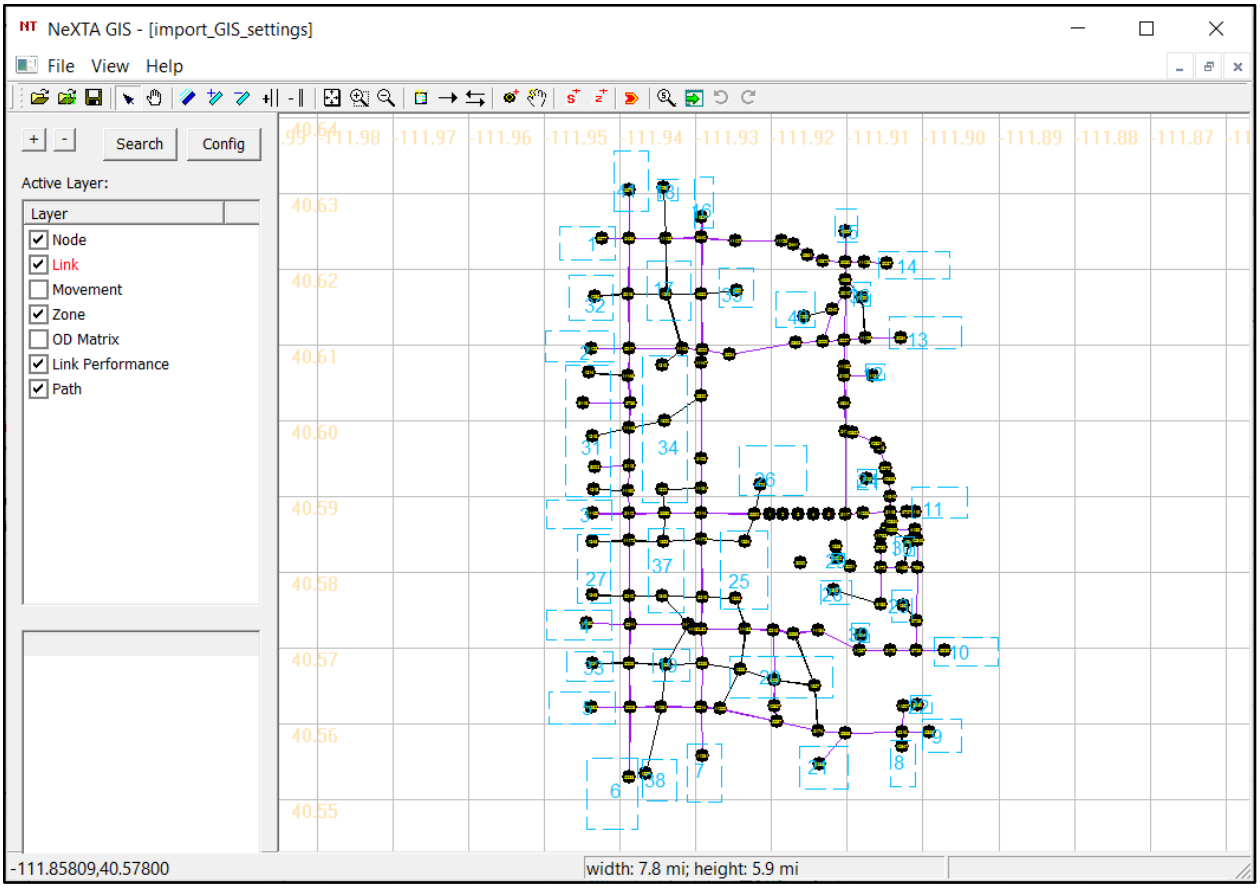


link_type	link_type_name	type_code	default_lane_ca...	default_speed_li...	default_number...	travel_time_bias...
1	Freeway	f	1800.000000	0	0.0000	1.0000
2	Principal arterial	a	1000.000000	0	0.0000	1.0000
3	Principal arterial	a	1000.000000	0	0.0000	1.0000
4	Major arterial	a	900.000000	0	0.0000	1.0000
5	Minor arterial	a	850.000000	0	0.0000	1.0000
6	Collector	a	650.000000	0	0.0000	1.0000
7	Local	a	600.000000	0	0.0000	1.0000
8	Frontage road	a	1000.000000	0	0.0000	1.0000
9	Ramp	r	1300.000000	0	0.0000	1.0000
10	Zonal connector	c	2000.000000	0	0.0000	1.0000
100	Transit link	t	1000.000000	0	0.0000	1.0000
200	Walking link	w	1000.000000	0	0.0000	1.0000

The file loading status will pop up, which gives all the information of networks such as number of nodes,

links, and zones.

You will see the following network on the display view.



Step 3-4: Go to the “File” -> “Save Project As” to save the imported network as a new text file and generate the necessary input files for the model.

Questions:

These sample questions try to evaluate and improve your understanding of this learning document.

1. How many GIS shape files do we need?

2. Does DTALite require sequential node numbers?

3. Are there any limitations on number of nodes, links and zones?

4. Can we only have link layer shapefile?

5. Is zone layer required?

6. Which shape file generates input_activity_location.csv?

7. Does the GIS projection system (e.g. WGS 84) matter?

8. Can one zone have multiple activity locations?

9. Can an activity location be shared by multiple zones?

10. Is zone boundary required for simulation?

11. Is the curve of links required?

12. Can we have the same upstream node and downstream node numbers for different links?

13. What is centroid?

14. What is centroid connectors?

15. How many layers is required by NeXTA and DTALite at minimum?

Questions with answers:

1. **How many GIS shape files we need?**
   Only the link shape file is required. The other shape files are optional (node, zone, centroid, connector).

2. **Does DTALite require sequential node numbers?**

   No, it can be arbitrary positive number, for example 1,4,7.

3. **Are there any limitations on number of nodes, links and zones?**
   No, there are no limitations on number of nodes, links, and zones

4. **Can we have only one link layer shapefile?**

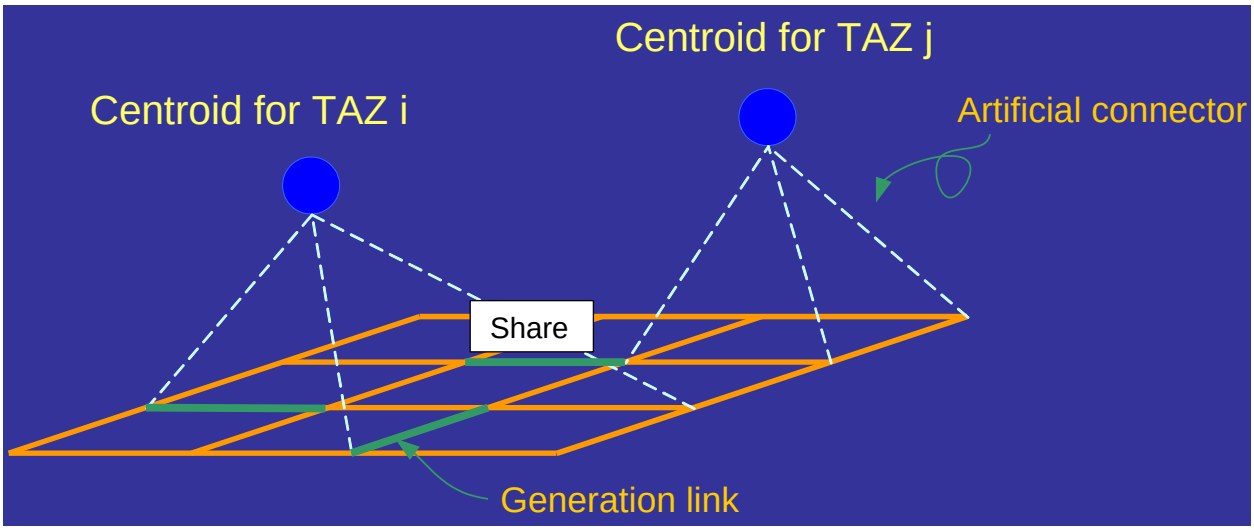
Link layer is the only layer that is required.

5. **Is zone or connector layer required?**
   No, it isn't.

6. **Which shape file generates input_activity_location.csv?**
   Node shape file.

7. **Does the GIS projection system (e.g. WGS 84) matter?**
   No, it doesn't. The node and link layers can use arbitrary coordinate system, but a WGS 84(long/lat) coordinate system is preferred to export data in Google Earth/ Google Map. FYI, In GIS, there are two types of coordinate systems: Geographic Coordinate System (GCS) and Projected Coordinate System (PCS).

8. **Can one zone have multiple activity locations?**
   Yes, in shown picture zone has 3 activity locations.



9. **Can an activity location be shared by multiple zones?**
   Yes. It's shown in above picture.

10. **Is zone boundary required for simulation?**
    No, just for display in NeXTA.

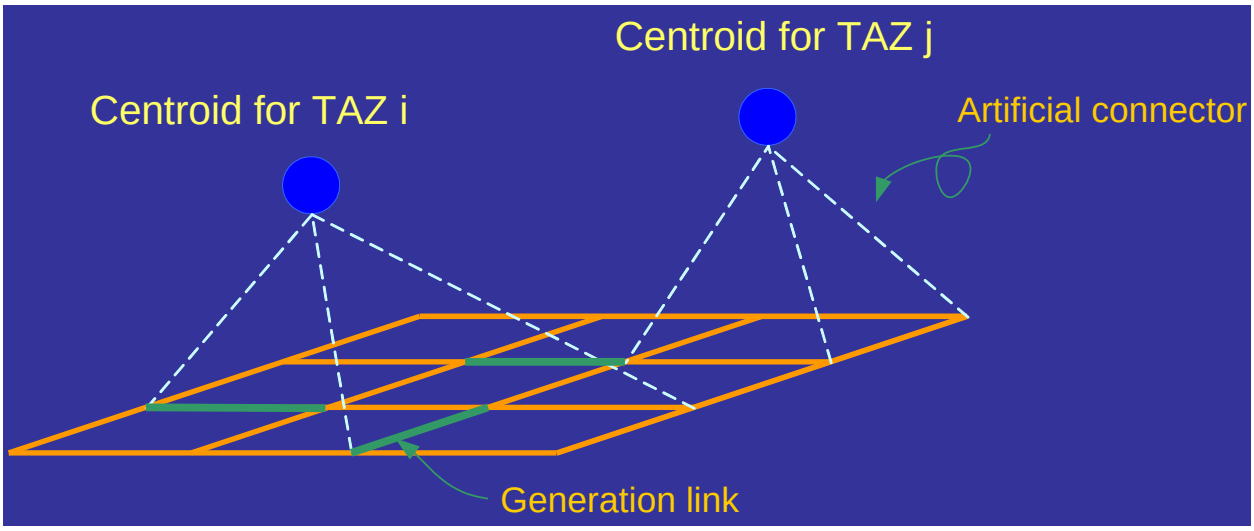
11. **Is the curve of links required?**
    No, we use straight line for display.

12. **Can we have the same upstream node and downstream node numbers for different links?**

No.

13. **What is centroid?**
    Points that identify the center of activity within a zone and connect that zone to the transportation network.

14. **What is centroid connectors?**
    Links that connect the centroids to the network.



15. **How many layers is required by NeXTA and DTALite?**
    4 Layers, node, link, zone, activity locations.