WisDOT Traffic Operations Infrastructure Plan

Appendix D – Data Processing

Bureau of Highway Operations
Wisconsin Department of Transportation

May 2008

For additional information, please contact John Corbin at john.corbin@dot.state.wi.us.

Table of Contents

List of Tables

List of Figures

1.0 Data Collection, Analysis, and Synthesis

This Appendix outlines the process by which the Meta-Manager data was utilized for the Traffic Operations Infrastructure Planning (TOIP) effort. Specifically it details the calculations needed to created the scoring process.

1.1 Data for Traffic Operations Infrastructure Planning Methodology

An important objective of the Study Team was to ensure that the data driving the TOIP process was easily accessible from standard WisDOT sources. This not only makes the process as straightforward as possible but also helps ensure that the methodology can be revisited and updated easily in the future as part of ongoing long-range planning activities.

To meet this objective, the study team worked closely with the WisDOT Program Development and Analysis Section to find the most relevant data that met the objective discussed above. Based on their input, the TOIP methodology was designed to run on three distinct datasets, the primary one being Meta-Manager data.

1.1.1 Meta-Manager Data

Meta- Manager, a comprehensive data repository for WisDOT, was developed by Division of Transportation Investment Management's Bureau of State Highway Programs to meet the data requirements for a variety of needs and performance analyses. The Meta-Manager Management System data is the best information currently available for evaluating system needs and measuring program impacts and serves as the major source of data for the TOIP effort. The data are currently maintained by the Program Development and Analysis Section whose assistance was critical to the development of the TOIP effort.

The Meta-Manager database is an excellent resource for assessing system condition, analyzing need and performance and supporting project development. Since the database is developed from several WisDOT databases and provides greater access to that data, users provide helpful feedback regarding the quality of the original corporate data.

The Meta-Manager geographically integrates a variety of data including pavement information, system deficiencies, safety, congestion and other information. The data also include future projections of physical condition data.

1.1.2 University of Wisconsin TOPS Lab Weather Data

Since the TOIP methodology is designed to provide guidance for operationally related projects along a corridor, operationally related data outside meta-manager is also needed to drive those decisions. Weather was one of the 10 criteria selected as part of the methodology and it was therefore critical to find an appropriate weather related dataset. The weather data proposed for use in the TOIP methodology was processed by the University of Wisconsin TOPS Laboratory and documented in the paper Application of Road Weather Safety Audit to the Wisconsin Highway System, (Qin, Noyce, Martin and Khan).

1.1.3 Wisconsin Event Data

Special events and their impact on the transportation system was another criterion developed as part of the Sketch Plan methodology. For this dataset WisDOT Bureau of Traffic Forecasting provided the Study Team with a list of the top eighty six special events around the state.

The remainder of this section details how the Meta-Manager, weather, and event data were utilized.

1.2 META-MANAGER ROLL-UP METHODOLOGY

The rollup process was developed to reduce the number of links representing any given corridor by aggregating series of consecutive links with similar characteristics.

This subsection describes the tools and techniques used in this process.

Several terms are presented initially here for reference:

- Sketch Plan link– a single link in the result data set that is comprised of one or more links from the Meta-Manager links database. The combined links are physically consecutive, on the same route and direction, and enable all of the criterion data important for this project can be properly aggregated.
- The Meta-Manager links comprising a Sketch Plan link are called the **member links**.
- Split Criterion – the decision parameters that determine when the “end” of a Sketch Plan link has been reached. As the process considers adding the next physically consecutive member link to the current link being rolled up, these criteria determine whether that link should be allowed to accumulate in the roll up, or if it should start the next rolled up link. The split criterion are outlined Section 1.2.2.

1.2.1 Data Preparation

The process relies heavily on the ability of the system to correctly traverse the Meta-Manager links in route, direction, and then physically consecutive order. Specifically, the system needs to be sure that the "next" link in the process is:

- On the same route,
- In the same direction, and,
- the one that starts at the same physical location as the current one ends.

It was not assumed that the order of the Meta-Manager links in the database would be as required, though in large part, the Meta-Manager data is in order on any given route and direction.

When the links progress through the rollup process, the three attributes above can safely be used as split criteria and the resulting data sets will consist of links that have unique combinations of route and direction. This is specifically intended to assist those who will be working with this data to follow the routes along a corridor.

The system goes through a process of reading through all the links in a corridor, ordering them and confirming their location through a test process.

In order to facilitate the ordering of the data in the rollup process, the Meta-Manager links are first passed through an ordering process and then inserted into a MySQL database in the order from which they are retrieved.

The route and direction values for each link are obtained from the HWYADIR field in the Meta-Manager data. This field contains three characters of route information and 1 character of direction information. It was noted that this data is very clean, and that there were no missing directions and no inconsistently provided route numbers (29 vs. 029, etc).

The physical location of each link was obtained via the GIS shape file which accompanies the Meta-Manager data. The order of the shapes within the shape file is the same order of the data in the Meta-Manager data . Thus, by traversing the shape file and the Meta-Manager data file at the same time, the location of every link was known.

Many of the shapes for a given link are represented with more than two points in the shape file. However, only the first and last point for each link shape were saved with the links.

Because these link end-points are represented with real numbers, when comparisons are done to determine any proximity to a point, it is done by checking that the comparison of two numbers are within a certain tolerance. This guards against errors in comparisons when one value may be a number such as 25.99999, and the other 26 – which is a situation often occurring with real numbers. Several different values were tried for this tolerance number, however, the Study Team found that one foot works fine.

After reading all links, the data preparation process then goes through all routes discovered in the data set, then through all directions discovered on that route, and saves the first link on that route and direction found and then saves the rest of the links (on that route and direction) in consecutive order – that is, it will only save a link in the database when that link's start point equals the last saved link's end point.

The process is repeated until all links are in the database, and results in 20,606 links.

1.2.2 Rolling up links

Aggregation

Once the data are used to follow each corridor and implement the three primary split criteria (route, direction, consecutive), the links can be combined into longer, representative links, thus providing fewer links to analyze.

The primary concern in combining links is to ensure that the data represented on the Sketch Plan link are properly aggregated from the member Meta-manager links. For example, it is not appropriate to take the average accident rate of the member link. Instead, it is more accurate to calculate the number of accidents for each member link (using the link's accident rate, AADT, and length), and accumulate this number. When the split criteria are applied and a new sketch Plan link is closed off the number of accidents is then used to calculate the accident rate for the rolled up link, this time using the total AADT and total length .

Fields that are aggregated in this process include:

- Accident Rate (RATE) – Weighted (by VMT) average of member accident rates. Note, the AADT used to calculate VMT is that for 2007 (AADTYR_1)
- Severity Index (SEVINDX) – This is a simple accumulation. It is understood that this value is actually the sum of a series of "counts" of certain types of accidents times a point value for each type of accident. As such, the proper accumulation is to simply add the value among multiple links.
- Percent Trucks (TRKDYR_1 and TRKYR_1) - Weighted average of member percent trucks
- AADT (AADTYR_1, AADT2030) – Recalculated by taking accumulated VMT (AADT * length) and dividing by total length
- Level of Service (LOSYR_1,LOS2030) – This value is accumulated as a simple average. However, the average itself is not strictly intended to be used in analysis. Instead, LOS values A – F are used and as such – all rolled up links will contain a constant value of A – F. This is an important split criterion and so that when LOS, as a letter, changes among links –a new roll up is initiated.
- Length (PDP_MILE) – The accumulated total length.

- Intersecting Street Name (INTS_NM) – This value is taken from the first member link that has a non-blank value in the field. It is intended to assist in locating the link on a map if necessary. However, because it may be contributed by a member link that occurs some distance from the start of the Sketch Plan link – it will not always name a crossing facility that is near the start.
- Other fields –It is noted here for reference that other fields in the Meta-Manager data that appear in the Sketch Plan link take their values from the first member link. Rather than delete these data, it has been kept in place should some manual reference need to be made back to the Meta-Manager data .

Split Criteria

The split criteria dictate when one Sketch Plan link ends and the next one begins. Many approaches were tried with the primary goal of keeping the aggregated data on the rolled up links valid, accurate and relatively homogeneous .

The roll up process allowed fewer links to be analyzed in the deployment threshold scoring process; as opposed to looking at each of 500 links along a corridor, for example, it would be easier to look at 50 "representative" links.

It was initially proposed to standardize the rolled-up links based on length, recognizing that there would be variation depending on roadway characteristics, traffic volumes and adjacent land uses. However, it became clear as work progressed that this was not an effective method defining TOIP links. Specifically, we find that there are interesting dynamics occurring in the Meta-Manager links, sometimes over short distances. As a result some of the rolled up links are shorter than what would be preferred, but are much more accurate in representing travel characteristics and the reality that they often vary significantly over a short distance. This variation will result in different levels of need for ITS and operational projects.

This results in some links that are short and for which there are only 1 or 2 Meta-Manager member links. Except for one minor case – these "short" links are necessary to meet the objectives and data requirements of the TOIP methodology. The exception noted is in transition areas where a two parallel links in opposite directions differ in length.

Special adjustment must be made when there is a short link parallel to longer links in the other direction.

Most of the Meta-Manager data (for non-divided facilities) is provided in one "direction" only. Here, direction means East or West, or North or South, but not that the facility supports one or two way travel. The exception to this is, of course, Freeway or Expressway system (in most, if not all cases).

In some cases, when a short, two or three mile, section becomes divided, there appear two links in the Meta-Manager data at that physical location. This is, of course, as expected. However, because it is necessary to traverse the links in route and direction order – the system doesn't encounter this part that has become divided out, until a later in the process. Thus, causing that short link (or set of MM links) to become it's own rolled up link.

This was considered at great length and ultimately it was agreed that the fact that the division was made in the Meta-Manager data already points up that the performance characteristics of the two sides are each important in their own right. Indeed, we also know that often opposite sides of such divided sections do have distinct characteristics. Therefore, after close inspection of this issue – we decided that it would be best to leave these links in the system.

Criteria for Defining TOIP Links

The entire set of criteria for splitting links is as follows:

- Route and direction – when either route or direction changes, a new link is started. In some cases, two routes may merge into one. When this happens and the Meta-Manager route name field takes it's value from the one merging in, we end up with a possible "early" split. However, this situation is not detectable in software and is believed to happen at most a few times in the entire system.
- Number of lanes –Number of lanes is an important roadway characteristic that will be useful in threshold scoring. Because we don't have a good way to aggregate this value, and averaging the number of lanes would be invalid, we must therefore split on it.
- Level of service Letter designation – Level of service is an assigned value for the Meta-Manager links. We can neither calculate or accumulate it. We investigated whether some relationship could be found amongst the other data at hand which would allow us to accurately accumulate or calculate it. However, we could not find a suitable relationship. As a result, this criteria is actually the driving criteria for splitting TOIP links. In other words, most new links start where a value in LOS has changed significantly. (Initially, we looked at splitting where the numeric LOS value (1 – 6+) changed by a certain percentage. However, the LOS letter designation became the data used in threshold scoring analysis and the split was changed to occur where this letter changes.)
- Seasonal Factor Group – This value is relatively constant through large sections of most corridors. However, it sometimes changes suddenly for a small number of links within a corridor. To minimize this effect (causing a large number of small links), we allow a limited number of member links to continue accumulating after the point of this change. Thus, a Sketch Plan link may comprise a few member links of a different SFG – ultimately, however, the SFG change in that region is correctly captured.

- Total length – when the accumulated length reaches beyond a certain designated length, the next Sketch Plan link is started. The length specified, however, is very long (100 miles) and this rarely becomes a split point due to the presence of other factors.
- Contiguity – when the start point of a given link is further away than a specified value from the end point of the (current) last link in a rolled up link, a new Sketch Plan link is started. This proximity is currently represented by a distance of 10 feet .

Results

There are several products of the roll up process besides the TOIP links themselves. This includes cross reference information that can be used to relate the rolled up links back to their member Meta-Manager links and vice-versa.

The output files of the process contain the following:

- The rolled up links. Each corridor is represented by its own set of files. These include:
    - <Corr>.dbf – a dBase file containing the rolled up links. This data is a subset of the Meta-Manager data set. Except for fields we've added for our purposes, the format of this file is exactly that of the raw Meta-Manager dBase file.
    - Here <Corr> represents the root name of the file, which is an abbreviated form of the corridor name. A table is presented below which links these names to corridors .
    - <Corr>.shp, <Corr>.shx – the ESRI format shape file containing the geographic information of the links in the <Corr>.dbf file – in the same order.
    - <Corr>_Raw.dbf – a dBase file, in the same Meta-Manager format, containing all the links in the corridor directly from the Meta-Manager main file.
    - <Corr>_Nodes.dat, <Corr>_Links.dat – Generated node Ids and coordinates, as well as link information for use in importing the corridor links into IDAS (ITS Deployment Analysis System) which was used extensively in the development and validation of the rollup process (to visualize the created links)
- Summary information about all links:
    - The xRef.dbf file contains one record for every Meta-Manager link (20, 606 records). This contains the Meta-Manager's id value (META_MANAG) and the id of our Sketch Plan link (which is the Segment ID and is a contiguous integer value starting at 1)

- All.dbf contains all (rolled up) links from all corridors. It is essentially the concatenation of all the individual corridor files. Note that more than one section in the file may have any given route and direction. All other corridor link files can be considered to be "sorted" by route and direction.
- Results.dbf contains the statistics of the rollup process. Primarily, how many links are in the corridor in the raw data, and how many links are in the rolled up data.

The results of the rollup process are indicated in the following table. Note, since the process was concluded before the corridor names were changed by the department they differ slightly. The correct data sets were used in the final analysis.

Table 1.1 Roll-up Results

	CORRIDOR	# in Source	# in Result	% Change
ALP	Alpine Valley	204	64	69%
BAD	Badger State	543	110	80%
BLA	Blackhawk	377	145	62%
CAP	Capitol	766	299	61%
CHE	Cheese Country	198	55	72%
COR	Cornish Heritage	327	92	72%
COU	Coulee Country	211	54	74%
CRA	Cranberry Country	133	45	66%
DOO	Door Peninsula	96	35	64%
84T	84th Division Railsplitters	97	47	52%
FOX	Fox Valley	512	151	71%
FRA	Frank Lloyd Wright	207	61	71%
FRE	French Fur Trade	98	28	71%
GEN	Geneva Lakes	129	53	59%
GOP	Gopher Connection	476	149	69%
HIA	Hiawatha	437	180	59%
IND	Indian Head Lakes	242	63	74%
KET	Kettle Country	108	34	69%
LAS	Lake Superior	183	47	74%
LAT	Lake To Lake	130	46	65%
LUM	Lumber Country Heritage	149	32	79%
MAR	Marshfield - Rapids Connection	233	77	67%
MIS	Mississippi River	441	115	74%
NOR	North Country	331	72	78%
PEA	Peace Memorial	381	72	81%
PES	Peshtigo Fire Memorial	122	30	75%
POT	Potato Country	236	73	69%
POW	POW/MIA Remembrance	178	38	79%
ROC	Rock River	174	73	58%
SOU	Southern Tier	279	106	62%
TIT	Titletown	495	124	75%
TRE	Trempealeau River	116	26	78%
WAU	Waukesha Connection	210	82	61%
WIL	Wild Goose	304	99	67%
WIS	Wisconsin Heartland	446	61	86%
WIV	Wisconsin River	496	93	81%
WOL	Wolf/Waupaca Rivers	180	45	75%

All the raw Meta-Manager, weather and event data, detailed in the previous sections, has been complied into a single dataset within an excel spreadsheet for portability and to allow for easy analysis. This spreadsheet also contains the weights and thresholds discussed above and in the main document. These two elements constitute all the information needed to execute the TOIP methodology. Since there are over 50,000 TOIP links in the dataset, an automated macro was developed within Excel which will score al the links on all the corridors.

This automated process also produces the xReff files needed to illustrate the results in GIS utilizing the display standards presented in the final document. The GIS related output files are produced for each corridor and reside within the spreadsheet as well. The native Meta- Manager links are also provided within the spreadsheet should further analysis be needed at this basic link level. It should be noted that this approach allows the capability to modify the methodology (including weights and thresholds) should sensitivity analyses be required or if in the future these elements need to be modified based on changing conditions. Finally, a GUI has been developed that allows the user to select a specific corridor, execute the methodology, and receive summary statistics. A draft of this GUI is presented in Figure 1.2. This macro has been delivered to the WisDOT and is being maintained by the TOPS Lab.

Figure 1.1 TOIP Methodology GUI