

A Technique for Merging State and Non-State Linear Referencing Systems

Prepared By:

Zach Ryals

Department of Civil and Environmental Engineering

University of Alabama – Tuscaloosa

Phone: 1-205-616-2587

E-mail: ryals006@crimson.ua.edu

Submitted to:

AASHTO GIS-T 2011 Student Paper Competition

December 6, 2010

Table of Contents

Abstract	iv
1. Introduction.....	1
2. Literature Review and Background.....	2
2.1 Linear Referencing System Conceptual Model	3
2.2 Background	5
2.3 Data Merging Issues	6
3. Methodology	7
3.1 STN and WISLR Systems	8
3.2 Link_link Table Design and Population	10
3.3 Flag Columns	12
3.3.1 Turn-lane Flag	13
3.3.2 Median-Crossover Flag.....	14
3.3.3 Gore Point Flag.....	14
3.3.4 Weigh-station Flag.....	15
3.3.5 Problem Flag	15
3.4 Link_link Coding Tool	16
3.5 STN Point Generator and the Point Moving Program.....	17
3.6 Quality Assurance / Quality Control (QA/QC).....	18
3.6.1 STN Link Check.....	18
3.6.2 WISLR Link Check	19
3.6.3 WISLR Link Visual Check	19
3.6.4 Access Point Check	20
3.6.5 Gore Point Check.....	20
3.6.6 Point Moving Program Output Check	21
3.6.7 XY Connector Line Check.....	21
3.6.8 Reversed Link Identification Tool.....	22
4. Pilot Study	23
5. Results of the Pilot Study	24
6. Conclusion and Future Work.....	25
6.1 Conclusion	25

6.2 Future Work	25
7. References.....	26

List of Figures

Figure 1. Generalized LRS model a.k.a. NCHRP 20-27(2) model (Vonderohe et al., 1997).....	3
Figure 2. Visualization of 20-27(2) conceptual model (Scarponcini, 2001)	5
Figure 3. STN and WISLR conceptual model (Graettinger, 2008) and functional join example ..	6
Figure 4. STN and WISLR systems in downtown Madison	8
Figure 5. STN and WISLR are very different in logical design in this interchange area	9
Figure 6. Typical link_link coding example.....	11
Figure 7. Access Points provide calibration for the link_link table.....	12
Figure 8. Turn-lane and median crossovers are not represented in WISLR by links	13
Figure 9. Aerial image of a gore point alongside STN and WISLR	15
Figure 10. Link_link Coding Tool UI	17
Figure 11. WISLR link visual check method	19
Figure 12. Gore point at an off-ramp	20
Figure 13. XY connector lines showing correctly and incorrectly coded links	22
Figure 14. Simplified examples of correct and incorrect XY lines	23
Figure 15. Pilot study area with RP crashes displayed	24

List of Tables

Table 1. Names and descriptions of the six basic link_link columns	10
Table 2. Example link_link table coded using ratio calculations	11
Table 3. Link_link table coded using Access Points.....	12
Table 4. Flags columns and descriptions.....	13

Abstract

The Wisconsin Department of Transportation (WisDOT) has two separate linear referencing systems (LRS) for managing state roads and local roads. The State Trunk Network (STN) only represents statewide routes such as interstates, state highways, and county highways. The Wisconsin Information System for Local Roads (WISLR) system includes all roads, but focuses on local roads. These systems are not related; link IDs (road segments) and node IDs (intersections) are not the same between the systems and even the lengths of roads in the two systems are measured in different units. Therefore, the focus of this project was to develop a technique to relate the two systems.

Merging STN and WISLR is advantageous because it allows the transfer of transportation data between the two systems. To move this data, a table, called the link_link table, was created that relates the two systems based on link IDs and lengths of links. Point data in STN (in the form of link IDs and offset distances) can be programmatically moved using the link_link table to create an equivalent table of WISLR link IDs and offset distances. A pilot study was conducted to test and confirm the procedures for coding the link_link table. This paper presents the link_link table design and coding methodology, quality checks for the link_link table, the results of moving data from STN to WISLR, and recommendations for future work related to this project.

1. Introduction

Since the advancement of computers and computer software, transportation data storage and analysis methods have drastically changed. The introduction of Geographic Information Systems (GISs) has revolutionized not only the way people display and process data, but also the amount of data that can be handled at one time. GISs are a genre of software similar to computer-aided drafting (CAD) programs and database management programs. GISs allow visualization, storage, and spatial analysis of location-related data, while providing robust storage and analysis of attribute data that may or may not be directly related to the spatial data. GISs provide tools that can be applied to many different fields, especially transportation.

The past two decades have witnessed many studies, conferences, and workshops that have focused on the development of a GIS that can efficiently utilize and display transportation data (Miller and Shaw, 2001; Vonderohe et al., 1997). This type of GIS is known as GIS-T, which is designed to handle transportation-related data such as networks, pavement types, bridge locations, crash data, and other Department of Transportation (DOT) data.

Initially, state DOTs maintained data about statewide routes within their jurisdiction; these included both interstates and state roads. Local roads have, for the most part, been maintained by county and city DOTs, who essentially ignored data regarding the statewide routes. Recently, the federal government has called on state DOTs to begin mapping and analyzing local roads in addition to the state routes to allow more complete understanding of statewide transportation data. This presented a problem for most states since local road systems had either been developed in a county-by-county or city-by-city manner or not at all.

The Wisconsin DOT (WisDOT) built a Linear Referencing System (LRS) called the State Trunk Network (STN) in the mid 90's to satisfy the business needs of WisDOT relative to statewide routes. About a decade later, a second LRS was developed called the Wisconsin Information System for Local Roads (WISLR) that focused mainly on local roads and maintained little or no data regarding the state roads. Each of these systems was designed to meet the business needs of WisDOT, which is well achieved within each system; however, since STN and WISLR were designed and implemented separately, there is little correlation or data sharing between the two systems. Therefore, data that is maintained in one system has no way of being transferred to the

other system. Given that there are approximately 12,000 miles of pavement that overlap in each system, which encompasses some of WISLR and all of STN, there exists a great need to develop a bridge between the two systems that affords efficient data sharing and transferability (Graettinger et al., 2008; Graettinger et al., 2009).

Unfortunately, since these two systems were developed in a disconnected environment, there are inherent differences in the technique for drawing complex transportation features. STN describes transportation entities such as intersections in much greater detail than WISLR. STN accounts for these intersections by placing several links to represent all of the different directions and turns that can be made; however, WISLR may draw the same intersection as a single node.

This paper describes a method that effectively relates the two systems without the need to interrupt business practices at WisDOT. In this join, neither system is required to adjust or change any attribute information or spatial features (with the exception of incorrect areas that are occasionally encountered when processing the systems). The paper details methods of mitigating problems with inconsistent drawing methodologies, quality control practices to ensure the accuracy of the join product, and the process of moving data from STN to WISLR. The paper concludes with recommendations for future work regarding possible methods for moving data from WISLR to STN (which requires moving data from a low-resolution system to a high-resolution system), as well as maintenance issues that may be encountered with this join method.

2. Literature Review and Background

LRSs, sometimes referred to as linear location referencing systems, were developed to map transportation business data. The main advantage of a LRS is that it does not necessarily require a spatial reference as do geographic datasets. Two-dimensional and 3-D data can now be displayed on a 1-D platform that is not necessarily related to any specific location (Scarponcini, 2002). Unfortunately, the LRSs of the past years were evolved through business needs, and were never properly designed (Vonderohe and Hepworth, 1998), which increases the difficulty of designing and implementing a merge strategy.

2.1 Linear Referencing System Conceptual Model

A properly designed LRS is made from a set of specific components, data and tools that comprise a system where the goal is to locate transportation data. The LRS data model has been studied with the intent of developing a generic model that encompasses all business needs of federal, state and local DOTs. The conceptual model in Figure 1 shows the LRS as a compilation of three main parts: a datum, network (s), and linear referencing methods (LRMs).

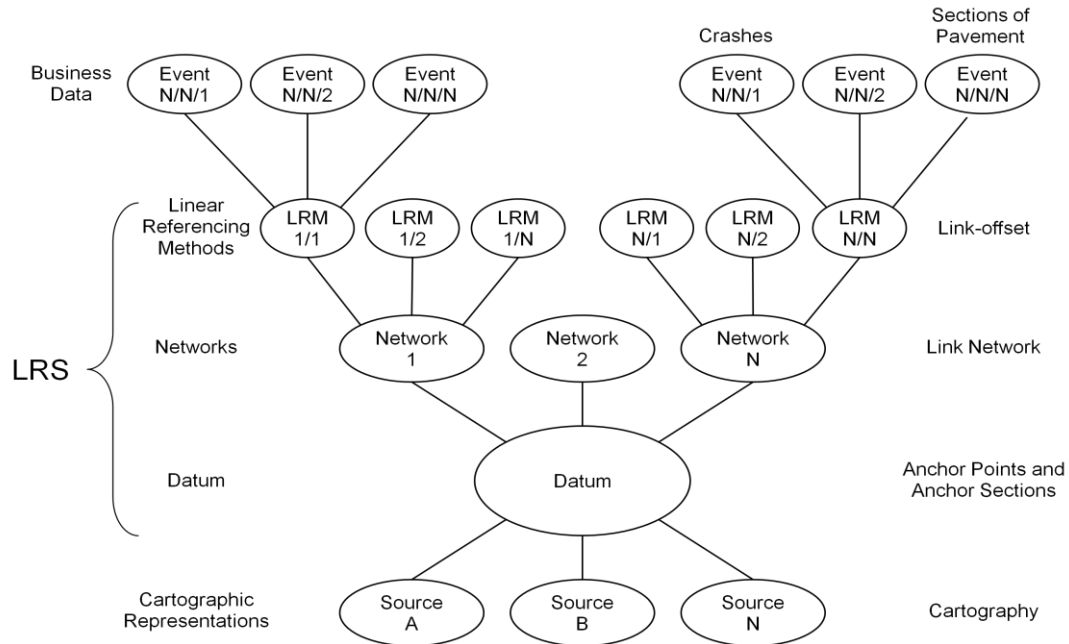


Figure 1. Generalized LRS model a.k.a. NCHRP 20-27(2) model (Vonderohe et al., 1997)

The datum is an absolute set of anchor point and anchor sections. These anchors are points and segments are related to real locations on the earth, and act as a platform for movement among the other parts of the conceptual model. Anchor points are zero-dimensional locations and require some detailed explanation of location in the field, which can be quantitative and/or qualitative. Anchor sections are non-branching lines that are solely a connection between two anchor points. The length of the anchor section can be calculated in the field to provide an accurate relationship between the anchor points (Vonderohe et al., 1997).

A network is best described as a means for interaction and movement among point locations (Miller and Shaw, 2001). There are several types of networks, and these different types of networks can be present simultaneously through a common datum that is associated with each LRS, as shown by Network N in Figure 1. A common network type is a link-node system,

where links are directional and act as conduits for flow, and nodes are locations where conduits meet. Vice-versa, nodes can be described as locations where flow can change, and the links simply connect certain nodes, as described in WisDOT's Location Control Management Manual (1995). The network often, but not always, describes the type of LRM that can be used in that system of features.

A LRM is a method for describing the location of transportation data on a given network. Multiple LRMs can branch from a single network, as seen in Figure 1 as LRM 1/N (the 1/N signifies that this is the Nth LRM on Network 1). Some common LRMs are street address; route-milepost; stationing; on/at, distance, direction; and link-offset, sometimes called reference point (Scarponcini, 2002; WisDOT LCM Manual, 1995; Graettinger et al., 2008; Graettinger et al., 2009). The link-offset method is employed by WisDOT in the STN system. This method states the directional link on which the transportation data is located, as well as the distance down link that must be traveled from zero to the event.

Notice that the LRS does not encompass the entire model; it is comprised of only the essential components and all other data either sits on top of or hangs off the LRS. These additional pieces of data are cartographic representations and business data, at the bottom and top of Figure 1, respectively.

Events are the visual product of processing business information through a LRS and are at the center of geographic analysis. Events can be points or lines, depending on the data format and the needs of the user. In a link-offset LRM, points will be represented by a link ID and an offset measure, while linear events will be represented by a link ID, a start measure, and an end measure (Vanderohe and Hepworth, 1998). Bridge locations and segments of pavement are tangible data events, while crash points and project reference lines are abstract data events. One important point to remember, however, is that events are generated solely through a LRS, and since the LRS is an abstract representation of reality, the locations of events will not always correspond to their actual location in the field.

Cartographic representations can change over time and show varying levels of detail depending on the specific needs of the DOT, therefore, several maps can be related to the LRS based on the virtual anchors of the linear datum. This cartography is not necessary to the functionality of the

LRS, but can provide visual perspective to better understand the relationships of the network(s) and event data.

A graphic version of the levels of the conceptual LRS data model is shown in Figure 2. Starting from the top of Figure 2, there are Events located by the LRM, the LRM works by referencing the Network, the Network is located on the earth's surface by the Datum, and the cartography is overlaid onto the Datum for visualization.

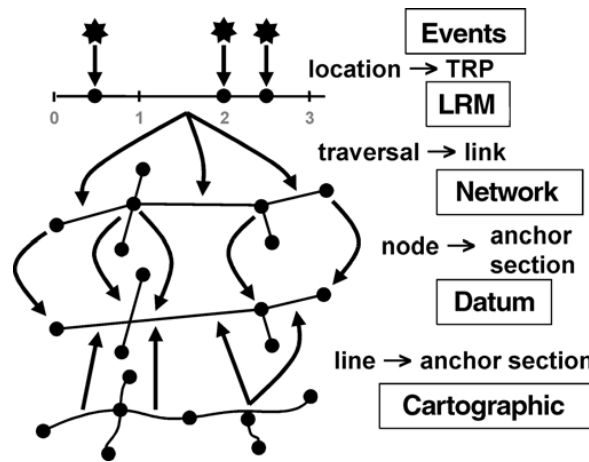


Figure 2. Visualization of 20-27(2) conceptual model (Scarponcini, 2001)

2.2 Background

STN and WISLR are two separately designed LRSs that were developed independently of one another. STN was developed in the 90's to satisfy WisDOT's business needs for maintaining Wisconsin's interstates and state roads. WISLR was developed about a decade later to encompass all roads, but focused on local roads. STN was fully functioning by the start of WISLR development, and it would not have been practical to suspend the utilization of STN in order to create the WISLR system. This separation is acceptable to WisDOT considering that each system performs the job for which it was built. However, each system contains some information that the other does not, which decreases the usability of the systems on a macroscopic scale. WISLR shares approximately 12,000 miles of linework with STN (Graettinger et al., 2008; Graettinger et al., 2009), but these LRS features are almost completely unrelated, other than representing the same actual features. Business data is generated in both

systems, but cannot be analyzed effectively relative to both systems, so the need for a merge exists.

WisDOT chose to keep STN and WISLR separate in almost all respects, even when devising a linear datum. Figure 3 details this problem from a conceptual standpoint. The center of the figure shows a possible design of the STN and WISLR systems, where the separate networks are related to a common linear datum. However, on the left and right side of the Figure 3, it is apparent that the systems are inherently different. The systems are similar in that both lack a distinct datum; rather, each system has the datum embedded in the respective network, which is to say that the network and datum are functionally equivalent (though in WISLR, the cartographic chains are part of the network, and there are no driven distances, so there is technically no functional datum). The chains at the bottom of the figure are the cartographic entities that represent the state routes and local roads. Another difference between STN and WISLR is that each system employs a different LRM (Graettinger et al., 2008; Graettinger et al., 2009).

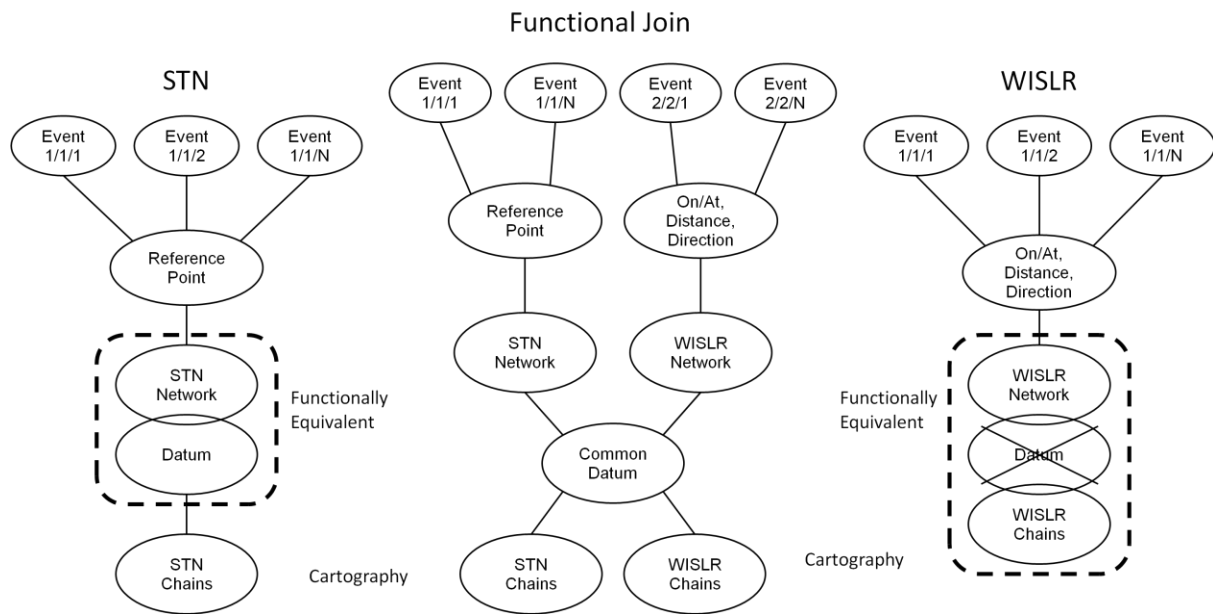


Figure 3. STN and WISLR conceptual model (Graettinger et al., 2008) and functional join example

2.3 Data Merging Issues

The solution to this problem is a merger between STN and WISLR that takes an approach unique to previous methods. The problem with transportation data sharing is typically caused by the

diversity of transportation data formats, which leads to inconsistencies and duplications. The major challenge is to develop a means of exchange between systems for not only the transfer of event data, but also completeness relative to updating and maintenance issues (Dueker and Butler, 2000). Some have devised approaches to merging spatial objects in the past, using both top-down and bottom-up approaches; however, these methods require some form of aggregation of the data (Sester et al., 1998). Others have developed scripting functions to translate the location of the data before mapping occurs (O'Neill and Harper, 1997).

The solution is the construction of a relationship between topological objects of the STN and WISLR system, respectively. This relationship is defined and built upon within a hand and custom tool populated table that describes links in either system by ID, start measure and end measure, so that each record of the merge table relates a section of one system to its corresponding link length in the other system.

3. Methodology

The previous section discussed LRSs and various issues inherent with data sharing between different transportation GISs. The merge method chosen and implemented for this project was a link-to-link method, where a section of a link in one system is matched to a section of a link in the other system. These sections of links are then placed into a table called the link_link table. This table was populated on a county-by-county basis to reduce the amount of simultaneous data processing, and also to segment the population among several people, so that data could be easily pieced together after completion. This section will detail the methodology of populating the link_link table, discuss issues that arose and the mitigation of these issues, and also describe the process of error-checking the table for each county.

3.1 STN and WISLR Systems

As described, the link_link table is simply a relationship of sections of GIS features in STN and WISLR. Figure 4 shows the STN system (orange lines) overlaid on the WISLR system (black lines).

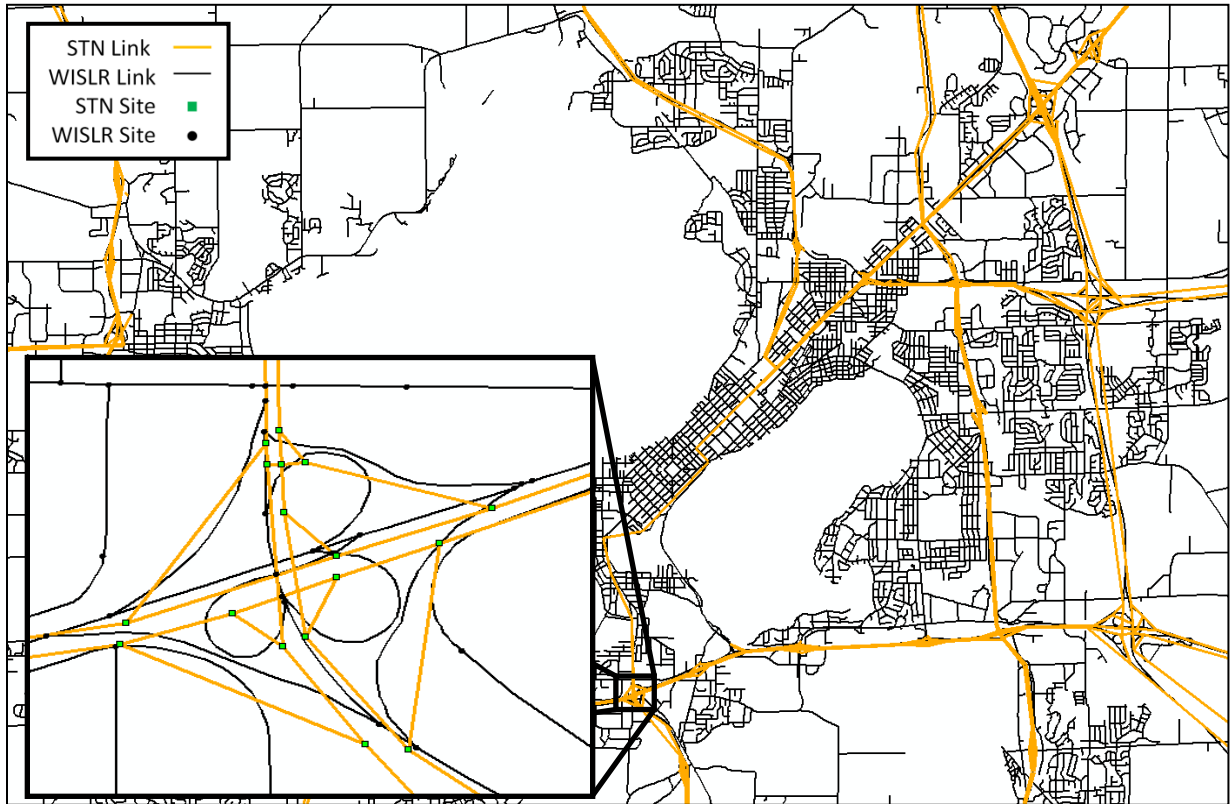


Figure 4. STN (orange lines) and WISLR (black lines) systems near Madison, Wisconsin

The structure of the STN system, shown in Figure 4, consists of links and sites (or nodes), along with cartographic chains (not shown). Each direction in STN is represented by a link that has an ID, length, from-site, and to-site. The from- and to-sites establish directionality and are stored in the attribute table; however, directionality is a function of topology that is established when the system was drawn. STN links are measured in thousandths of a mile, although links are only measured to a hundredth of a mile (e.g., the smallest length of an STN link is ten units). Notice that the STN links do not follow the cartographic representation exactly; rather, they are drawn via the STN sites as straight lines, and only represent connectivity between sites. The STN sites are shown in the inset of Figure 4 as green squares. Each site contains an ID number and an on/at description. The STN chains are the cartographic representation of state routes and are

utilized as a visual check since straight-line representations can reduce spatial accuracy between nodes.

The WISLR system, also shown in Figure 4, consists of black lines as the links, and black dots as the sites (or nodes). The WISLR links contain the same attributes as the STN links, though WISLR lengths are measured in feet rather than in thousandths of a mile. A key difference between the two systems is that WISLR links are a cartographic representation of the actual transportation features.

Contrary to the STN structure where each direction is represented by one link, WISLR does not follow this structure. WISLR was created by processing digitized lines and inserting a node at any location where a line ended or intersected another line. This process automatically created two links for every line, regardless of whether the link represented a one-directional or two-directional street.

Another significant difference between the two is the resolution of complex areas such as divided highway intersections, as shown in Figure 5; sometimes WISLR does not display key features of intersections such as median crossovers and turn-lanes.

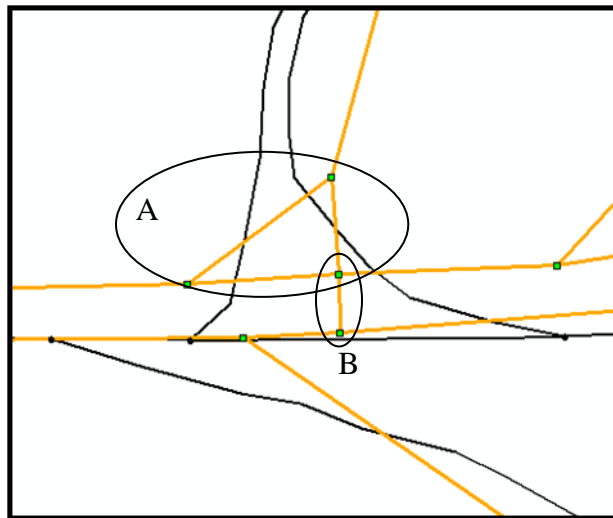


Figure 5. STN and WISLR are very different in logical design in this interchange area

In Figure 5, there is an east-west divided highway represented in STN as two approximately parallel sets of orange links, while WISLR represents this highway as one set of overlapping black links. This location is an interchange showing three on- or off-ramps. The off-ramp coming from the top of the figure at location A splits into two turn-lanes; however, WISLR does

not show this split. Another extra link that is present in STN but missing in WISLR is the vertical link in location B that connects the through-lane to the opposite side of the divided highway. This link is known as a median crossover and represents a situation in which a driver can cross the divided highway and make a left turn. WISLR represents the median crossover with a single node.

3.2 *Link_link Table Design and Population*

A key design element of the link_link table is simplicity without sacrificing functionality. As described before, the merge methodology is essentially matching sections of links in each system. Table 1 shows a layout of the basic link_link table. Each row in the table represents what is believed to be one segment of pavement in reality, but described in two separate LRSs. The result of the design process was a table containing six main columns.

Table 1. Names and descriptions of the six basic link_link columns

STNid	STNstart	STNend	WISLRid	WISLRstart	WISLR end
Unique identifier for the STN link	Start measure for the STN section	End measure for the STN section	Unique identifier for the corresponding WISLR link	Corresponding start measure for the WISLR section	Corresponding end measure for the WISLR section

The first three columns of the link_link table describe a portion of the STN system, while the last three columns describe a corresponding portion of the WISLR system. To demonstrate the link_link table population procedure, an example section of roadway is presented in Figure 6 and coded into Table 2. For simplicity, only one direction will be coded in the link_link table. In Figure 6, STN link A represents four WISLR links a, b, c, and d. Note that the STN sites are not always spatially close to the corresponding WISLR sites.

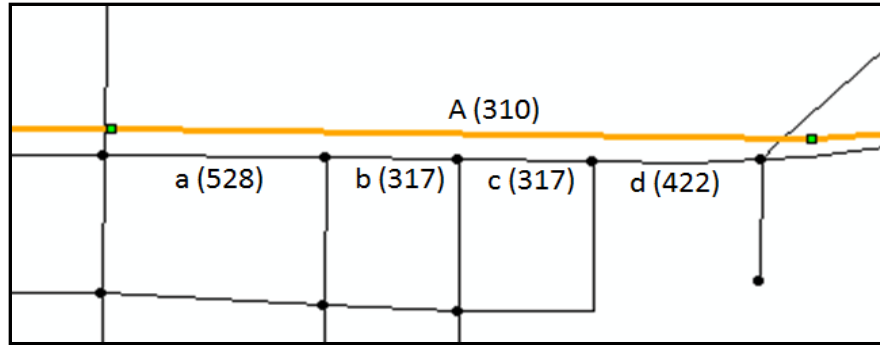


Figure 6. Typical link_link coding example

The STN link attribute table provides the STN link ID, which for this example is A. As shown in Table 2, the first STN start value is always zero, and last STNend value is the length of the STN link (310 units for this example), which is also in the STN link attribute table in the length field. The WISLRid column can be populated by extracting the link ID from the WISLR link attribute table. Given that the entire length of these four WISLR links are within the limits of the STN link, the WISLRstart values will all be zero and the WISLRend values will always be the length of the WISLR link, which is also in the WISLR link attribute table in the length field. This is the most common situation when coding, where a single record in the link_link table represents a single WISLR link.

Table 2. Example link_link table coded using ratio calculations

STNid	STNstart	STNend	WISLRid	WISLRstart	WISLRend
A	0	103	a	0	528
A	103	165	b	0	317
A	165	227	c	0	317
A	227	310	d	0	422

After entering the values that can be determined from the attribute tables, the next values to be entered are the STNstart and STNend values for the rest of the cells. These values essentially represent the location of the WISLR site on the STN link. A simple ratio calculation is performed to obtain these values (where STN part/STN full = WISLR part/WISLR full).

Notice when the ratio calculation is performed for STNstart and STNend values that these numbers are rounded to nearest thousandth of a mile, unlike the measurement found in STN, which is measured to nearest hundredth of a mile.

Another set of data in STN known as Access Points (black X's), shown in Figure 7, can be used in coding and calibrating the link_link table. Access Points provide references to local roads and other landmarks in the form of an STN link and an offset. Figure 7 displays the same STN and WISLR links as in Figure 6, and includes Access Points. At the request of WisDOT, Access Points are used when possible to aid in calibration and improve reliability. These Access Points are driven distances and contain the street names in a description field (note that the street names can be displayed on the WISLR links), which aids in determining which Access Point corresponds to which WISLR site.

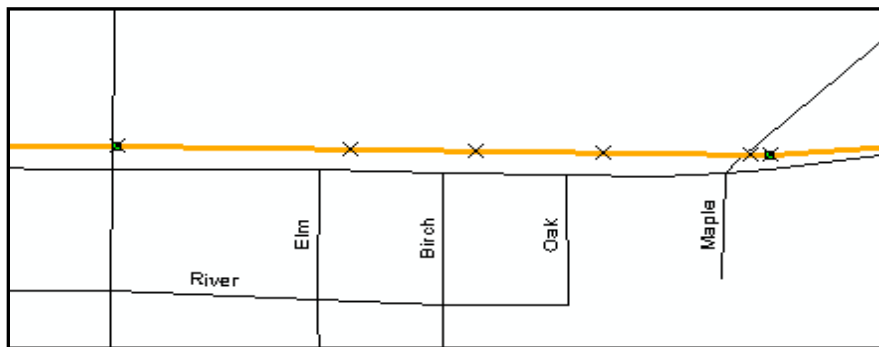


Figure 7. Access Points provide calibration for the link_link table

Repopulating Table 2 with Access Point information slightly adjusts the locations of the WISLR sites as shown in Table 3. The first Access Point was seven units from the ratio value, though the other Access Points were within five units of the ratio values. While this does not present a major difference in the final product, Access Points increase the confidence in the data by using driven distances to determine intersection locations rather than hand-calculated values.

Table 3. Link_link table coded using Access Points

STNid	STNstart	STNend	WISLRid	WISLRstart	WISLRend
A	0	110	a	0	528
A	110	170	b	0	317
A	170	230	c	0	317
A	230	310	d	0	422

3.3 Flag Columns

Coding the link_link table can be simplistic when STN and WISLR match up well, as shown in the previous example. However, as mentioned in Section 3.1 STN and WISLR Systems and shown in Figure 5, there can be significant differences in the representations of certain areas such

as intersections. Another cause of these differences can be areas that WISLR does not consider part of a road system, such as a weigh-station or a Park & Ride, which STN maintains. As shown in Table 4, a series of flag columns were conceived and included as a part of the link_link table. There are five flags: turn-lane flag, median-crossover flag, gore point flag, weigh-station flag, and the problem flag. Each flag uses one column in the link_link table, except for the problem flag, which also contains a “Comments” column.

Table 4. Flags columns and descriptions

Column Name	Column Description
Turn-lane	Turn-lanes represent straight and right-turning traffic lanes
Median Crossover	Median crossovers are representative of the distance to cross an intersection
Gore Point	Gore points occur at acute-angle pavement intersections
Weigh-station	Weigh-stations and Park & Rides are only maintained in STN
Problem	Problem flags are used to mark areas where STN or WISLR has errors
Comments	Comments about the problem flag are written in this column

3.3.1 Turn-lane Flag

Turn-lanes, as example of which is shown in Figure 8, are locations in STN where traffic flow splits into a through lane and right-turning lane. These links are not represented in WISLR; therefore, a flag is needed to mark the presence of this inconsistency.

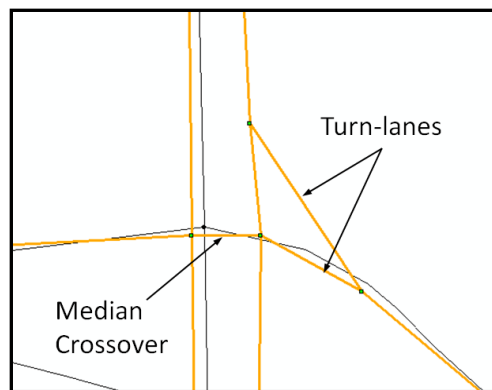


Figure 8. Turn-lane and median crossovers are not represented in WISLR by links

Turn-lane STN links are considered to represent a distance along a WISLR link; therefore, the turn-lane is coded from a location on the WISLR link to the end of that WISLR link. The distance used is calculated from the length of the STN turn-lanes. Typically, these links are ten

to twenty thousandths of a mile. The STN link length is multiplied by 5.28 to convert the length to feet and rounded to the nearest foot. This length is used as the WISLR segment in the link_link table. The link_link records for these STN turn-lanes are then flagged with a value of 1 in the “Turn-lane” column.

3.3.2 Median-Crossover Flag

Intersections are not point entities; rather, they cover an area that represents segments of pavement. However, logically, there are many ways to represent these intersections in an LRS. WISLR represents most intersections as nodes, while STN draws links to represent these same areas. The distance through an intersection (e.g. from the stop line to the edge of the median) is known as a median-crossover, an example of which is shown in Figure 8. Median crossovers are coded to a single point in WISLR (typically from zero to zero on the appropriate link, based on the direction of the STN link).

Median crossovers can occur when local roads or off-ramps cross a divided highway or “T” into a divided highway. In the example shown in Figure 8, the median crossover shows the ability to turn left or continue straight after traversing the link from east to west. In this case, the STN link could possibly go to two WISLR links, so a convention was created for this situation. The ID for the WISLR link that is not coded as the WISLRid value is placed in the “M” column. T-intersections are coded and flagged differently in the link_link table when compared to the through-street median crossover situation. With a T-intersection, only one WISLR link is available per STN link, so this median crossover is coded normally, and a value of 1 is placed in the “Median crossover” column.

3.3.3 Gore Point Flag

On-ramps and off-ramps are typically constructed with a sharp offset angle from the main road to allow the driver a safe and easy transition between the roads sections at a high speed. There is a beveled section of pavement that smoothes the transition between the pavement sections called a gore point, as shown in the aerial view of a gore point in Figure 9. STN and WISLR representations of this gore point are overlaid on the image. The inconsistency between STN and WISLR in this case is not visual; rather, the difference is in the method of measurement of the links connecting to this gore point. WISLR measures the ramp from or to the centerline of the

main road, whereas STN measures the ramp from or to the gore point. Gores can change lengths of roads significantly depending on the severity of the gore.



Figure 9. Aerial image of a gore point alongside STN and WISLR

A gore point can be imagined as a node attribute. Each link that connects to that node is going to be affected by the existence of the gore point. Gores are flagged with one of three different values: T, F, or B. The T (To) and F (From) values respectively indicate whether a link travels towards or away from a gore. If a gore exists on both sides of a link, the record is flagged with a B (Both) in the “Gore Point” column. The proposed idea is to adjust the STN or WISLR link by some factor on the side that connects to the gore point. Therefore, the flag was required to identify the location of the gore relative to the link in question.

3.3.4 Weigh-station Flag

Weigh stations, Park & Rides, and rest areas are maintained in the STN system, but are not represented in the WISLR system. The locations of these links must be arbitrarily assigned. Typically, rest areas and some parking lots will be drawn in STN, branching from a state route that is drawn in WISLR; therefore, the STN links will be coded to one point at this spot on WISLR. For example, if a parking lot is drawn as a single STN link, this link will be coded to the point on WISLR off of which the STN link branched.

3.3.5 Problem Flag

Problematic areas are not commonplace, but occur often enough to require modification to the systems. Problem areas are flagged with a value of 1 in the “Problem” column and a comment

describing the problem, including the source of the problem (STN or WISLR), is included. The link_link records that are marked as a problem will be assessed by WisDOT.

3.4 Link_link Coding Tool

In order to increase the efficiency of the coding process and reduce the amount of human error in the link_link table, a semi-automated tool was designed and developed. This tool, known as the Link_link Coding Tool, significantly increases the speed of coding simple areas (i.e. where problems are not present). Through coding experiments, the efficiency of the tool was determined to be approximately three times faster than coding without the tool.

The Link_link Coding Tool is only semi-automated because the coder is required to select links from the map. The tool then populates the link IDs and performs any needed calculations.

The user interface (UI) of the Link_link Coding Tool is shown in Figure 10. The tool removes the need for any calculation to be performed by hand. When all data is selected and entered into the tool, the tool will either calculate a ratio for the user automatically, or the user can select Access Points via the tool to be used in the link_link table.

As shown in Figure 10, one direction of the data has been entered into the bottom area of the UI, which shows a preview of the output. The data for the opposite direction is still in the boxes at the top of the UI, which is where the attribute data from the features is stored until the coder chooses to enter the data into the preview section (the relevant STN and WISLR sites must be selected so the program knows which links pertain to which direction). The “Calculate” button performs a ratio calculation when the data for both directions has been entered; however, if Access Points exist for the STN links, the Access Points will be used in lieu of a ratio. The coder clicks the “Commit” button to submit the records to the link_link table.

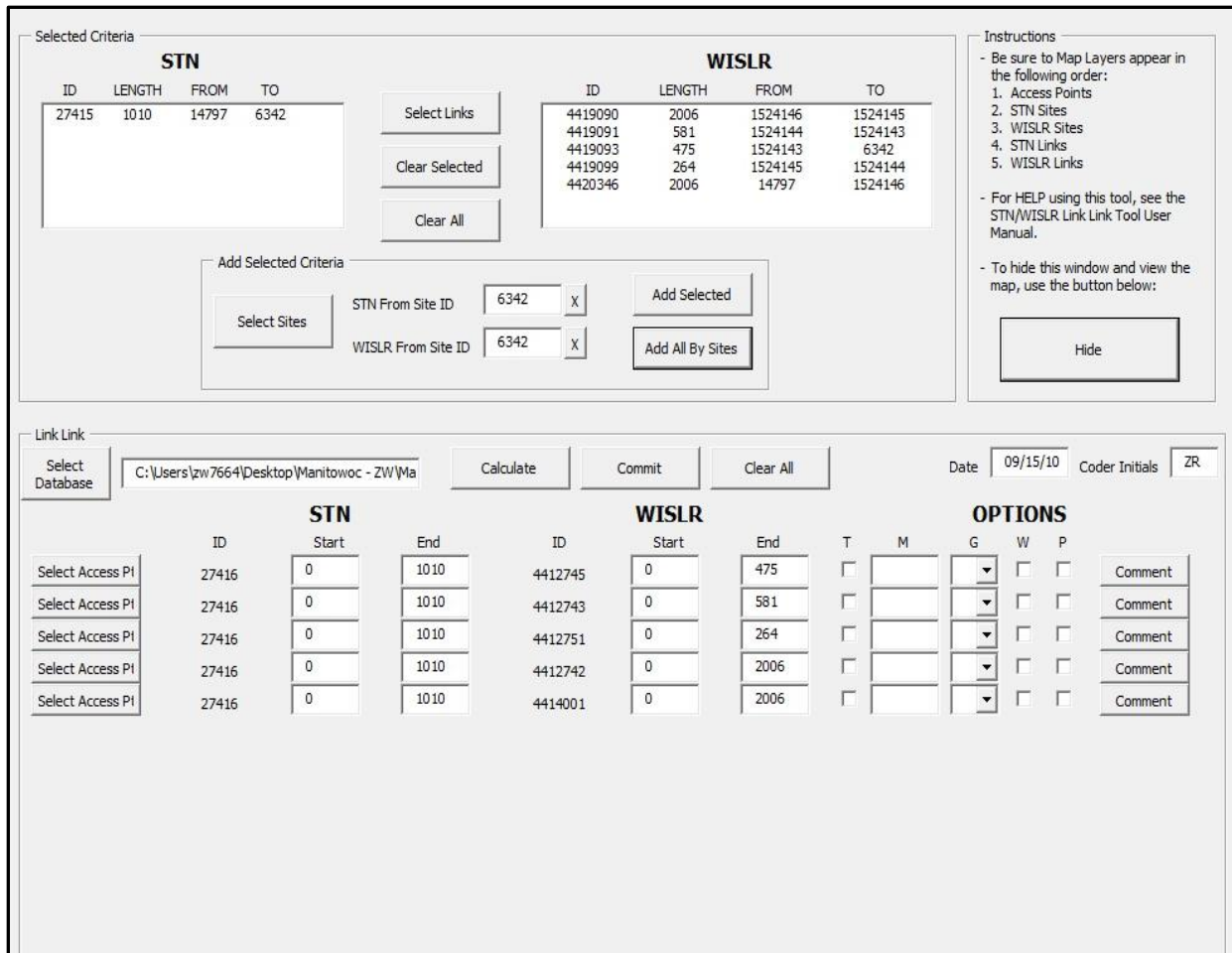


Figure 10. Link_link Coding Tool UI

3.5 STN Point Generator and the Point Moving Program

Two programs were written to facilitate data processing and testing of the link_link table: the STN Point Generator and the Point Moving Program. Creating and moving points is the most efficient and complete method for checking the quality of the link_link table. Moving data from STN to WISLR, which is currently performed by the Point Moving Program, ensures that the link_link table was populated correctly and is the primary objective of the LRS merge. These programs are used in the quality checking procedures that are described in Section 3.6 Quality Assurance/Quality Control (QA/QC).

The STN Point Generator uses an Excel spreadsheet containing the STN link IDs and lengths for a county to generate a STN points table. This table consists of a Unique ID, STN link ID, and

STN link offset. These offsets are from zero to the full length of the STN link, in hundredth of a mile increments.

The Point Moving Program uses the STN points table and the link_link to generate the WISLR points table, which is essentially equivalent to the STN points table in that the WISLR table contains a Unique ID (the program transfers the ID from the STN points table), a WISLR link ID, and a WISLR link offset. The program joins the tables based on the STN link ID in each table. The program reads an STN link ID and STN offset, and then finds which record of the link_link table contains that STN link ID and a STNstart and STNend range that encompass the STN offset value. Using the STNstart, STNend, WISLRstart, and WISLREnd values, the STN offset is translated into a WISLR offset value.

The STN points and the output WISLR points can be displayed in GIS along the STN and WISLR links, respectively, using a route event tool. The visualization of these points on the LRSs aids in quality checking and assures that the link_link table was coded correctly.

3.6 Quality Assurance / Quality Control (QA/QC)

One of the most important aspects of the link_link table population process is quality. WisDOT requested the utilization of a series of quality validation measures before the table was submitted. When merging two LRSs, confusion is occasionally unavoidable; therefore, QA/QC checks are an essential part of the population process.

After completing the initial population of the link_link table, a series of checks are applied by the coder. There are seven checks which include: STN Link check, WISLR Link check, Access Point check, Gore Point check, Point Moving Program Output check, WISLR Link Visual check, the XY Connector Line check, and the Reversed Link Identification Tool. There is also a program that finds erroneous WISLR links. Each check is briefly described in the following sections.

3.6.1 STN Link Check

The STN system contains a finite number of graphic elements that represent the interstate and state highways that exist in Wisconsin. All STN links in the state must be coded into the link_link table. This check is relatively simple, and only involves a table join within the GIS. The link_link table is added to the GIS from the Access database or Excel file, and is then joined

to the STN links attribute table for the current county. If there is a link in the attribute table that is not represented in the link_link table, null values will appear. These null values point the coder in the direction of un-coded STN links.

3.6.2 WISLR Link Check

The WISLR system contains a relatively large number of graphical entities that represent all roads in the state. Therefore, not every WISLR link will be used in the link_link. To assure that there are no mistyped IDs in the link_link table, every WISLR link ID used in the link_link table must be a valid ID. The check is performed by joining the WISLR links attribute table to the link_link table based on WISLR link ID, and then checking that there are no records in the link_link table that do not have a match in the WISLR links attribute table.

3.6.3 WISLR Link Visual Check

There is a large number of WISLR links in each county, but not all are used, as described previously. In order to ensure only correct WISLR links were used in coding the link_link table, a visual check is performed by displaying only the WISLR links that were used in the link_link table, so that any unnecessary links or gaps are easily visible to the coder.

Figure 11a shows a view of the STN and WISLR links before beginning this visual check.

Figure 11b shows only the STN links and the WISLR links that are included in the link_link table.

The removal of the unused links can be performed within ArcMap by loading the link_link table into the program, summarizing the WISLR ID column, then joining this to the WISLR links shapefile. Any link that is not used will have null values in key columns and can be removed from view.

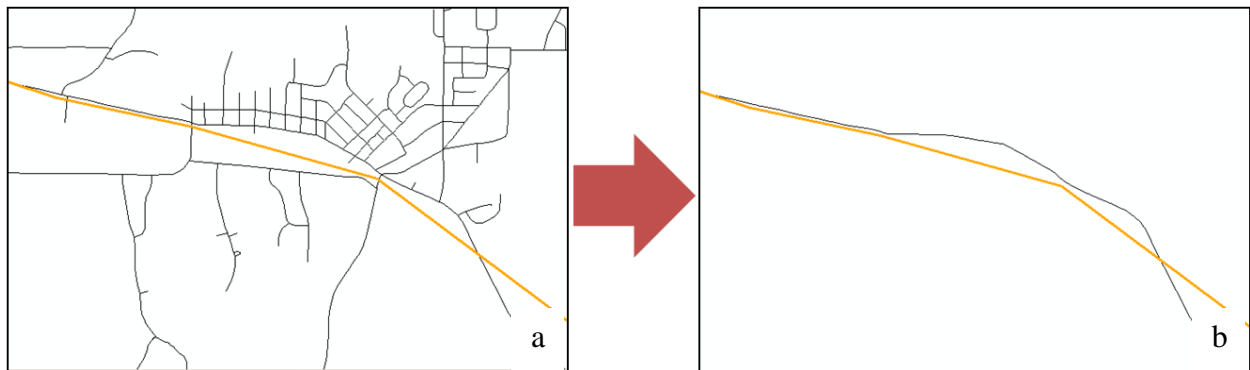


Figure 11 a) All WISLR links shown; b) Only WISLR in the link_link table are shown

3.6.4 Access Point Check

Access Points that describe an intersection in WISLR should be included in the coding of the link_link table. After the link_link coding process is complete, a check identifies all of the unused Access Points within the county. The coder then verifies each unused Access Point to guarantee that unused points were not mistakenly overlooked.

The Access Point check is performed by creating a unique column in both the link_link table and Access Point attribute table. This unique field is the link ID concatenated with the offset distance. The tables are then joined based on this column, and all unused Access Points appear as null values in the link_link table.

3.6.5 Gore Point Check

Gore points are typically a subjective aspect of coding that must be analyzed on a case-by-case basis. Typically, gores at intersections are easy to spot, but can sometimes be overlooked. This check assures that the coder looks at areas that typically contain gores and verifies that all links that should be marked in the “Gore Point” column are correctly marked (gores will be flagged with either a To, From, or Both flag, depending on the location of the gore with respect to the link).

As shown in Figure 12, for a single direction of travel, a gore point will always have at least three associated links. Each of these links will be coded with the correct letter in the “Gore Point” column (e.g. if a link is shown as traveling to a gore point, the correct flag is a To flag). Each record in the link_link table typically represents a whole WISLR link; therefore, the link_link table can be successfully joined to the WISLR links shapefile based on WISLR link ID.

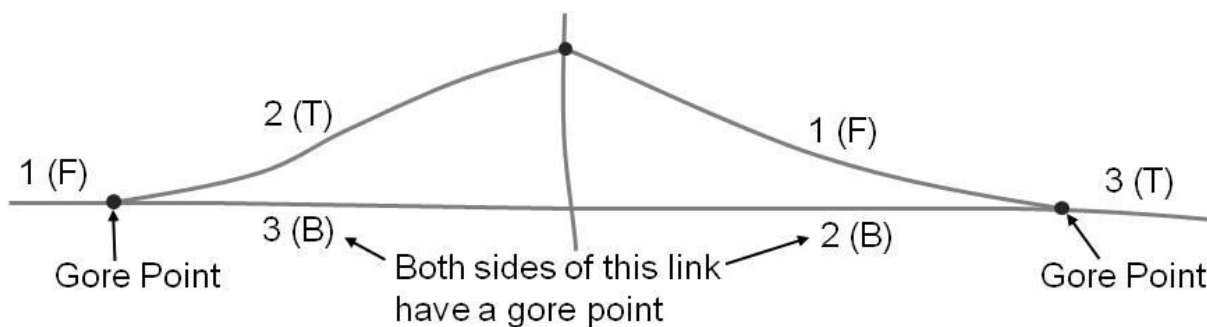


Figure 12. Gore point at an off-ramp

The symbology of the links is changed so that specific values in the “Gore Point” column are represented by changing the WISLR links into different colors, creating a visual check through which the coder can quickly evaluate and makes changes as necessary.

3.6.6 Point Moving Program Output Check

The Point-Moving program, which moves crash points from STN to WISLR, is executed as part of the quality control procedures. The input is an Access Database containing the link_link table and a table of STN points. The STN Points table can be any number of points that would be on one of the links coded in the link_link table. A unique ID number is provided for each STN point before the program is executed, and this unique ID is transferred to each corresponding WISLR point. The program reads the STN point, calculates where each point moves to in WISLR, and then creates a WISLR point table that contains one point for every STN point. A check is performed to verify that the number of input and output values is the same. If these numbers differ, then one or more records in the link_link table are incorrect.

3.6.7 XY Connector Line Check

The XY Connector Line check is a visual check that illustrates the output from the Point Moving program in a way that allows mistakes and improper coding to be displayed in an obvious manner. Each input point for the program contains a unique identifier that is carried over to the output. Using this identifier as a common field, the input and output can be joined together. Coordinates are calculated using the ArcMap Geometry Calculator to provide starting and ending locations for each set of points. These locations then have a line drawn from start to end using a set of tools known as Hawth's Analysis Tools for ArcGIS (Beyer, 2004), a free toolset that works within ArcMap.

An example of the connector lines when there are no problems is shown in Figure 13a. The lines are parallel and evenly spaced throughout an entire STN link. Figure 13b shows an example of a WISLR link that is drawn in reverse direction. Since the routes are created from the WISLR cartography, the links must be drawn in the direction of travel. The attribute data may be correct, but visually, the end of the link is the start and vice versa. This is apparent in Figure 13b where the connector lines cross one another; the STN starting point is transferred to the WISLR end point and the STN end point is transferred to the WISLR starting point.

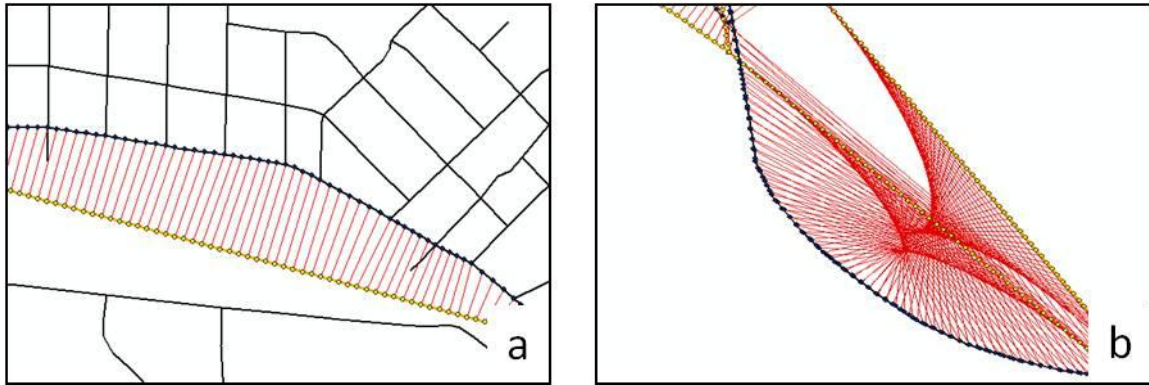


Figure 13. XY connector lines showing correctly and incorrectly coded links

The reversed WISLR links are only evident when this QA/QC check is performed; however, this does not constitute a problem in the link_link table, only a problem that must be addressed in the “Problem” flag column. This check is performed an average of three times during the QA/QC of the coding of a single county.

3.6.8 Reversed Link Identification Tool

This tool was developed to automatically determine if WISLR links are reversed as discussed in the previous subsection. This tool uses the cartographic lengths of the STN and WISLR links, as well as the lengths of the connector lines to find where the connector lines are not displaying in a parallel manner. Much of the information used by this program is calculated at the beginning of the XY Connector Line check, therefore it is advantageous to perform these checks simultaneously, so that problems may be discovered before the full visual check is performed.

A generic example of correct and incorrect XY lines is shown in Figure 14. Notice the considerable difference in lengths between the matching lines in the correct and incorrect illustrations (where line 3 is the exception). This difference is how the tool determines which links may be reversed.

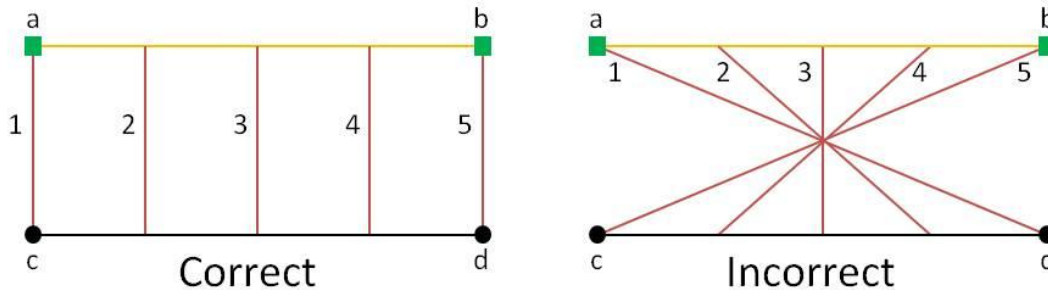


Figure 14. Simplified examples of correct and incorrect XY lines

Two checks are performed by the program. The first check looks at the distance between the STN start and the WISLR start, which, in the example in Figure 14, is line 1, and compares it to the graphical length of the STN link. This indicates a reversed WISLR link by showing that line 1 is longer than link a-b.

The second check method compares the length of line 1 to the distance between nodes a and d, as well as nodes c and b. If the length is approximately equal, a problem with this WISLR may exist. The WISLR link ID numbers are flagged as incorrect and output to a table that can be added to ArcMap and analyzed. After making changes where necessary, the Point Moving program is executed again and the XY Connector Line check is performed again.

4. Pilot Study

Prior to implementing the link_link coding process for the entire state of Wisconsin, a pilot study was performed on Dane, Iowa, and Grant counties. The objective of the pilot study was to develop and confirm the procedures, conventions, and methodologies for coding the link_link table.

Another objective of the pilot study was to present the ability of the link_link table to move actual data. A crash database was used to obtain STN crash data points that were moved with this system to WISLR. The crashes contain an accident number (ACCDNMBR) value, which was utilized as the Unique ID. Each crash record also contained a STN link ID and link offset, providing the essential columns for the program to run.

All crashes for 2006, 2007, and 2008 within the pilot counties were processed to create the STN points table to be used in the Point Moving Program. 14,455 crashes exist for these three

counties; however, several of these crashes occurred on retired STN links. Therefore, the 12,722 current crashes were used for the pilot study.

5. Results of the Pilot Study

The pilot study counties are shown in Figure 15. The crashes for the three counties are displayed on the STN links. 100% of the 12,722 crashes were moved from the STN system to the WISLR system.

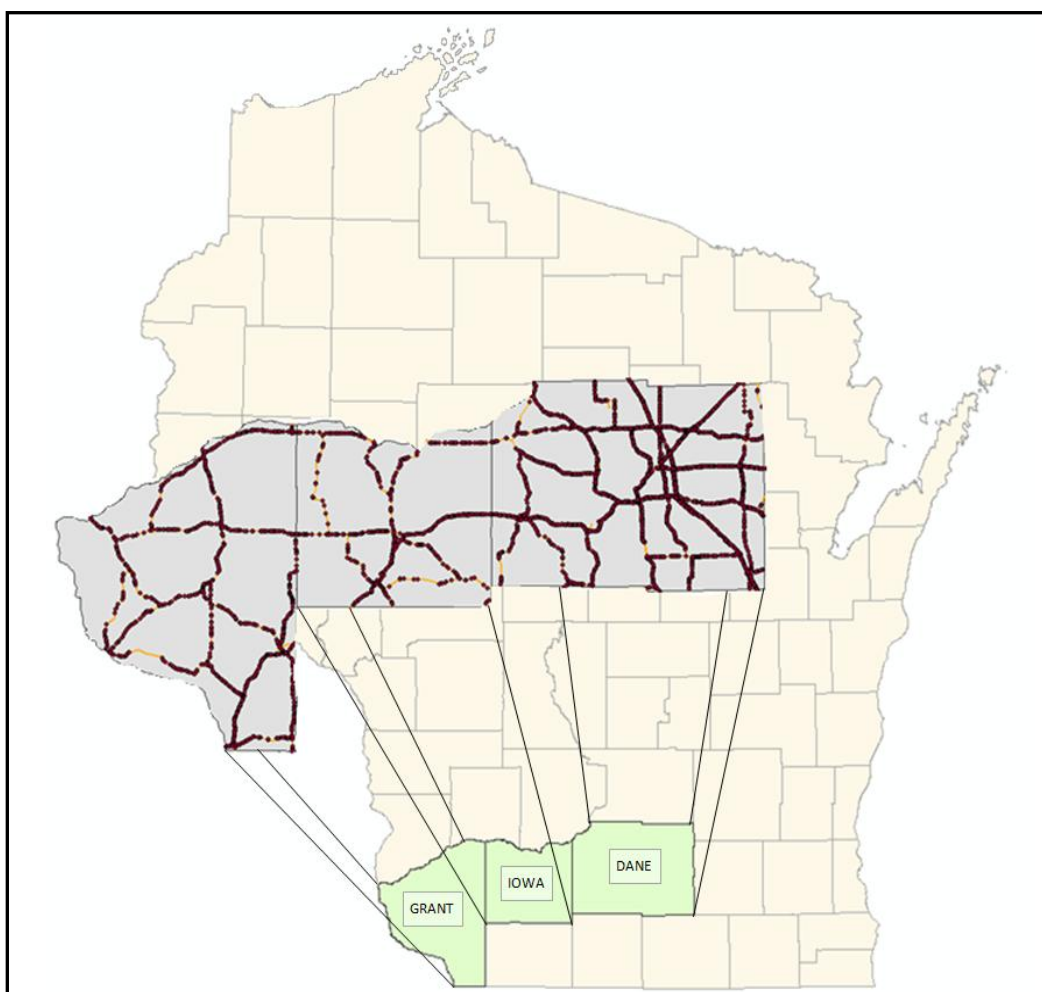


Figure 15. Pilot study area with RP crashes displayed

The pilot study demonstrated that the link_link table coding procedures can be successfully implemented throughout the entire state of Wisconsin.

6. Conclusion and Future Work

6.1 Conclusion

Though several studies have been performed to develop and analyze structures for transportation data sharing, this project presents a real application and data sharing methodology that has been successfully implemented using two robust LRSs, without requiring unnecessary changes or suspension of the currently functioning systems. The pilot study showed that after coding and QA/QC was performed, 100% of crashes could be transferred from one system to the other, essentially combining two logically different LRSs.

6.2 Future Work

An interesting notion to consider in the future is the ramifications of attempting to transfer data not only from STN to WISLR, but from WISLR back to STN. Because STN contains far more detail at intersections, data that occurred on a WISLR link around these intersections may have multiple possible destinations in STN. Future work needs to be performed to determine the best method to move data from a low-resolution system (WISLR) to a high-resolution system (STN).

Another factor that must be considered in the future is the retirement of STN and WISLR links. Links are annually updated and several STN and WISLR links become historic. A program or methodology should be developed that can use WisDOT's link-history table to reconcile the changes that must be made to the link_link table. This will increase the usability of the link_link table, regardless of future system updates.

7. References

- Beyer, H. L. 2004. Hawth's Analysis Tools for ArcGIS. Available at <http://www.spatial ecology.com/htools>.
- Dueker, K. J., and J. A. Butler. 2000. A geographic information system framework for transportation data sharing. *Transportation Research Part C-Emerging Technologies* 8(1-6): 13-36.
- Graettinger, A. J, Qin, X., Spear, G., Parker, S. T., and Forde, S. 2008. State and Non-State Network Mapping Integration. *Proceedings of the 2008 Mid-Continent Transportation Research Forum*, Madison, WI.
- Graettinger, A. J., X. Qin, G. Spear, S. T. Parker, and S. Forde. 2009. "Combining State Route and Local Road Linear Referencing System Information" *Journal of the Transportation Research Record*, Vol. 2121, p. 152-159.
- Miller, H.J. and S. L Shaw. "GIS-T Data Models," Oxford University Press, 2001.
- O'Neill, W. A. and E. A. Harper. 1997. Linear location translation within GIS. *Proceedings of ESRI User Conference*.
- Scarponcini, P. 2002. Generalized model for linear referencing in transportation. *GeoInformatica* 6(1): 35-55.
- Sester, M., Anders, K., Walker, V. 1998. Linking objects of different spatial data sets by integration and aggregation. *GeoInformatica* 2(4), 335-358.
- Vonderohe, A. P., Chou, C.L., Sun, F. and Adams, T.M. 1997. A generic data model for linear referencing systems. *NCHRP Research Results Digest 218*, National Cooperative Highway Research Program, Transportation Research Board, Washington, DC.
- Vonderohe, A. P., and T. D. Hepworth. 1998. A Methodology for Design of a Linear Referencing System for Surface Transportation, Final Report. Project AT-4567, Sandia National Laboratories.
- Wisconsin Location Control Management Manual. 1995. Wisconsin Department of Transportation, Madison, WI.