Effect of spatial data aggregation on highway safety analysis

by

Justin Cooper Jackson

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee:
Reginald Souleyrette, Major Professor
Thomas Maze
David Plazak

Iowa State University
Ames, Iowa
2006

UMI Number: 1439923

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

ACKNOWLEDGMENTS

First I would like to thank my major professor, Dr. Reginald Souleyrette, and my committee members, Dr.Thomas Maze and Mr. David Plazak, for allowing me to draw from their knowledge and expertise. Without them I would have never been able to complete my thesis.

I would also like to thank Mr. Zach Hans for all the help and guidance that he provided me in completing my thesis. Also the staff and students of the Center for Transportation Research and Education for the knowledge and support they provided during this project.

Finally I would like to thank my family and friends for giving me support during this journey. Without them I would not be who I am. Again thanks to all who have helped me get to where I am and for their continued support on the rest of my endeavors.

CHAPTER 1 INTRODUCTION

There are 43,000 fatalities that occur on the nation's road system each year (FARS 2006). In Iowa alone there are approximately 425 fatalities and 60,000 crashes per year. Reducing the number of crashes and injuries is partially the domain of highway safety engineer. There are four basic strategies, which can be applied to improve highway safety; engineering, education, emergency response, and enforcement (4 E's) (FHWA 2006b). In order to determine which strategy is most appropriate, locations must be identified that are in need of safety improvements.

A typical approach to highway safety is to identify locations with a high, disproportionate number or severity of crashes. Once the high crash locations have been identified, prioritization of these locations allows for the potential of maximizing safety funding by identifying which locations will have the greatest benefit for reducing the number and severity of crashes. Also states are required by the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) to identify the top 5 percent of locations on the primary system with the highest number or severity of crashes (FHWA 2006a).

Previous studies have analyzed the prioritization and filters associated with identifying high crash locations. In the studies, crashes had already been assigned to a location. Little is stated about the methodology of how the crashes were assigned to locations. There is also a lack of discussion of how to prepare and format data so that different types of locations can be identified. For instance, can the analysis identify intersections and segments, or is the analysis constrained to identifying only one type of location.

Linking crash and roadway data is essential to the identification of high crash locations. As crashes are irregularly distributed rare events, location may greatly affect the results of crash analysis. Aggregation of crashes is also affected by the quality and format of data.

Quality data are essential to better identify high crash locations. To reduce errors and improve the quality of crash data, SAFETEA-LU provides safety data improvement grants. To receive a grant, a state must quantify data quality and set goals for improving this quality. The safety data improvement grant guidelines are listed in SAFTETEA-LU Section 2006 Appendix 3 (NHTSA 2006). The legislation identified six data quality:

- Timeliness
- Consistency
- Completeness
- Accuracy
- Accessibility
- Data integration

As the quality of each category increases, the ability to identify and consequently mitigate crash problems should increase.

The format of the data is as important as the quality. The process of identifying high crash locations is also dependent on the format of the data. The level of aggregation of data is constrained by the format of available data. For example, an agency will not be able to identify intersections with a high number of crashes if crashes can not be reliably associated with intersections.

Assigning crashes to segments and intersections is the foundation for identifying high crash locations. High crash locations can span from being a spot location to an entire corridor. Having data capable of performing a range of analyses is limited by the preprocessing procedures. Preprocessing procedures include segmenting the road network and assigning crashes to intersections. Determining segmentation and intersection crash assignment methods have an impact on the identification of high crash locations.

In this thesis three analyses that encompass distinct aspects of using data to identify high crash locations were performed. To demonstrate the impact of crash assignment on safety studies, this thesis explores the sensitivity of standard crash rating schemes to assignment methodologies or processes. Three such processes are tested:

Intersection Crash Assignment

- What are the effects of assigning crashes to an intersection based on spatial proximity of the crash to the intersection?
- Test the effect of using crash attribute data along with spatial proximities in assigning crashes to intersections.
- Should the spatial proximity used in assigning crashes to an intersection be based on traffic volume, approach speed, intersection geometry, or other factors?

Road Segmentation

- What effects does different segmentation strategies have on the identification of high crash locations?

- How does segment length affect the ranking of segments for identifying high crash locations?

Crash Rate versus Vehicle Rate

- Test the effect of using crashes versus number of vehicles involved in crashes on safety studies.

CHAPTER 2 LITERATURE REVIEW

Signed into law on August 10, 2005, SAFETEA-LU Section 1401 describes what programs state DOT's must have in place to receive federal funding for safety improvement programs. DOTs must systematically identify hazardous sites and prioritize them in an order to identify locations with the greatest potential for reducing the number and severity of crashes. Each state is to have a Hazard Elimination Program (HEP). To ensure that these programs are carried out in an organized, systematic manner where the greatest benefits can be achieved, a formalized Highway Safety Improvement Program (HSIP) has been established.

In 1979, the HSIP was formally defined in Federal-Aid Highway Program Manual, Volume 8, Chapter 2, Section 3 (FHPM-8-2-3). The primary purpose of FHPM 8-2-3 was to establish the policy for the development and implementation of a comprehensive highway safety program to reduce traffic fatalities and serious injuries on public roadways in each state. The FHPM 8-2-3 includes guidelines and responsibilities for states to follow for their highway safety programs. SAFETEA-LU also mandates the development of a state Strategic Highway Safety Plans (SHSP). States are required to identify their top 5 percent hazardous locations.

As directed by in SAFETEA-LU, states are to develop and implement, on a continuing basis, a highway safety improvement program with an overall objective of reducing the number and severity of crashes and decreasing the potential for crashes on all highways. The components of HSIP are planning, implementation and evaluation. Although SAFETEA-LU does not provide any guidelines or recommendations on how any of the three components are performed, Federal Highway Administration (FHWA) has produced a series of advisory guidance for states to follow (FHWA 2006b). This leads states to the tasks of deciding how to use these guidance with their current system. Many states will have to revamp procedures and systems to address this new mandate.

Planning is a process of identifying locations and ranking those locations with the highest benefit potential. To identify and prioritize locations, a well-maintained data network consisting of crash, traffic and roadway data is needed. States must also develop criteria for identifying and prioritizing locations, and availability of data constrains the selection of criteria. Some of the processes used by states to prioritize locations include (Hallmark 2002):

- Crash frequency
- Crash rate
- Frequency-rate method
- Quality control
- Crash severity
- Index
- Combination of the above methods

Minnesota uses a benefit-cost ratio while Maine uses three years of crash data to calculate critical rate factors. Washington uses a cost-effectiveness ratio and Oregon uses a priority index system based on site crash data, frequency and rate of crashes, and measure of crash severity (Agent 2003). Clearly there are a variety of state crash data systems and prioritization mechanisms.

Intersection Crash Assignment

Each year more than 2.8 million intersection crashes occur in the United States, accounting for more than 44 percent of all crashes (FHWA 2004). In Iowa 55 percent of all urban crashes occur at intersections (Iowa DOT 2003). Intersections constitute only a small part of the overall highway system, yet intersection related crashes constitute more than 50 percent of all crashes in urban areas and over 30 percent in rural areas and account for 21 percent of all fatal crashes (NCHRP 2006). It is not unusual for crashes to be concentrated at intersections, because intersections are the point on the roadway system where traffic movements most frequently conflict. Intersections create opportunities for the most severe types of crashes as well.

Assignment of crashes to an intersection is not often discussed and defined. In most studies that used intersection related crashes, the crashes had already been identified as intersection related and the studies gave no definition of how this assignment was determined. Assigning crashes to an intersection can be done by direct assignment (indicated by reporting officer), or “post-processed” using spatial proximity, attributes query, or a combination of the two. Little published literature could be found regarding the process of assigning crashes to intersections based on crash attributes. For example, while Hallmark (2002) and Knox (2005) indicated they assigned crashes to intersections by crash attribute query, they did not expand on how and why those crash attributes were used in assigning crashes to intersections.

A set of criteria for assigning crashes as intersection related was prepared by Bellomo-McGee, Inc. in a memo to FHWA dated March 26, 1998. The criteria were: (1) crashes must occur within 250 feet of the intersection center and (2) they must be (a) vehicle-pedestrian crashes; (b) crashes in which one vehicle involved in the crash is making a left turn, right turn, or U-turn prior to the crash; or (c) multiple-vehicle crashes in which the accident type is either sideswipe, rear end, or broadside/angle (Vogt 1999). There was no rationale provided regarding the choice of those attributes.

Most agencies use a spatial proximity in assigning crashes to an intersection in order to identify high crash locations. Kentucky Transportation Cabinet uses in urban areas 0.02 mile (105.6 feet) radius and in rural areas uses a 0.05 mile (264 feet) radius to assign crashes as being intersection related (Green 2003). Mankato, MN assigned crashes to intersections by spatial proximity in a buffer from the intersection a distance of 500 feet for speed limits 50 mph or greater and 250 feet for speed limits less than 50 mph (Mankato 2003). The buffer distance was chosen because the distances are usually the length of the turn lanes. The shapes of the buffers were adjusted to avoid overlapping and double counting of closely spaced intersections. Florida DOT uses a buffer distance of 100 feet to assign crashes to intersections (Thobias 2006). Iowa DOT uses a spatial proximity of 75 feet in urban locations and 150 feet in rural locations to assign crashes to intersections (Iowa Dot 2006).

Segmentation

More literature is available on assignment of crashes to road segments. Typically, segments are defined in two fundamental ways with respect to composition and length.

Usually, segments are defined to be homogenous with respect to road geometry, traffic characteristics, safety, and other roadway characteristics. This results in variable lengths. Defining segments by fixed length may allow variation of roadway characteristics and other features within the segment or arbitrary breaking of homogenous segments into small sections. A variety of approaches have been implemented to define roadway segments for identifying high crash locations. Use of many criteria for defining segments, allows testing of specific attributes as predictors of safety performance. However, as segment length decreases, the number of segments containing zero crashes increases. Although there are statistical methods for handling low or zero crash frequencies, shorter segments increase the likelihood that crashes will be geocoded to the wrong segment. Studies suggest that risk conditions can vary rapidly over a fairly short highway length (Markos 2002). A longer segment is more appropriate when conditions are fairly constant over an extended distance, or where cartographic representation on a small scale is desired.

Four types of segmentation may be considered in two groups, predetermined length and sliding scale segmentation. In each group of segmentation there may be either fixed or variable length. Predetermined segmentation results in every portion of the roadway being included in a unique segment. Figure 2.1 provides an example of a predetermined fixed length segmentation in which every segment is the same length. If a roadway's total length is not an increment of the segments length, the remainder length may be proportionally added to each segment. Figure 2.2 provides an example of predetermined variable length segmentation. As is the case in Iowa, the roadway network is segmented into various length segments based on homogenous attributes of the segments. Segments range in length from very short segments to considerable long segments.

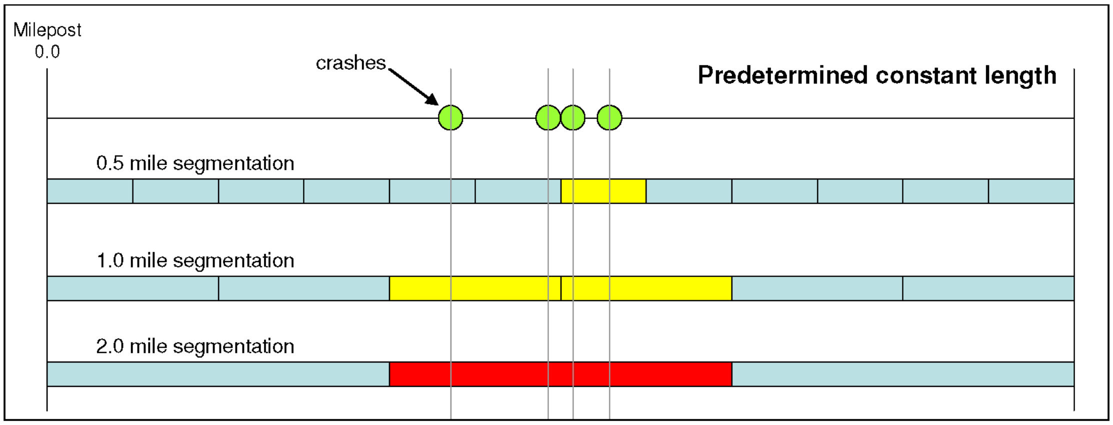


Figure 2.1 Predetermined constant length segmentation

Using predetermined variable length segmentation, short segments may be aggregated. A user can prescribe a minimum segment length. If a segment's length is less than the predetermined length, the next adjoining segment is added to that segment until the new segment's length meets or exceeds the predetermined length. Washington uses 0.1 mile or less segments and New York uses 0.3 mile segments (Geyer 2005).

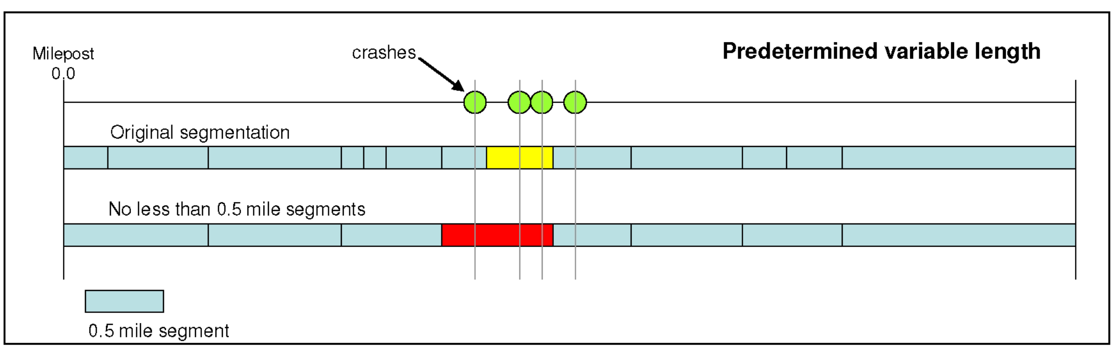


Figure 2.2. Predetermined variable length segmentation

Sliding scale segmentation uses a moving window that slides along the roadway. Again there are two types of sliding scale analysis in terms of the length of the moving window, fixed length and variable length. The portion of the roadway that is inside the moving window is the segment that is analyzed. If the segment meets or exceeds user's definition of a segment, the segment is defined and included in an output file. If not, the moving window is advanced along the roadway an incremental length and the next segment is then analyzed. Utah uses mile long segments but has the ability of using sliding scale window (Geyer 2005). Florida DOT has the capability of performing sliding scale analysis (Thobias 2006). The following are two, more detailed examples of sliding scale segmentation.

Fixed Length Sliding Scale Analysis, State of Practice

California DOT (Caltrans) currently uses a fixed length sliding scale in the analysis of roadway segments with high numbers of crashes. The analysis of a roadway starts at the beginning of the roadway at milepost 0. The first 0.2 mile segment of the roadway is then analyzed. If the segment exceeds a predefined number of crashes, the segment is added to an output table. If not the 0.2 mile segment advances along the roadway by an increment of 0.02 mile and this portion of the roadway is analyzed. The segment keeps sliding along the roadway until a segment is found to be significantly unsafe. When a segment exceeds a predefined number of crashes it is added to the output table. The next segment to be analyzed is started at the end of the segment that was added to the output table (Geyer 2005).

Shown in figure 2.3 is an illustrated example of the Caltrans system of sliding scale analysis. The first segment is analyzed from milepost 0.0 to milepost 0.2. The segment is not found to have a significant number of crashes. The segment then advances 0.02 mile and the segment from mileposts 0.02 mile to 0.22 mile is examined. This segment also does not have a significant number of crashes. The process is repeated twice until a segment is found to have a significant number of crashes. The identified segment is then added to an output table. The next segment to be analyzed is started at the end of the identified segment. A problem identified by Caltrans is that segments containing the highest number of crashes may not be identified. This is also shown in figure 2.1 as the red unidentified segment.

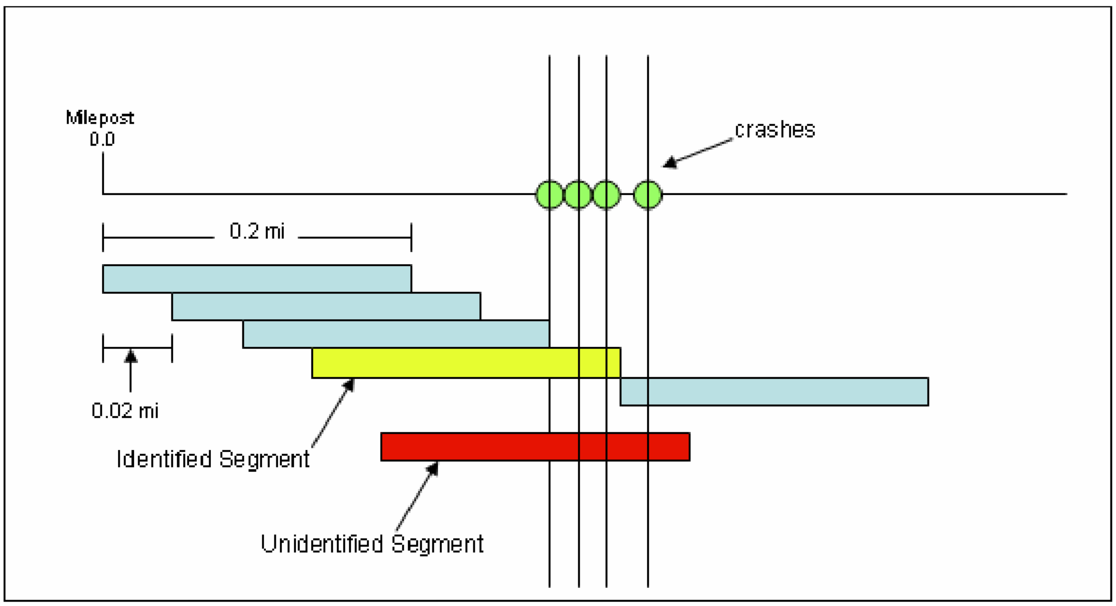


Figure 2.3 Caltrans sliding scale analysis

Wisconsin DOT (WisDOT) also identified the problem that sequential segmentation would have the bias of not identifying high crash concentrations at either side of a common border of two sequential segments. To reduce this potential, WisDOT developed a floating highway segment algorithm, PRÈCIS (Drakopoulos 2005). The process begins by segmenting the highway into 1/100th-mile segments which include attribute data for each segment. The segment's attributes include number of crashes, ADT, number of lanes, and rural or urban location. The user then defines the length of the floating highway segment. Once the segment length is determined, the program advances from the beginning point of a highway until it reaches the first 1/100th-mile segment that contains at least one crash. From this segment, the first section of road to be analyzed is defined by moving half the user defined segment length down stream and half the user defined segment length upstream. After this section is analyzed, the next section to be analyzed is found by moving downstream from the last 1/100th mile segment that contained crashes until the next 100th mile segment that contains crashes. The process is repeated until all roads have been analyzed. For each section analyzed, a crash rate is calculated and compared to the statewide average. If the section's crash rate is above the state's average crash rate, the section is considered as a site eligible for safety treatment.

Kentucky uses a visual basic program that allows the user to select segment length and user defined minimum number of crashes per segment. The users defined segment length is divided into two definitions: spot analysis or section analysis. Spot analyses are 0.1, 0.2, and 0.3 mile segments and section analyses are 1mile and 5 mile segments. The program starts at the beginning of the route and advances along the route until the location of the first crash. From that location, the segment to be analyzed is the user defined length. For example if the user defines the length to be 0.3 mile and the first crash is located at milepoint 10.2, the segment to be analyzed is from milepoint 10.2 to 10.5. If the segment's number of crashes meets or exceeds the user defined number of crashes, then the segment is exported into an output table. The program then advances from the first crash identified to the next crash along the route. Allowing the program to start the next segment analysis from the next crash location will ensure that the segments with the highest number of crashes will be identified (Agent 2003).

Variable Length Sliding Scale Analysis, State of Practice

Highway Safety Information System (HSIS) has developed a variable sliding scale analysis tool for identification of high crash roadway segments. This analysis tool is similar to the Caltrans fixed length sliding scale analysis tool but instead of the sliding segment having a fixed length; the HSIS sliding scale has a variable length. The HSIS variable sliding scale analysis tool allows the user to define the segment length and the incremental length. The first segment of the roadway is analyzed. If the segment does not meet the user defined crash rate threshold, the segment is advanced an incremental distance along the roadway. If the segment meets or exceeds the user defined crash rate threshold, then the segment is increased in length by user defined incremental length. The new segment's crash rate is calculated and compared to the user defined crash rate. If the new segment's crash rate falls below the user defined crash rate, then the incremental length is removed and the previously analyzed segment is exported to an output file. If the new segment's crash rate meets or exceeds user defined crash rate then the segment is increased by the incremental length. The process is repeated until the segment reaches defined maximum milepost or when the user defined maximum number of extensions is reached. When either of these two criteria is met, the segment is defined and exported into an output file.

An example is shown in figure 2.4. For this example the user defined segment length is 0.5 mile and the incremental length is 0.2 mile. First the segment from 0.0 to 0.5 mile is analyzed. The crash rate on this section is less than the user defined crash rate so the section to be analyzed is advanced along the roadway the incremental distance. This section's crash rate is also less than the user defined crash rate, so the section is again advanced along the roadway. The third section that is analyzed has a crash rate equal or greater than the user defined segment crash rate. The segment is then increased in length by the user defined incremental length. This new segment's crash rate is then calculated and its crash rate is greater than the user defined crash rate. Again the segment is increased in length by the incremental length. This process is continued until the user defined maximum number of extensions is applied, user defined maximum length is reached, or the segments crash rate fails below the user defined crash rate. At that point the segment is then defined and exported into a database of segments with crash rates equal or greater than the user defined crash rate. The next section to be analyzed is from the end of the defined segment (GIS/Trans, LTD 2000).

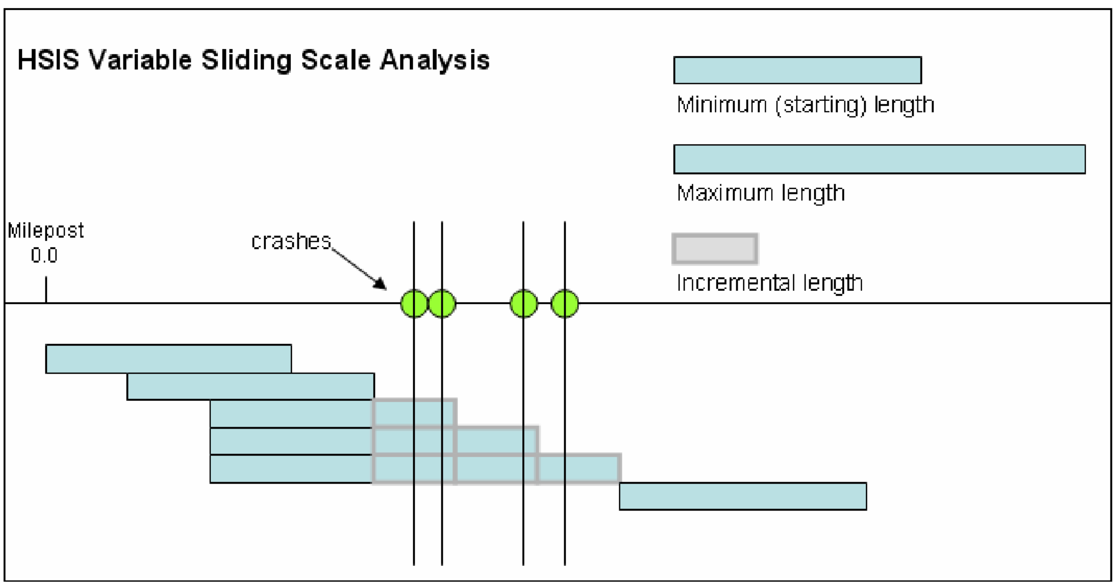


Figure 2.4 HSIS variable sliding scale analysis

In talking to traffic safety officials and reading case studies, there is a deficiency in defining segmentation and assigning crashes to intersections. No studies quantify the effect of segmentation on the identification of high crash locations. Also with assigning crashes to intersections, there are no studies stating the effect of assigning crashes to intersections using different spatial proximities. No literature was found on assigning crashes to intersections based on crash attributes. Stated in several reports is the need for segmenting roads in a fashion that segments can overlap one another in order to be able to identify segments with the highest number of crashes.

CHAPTER 3 METHODOLOGY

To test the effects of assigning crashes to an intersection based on spatial proximity of the crash to the intersection, five distances were chosen to use as buffers when assigning crashes. Those five distances were 75, 100, 150, 200, and 250 feet. A buffer distance of 75 feet produced the same results as the 100 foot buffer and 250 foot buffer produced the same results as 200 foot buffer, so for simplicity the 75 and 250 foot buffers were not listed in this thesis. Also the effect of using crash data attributes along with the three spatial proximities was tested. The effects of segmentation were limited to testing three predetermined fixed length segments because there was limited access to processing tools that are able to work with the network and crash data in its current format. The format of the data could have been changed in order to use these tools but would have been too extensive for this thesis. Also the effect of using the number of vehicles involved in crashes as a vehicle rate instead of the number of crashes as a crash rate was tested. Testing the effects of intersection crash assignment, segmentation, and crash rate versus vehicle rate was done by using the Iowa crash database in Geographical Information System (GIS) format.

GIS Based Approach

Iowa maintains all reported crashes from 1991 to present geo-coded in a GIS database. These geo-coded crash records include attribute data from the original crash report form. Although reporting thresholds and forms have changed throughout the years, these changes have not lead to significant differences in the process of identifying high crash locations.

A problem associated with using a GIS based system is as cartography of the highway system improves the topology is not always clean enough to automatically produce results. As in the case of Iowa, crashes from 1991 to 2000 were geo-coded based on a link node system and since 2001, crashes were digitized on the existing cartography. The procedure of geo-coding crashes on a link node system is accurate on tangent sections, but not on curve sections.

When cartography previously used to geo-code crashes is improved, crash locations should be updated to new coordinates. This has not been done in Iowa and crashes remain at their original location. Figure 3.1 shows a location where the cartography changed and the new intersection location is fifty meters from the previous intersection location.

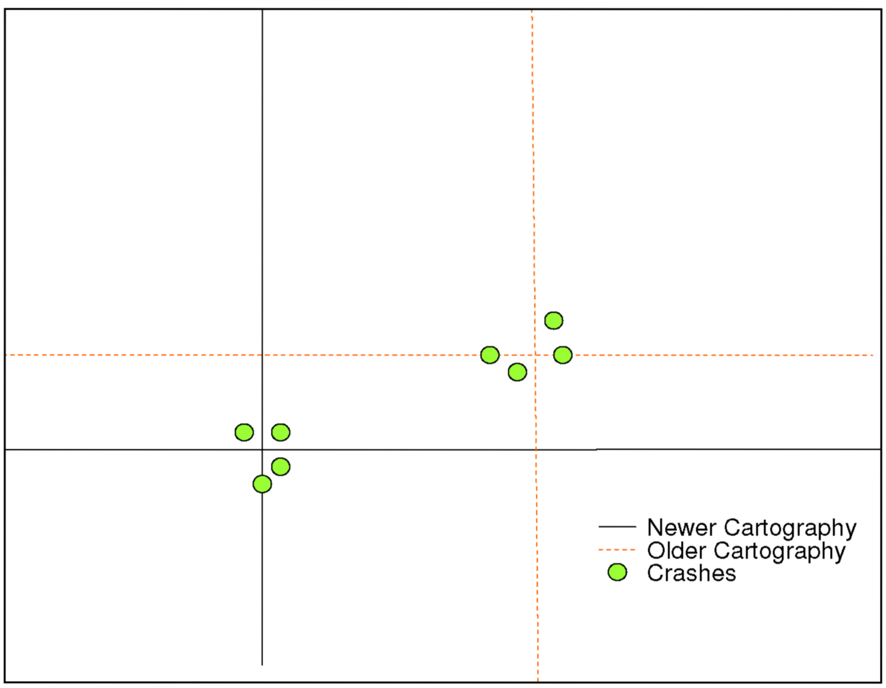


Figure 3.1. Cartography changes

Cartography problems are compounded for intersection locations (as opposed to segments). This problem is worsened if crash data do not contain information on whether the crash was intersection related or to what segment it is associated with. Also, as the cartography is updated and shifted it becomes more difficult to discern, by location, if the crash was intersection related or the crash happened on the main line. Fewer crash location discrepancies are expected now that Iowa DOT has improved all of its cartography to LRS standards using USGS orthophotography. However until all previously located crashes can be reassigned, the problem will remain.

The Iowa crash database was used in all analyses performed in this study. To insure quality analysis not all nuances of the data will be discussed but appropriate measures were taken so that results were not bias. The analyses performed were intersection crash assignment, segmentation, and exposure rate measure. In the next sections, the methodology of each process is explained.

Intersection Crash Assignment

The process of assigning crashes to an intersection by spatial proximity can be problematic and was not well defined in the literature. An analysis was conducted to test the effects of using three different spatial proximities for assigning crashes to intersections. The northwest portion of Iowa was chosen as the study area for intersection crash assignment. The area evaluated was north of Interstate 80 and west of Interstate 35. For this particular task, rural intersections on the primary system were selected. This was done by selecting any intersection within fifty meters of the primary highway system and at least fifty meters from an incorporated city boundary. A distance of fifty meters was chosen because the intersections are not part of the roadway network shapefile but are a different shapefile and the intersections locations are not updated with the cartography. Of the rural primary intersections, 2,750 were identified in the study area as shown in figure 3.2. The selected intersections were examined to assure the location of the intersection node was located on the newest cartography.

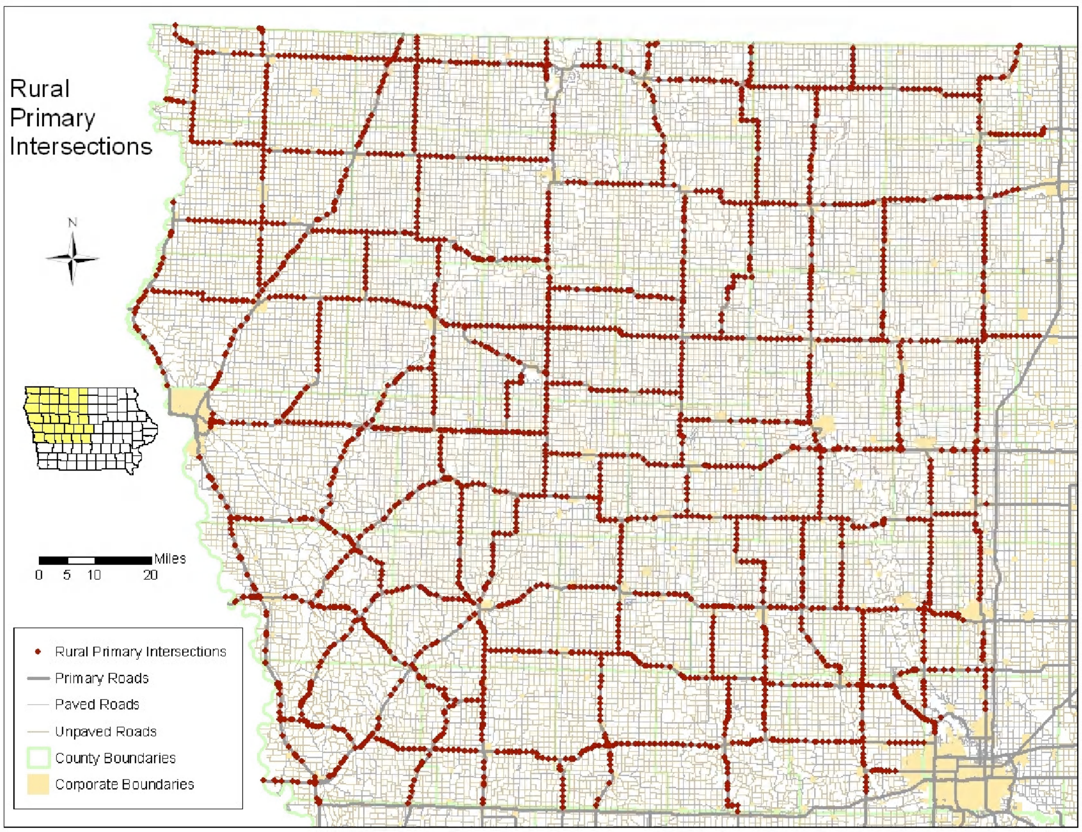


Figure 3.2. Northwest Iowa rural primary intersections

After the intersections were identified, the crashes were then assigned to them using three different spatial proximities. As the Iowa DOT uses a 150 foot spatial proximity to assign crashes to rural intersections, 100 feet, 150 feet, and 200 feet were investigated in a sensitivity analysis. The rank of intersections using crashes assigned at 150 feet was used as a baseline of comparison for corresponding intersection ranks using spatial proximities of 100 and 200 feet. To determine the rank of intersections, the Iowa DOT prioritization procedure was used.

A three year period (2003 to 2005) was used in this study. During this time, 177,125 crashes were reported statewide. At the 2,750 study area intersections, 1,725 crashes were assigned using 200 feet, 1,602 crashes were assigned using 150 feet, and 1,456 crashes were assigned at 100 feet.

After crashes were assigned to intersections for the three spatial proximities, intersections for each were then ranked according to the Iowa DOT's prioritization procedure, the first step of which was to rank in descending order of crash frequency.

The crash rate was then computed by dividing the number of crashes occurring during the analysis period by the number of vehicles entering the intersection. Crash rate was then multiplied by 1,000,000 to produce a crash rate per million entering vehicles (MEV), see equation 3.1. Once the crash rate for each intersection was calculated, the rates were then ranked in descending order.

$$\text{Crash Rate} = \frac{\sum \text{Crashes} * 1{,}000{,}000}{\text{DEV} * \#\,\text{Years} * 365}$$

Equation 3.1. Intersection crash rate

To obtain the number of vehicles that entered the intersection during the analysis period, the daily entering volume (DEV) of the intersection was calculated. The DEV for each intersection was computed using an ArcView 3.3 script using the data associated with the road network GIS files (Hallmark 2002). The script sums the Average Annual Daily Traffic (AADT) for each leg of the intersection and divides the sum in half to compute the DEV as shown in equation 3.2.

$$DEV = \frac{\sum \text{Approach AADT}}{2}$$ Equation 3.2. Daily entering vehicle

A value loss was then computed for each intersection based on the severity of injuries of crash victims. The value loss of crashes at an intersection was a composite score of values assigned to each type of injury severity as seen in table 3.1. While computing the value loss for each intersection, the first fatality at an intersection was assigned the value of a major injury.

Table 3.1. Value of injury severity in value loss ranking

Injury Severity	Value
Fatality	200
Major Injury	100
Minor Injury	10
Possible Injury	1
Unknown Injury	1

As with the frequency and crash rate ranking steps, the value loss was used to rank each intersection in descending order. After frequency, rate, and value loss rankings were compiled, a composite score for each intersection was computed as a weighted combination of the three rankings. The individual ranks are weighted to compute the composite score as follows: 20 percent frequency, 20 percent crash rate, and 60 percent value loss (Iowa DOT 2006). Once the composite scores were calculated, the scores for each intersection were ranked in ascending order to get the final ranking of each intersection.

Segmentation

To test the effects of segmentation length a sensitivity analysis was performed, again using three different values. Three lengths (2-mile, 1-mile, and one-half mile) of segments were defined and ranked according to the Iowa prioritization procedure. The rank of each segment was compared to its corresponding segments rank.

The rural primary system in the northwest portion of Iowa used for intersection assignment sensitivity analysis was also used for segmentation. This portion of the Iowa system was first segmented into two-mile segments using dynamic segmentation in ArcGIS 9.1. Within the two-mile segments, concurrent one-mile and half-mile segments were also defined. The segmented rural primary system is shown in figure 3.3. Of the 1,535 two-mile segments defined, there are 3,016 concurrent one-mile segments and 5,870 concurrent one-half mile segments, as some of the segments did not have equal number of concurrent segments due to network topology. Segments for example were terminated at corporate boundaries and other jurisdictional boundaries.

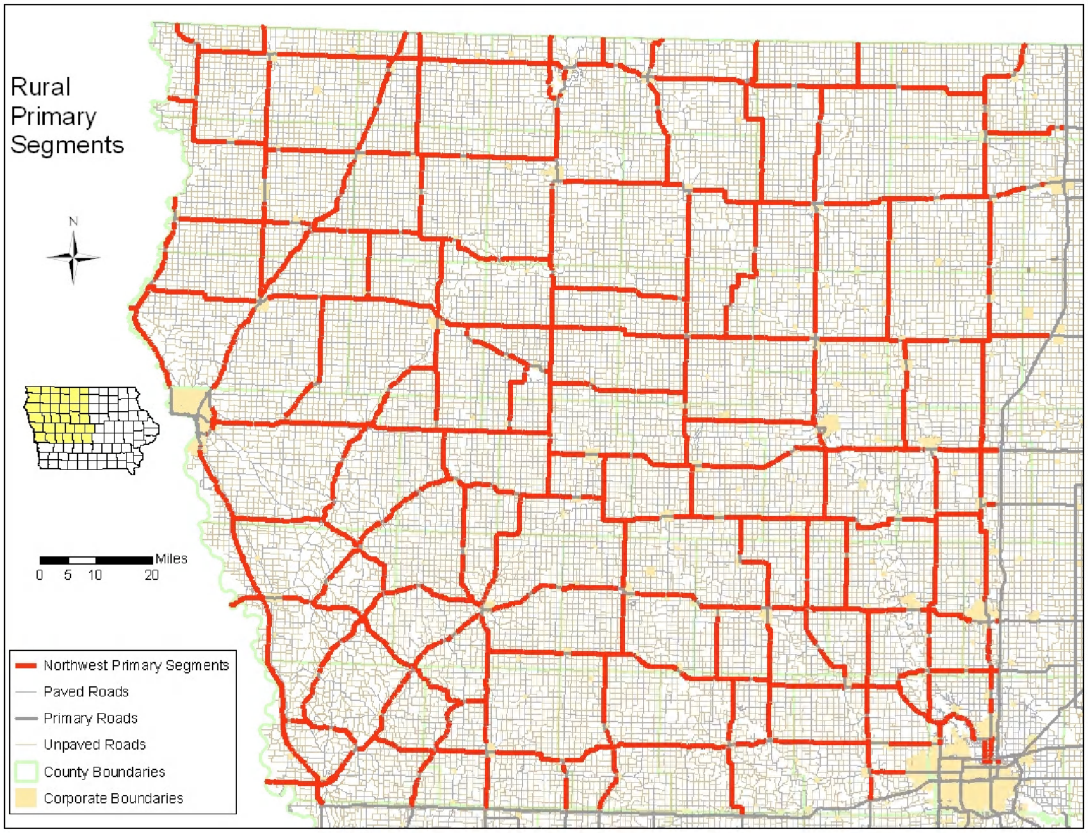


Figure 3.3. Northwest Iowa rural primary segments

As segments were identified, crashes were then assigned. Crashes were assigned by selecting crashes located within 50 meters of any segment outside corporate limits. Once the crashes were selected, the crashes were assigned to the nearest two-mile segment by a spatial join. The process was repeated for one-mile and half-mile segments using the same crash selection criteria used in assigning crashes to their longer counterparts. Once the crashes were assigned to segments, the segments were ranked in each analysis using the Iowa DOT prioritizing procedure as applied in the intersection crash assignment sensitivity study.

The process used for ranking segments was the same as intersections except for the crash rate calculation. Instead of calculating a crash rate based on MEV, the rate calculation used to rank segments was based on vehicle miles traveled (VMT). To calculate VMT, the length of the segment was multiplied by the number of vehicles that traveled the segment. Shown below in equation 3.3 was the formula used to calculate VMT and equation 3.4 was the equation used to calculate crash rate for segments. In the crash rate equation there is a multiplier of 1,000,000 to obtain a crash rate based on one million vehicle miles traveled (MVMT).

$$\text{VMT} = \text{AADT} * \text{Segment Length}$$ Equation 3.3. Vehicle miles traveled

$$\text{Crash Rate} = \frac{\sum \text{Crashes} * 1{,}000{,}000}{\text{VMT} * \#\text{Years} * 365}$$ Equation 3.4. Segment crash rate

For each length, the segments are assigned a final ranking from the composite score.

Crash Rate versus Vehicle Rate

To test the effects of the selection of exposure measure for drivers, the current procedure of calculating crash rate used in the Iowa prioritization procedure was examined. The crash rate exposure measure as shown in equation 3.1 was compared to an alternative exposure measure of vehicle rate as shown in equation 3.5. The new exposure measure (vehicle rate) used the number of vehicles involved in crashes (as noted by $\sum$Vehicles Involved) as opposed to the summation of crashes. The two exposure measures were used to compare final rankings of intersections, with crashes assigned at a spatial proximity of 150 feet. The Iowa prioritization procedure was then used to produce final rankings with the use of the new exposure measure to compute rate. A separate analysis of high crash locations using only these two rates was performed as the Iowa prioritization procedure composite score only uses 20 percent of the rate rank. At intersections, in particular, the use of vehicle rate is thought to be a better measure of exposure due to the number of conflicting traffic movements. Use of vehicle rate will also highlight locations with a higher proportion of multi-vehicle crashes.

$$\text{Crash Rate} = \frac{\sum \text{Crashes} * 1{,}000{,}000}{\text{DEV} * \#\,\text{Years} * 365}$$ Equation 3.1. Intersection crash rate

$$\text{Vehicle Rate} = \frac{\sum \text{Vehicles Involved} * 1{,}000{,}000}{\text{DEV} * \#\,\text{Years} * 365}$$ Equation 3.5. Intersection vehicle rate

CHAPTER 4 ANALYSIS

For both intersection crash assignment and segmentation, a sensitivity analysis was performed to test the effects of two data aggregation methods (DAM) on identifying high crash locations, cross section analysis and before and after studies. For intersection assignment, the ranks of intersections were compared using three spatial aggregations, 100, 150, and 200 feet. For segmentation, three spatial aggregations were investigated, 2-mile, 1-mile, and one-half mile. Descriptive statistics were calculated to describe the differences in rankings for both intersection crash assignment and segmentation. These include: magnitude of ranking shifts, absolute value of ranking position change, lowest ranking, maximum shift in rank and quartile rankings. Finally, the effect of crash exposure was investigated. Specifically, the use of two exposure metrics, crashes or vehicles involved, was examined for its effect on crash analyses.

Intersection Assignment

Three Spatial Proximities

A sensitivity analysis was performed on three spatial proximities for assigning crashes to intersections. The selection of the three distances was based on 150 feet that Iowa uses for assigning crashes in rural areas. Both smaller and larger distances were tested (100 feet and 200 feet) with 150 feet as the baseline. Descriptive statistics were used to quantify shifting in ranking of intersections.

The ranks of intersections using the three buffer distances are shown below in figure 4.1. The baseline rank corresponds to the 150 foot buffer. Only three intersections dropped in ranking below 60 when the buffer distance was reduced to 100 feet. For these, composite ranks are listed next to the symbol. Visual inspection of the graph reveals that most of the intersection ranks change relatively little.

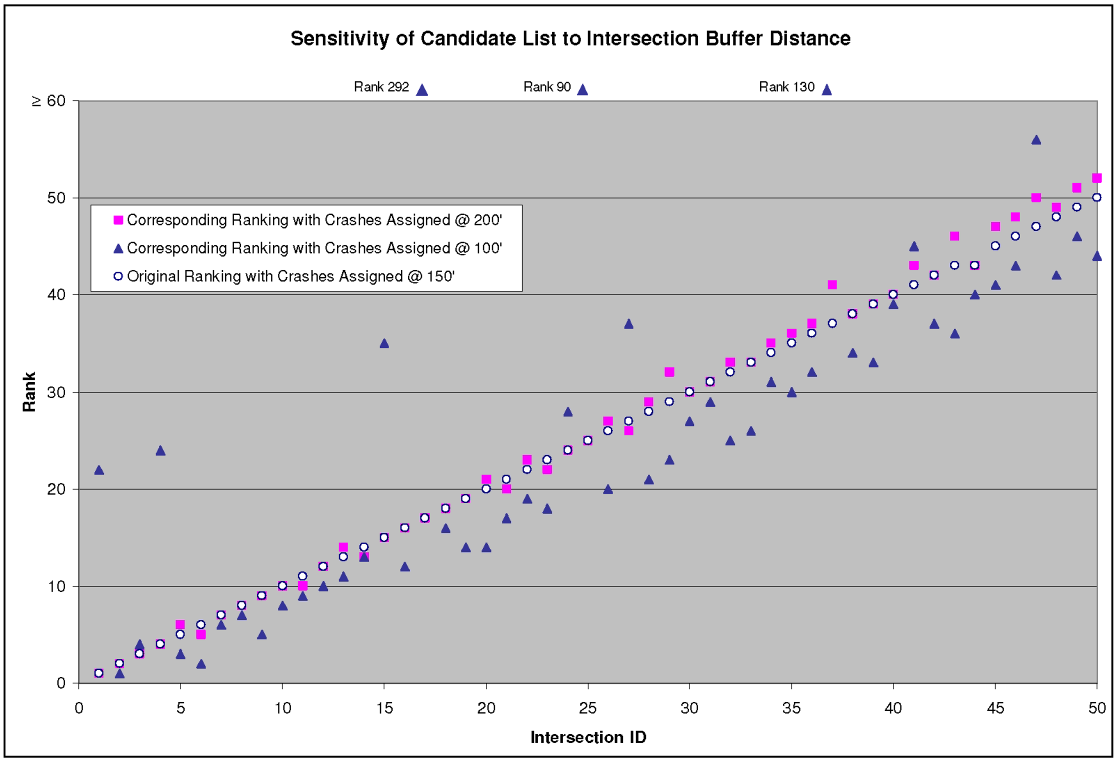


Figure 4.1. Sensitivity of candidate list to intersection buffer distance

Magnitude of Ranking Shifts

Table 4.1 shows the number and percentage of locations that shifted out of the top 50, 100, and 200 locations when 100 foot or 200 foot buffers were used to assign crashes. The portion of top ranked sites shifting out of the top 50, 100, and 200 lists range from four to ten percent.

Table 4.1. Spatial proximity top locations that shifted rank

Spatial Proximity	Top 50 Locations		Top 100 Locations		Top 200 Locations	
	Shift out	Percentage Shift	Shift out	Percentage Shift	Shift out	Percentage Shift
100 feet	4	8%	8	8%	20	10%
200 feet	2	4%	6	6%	18	9%

Absolute Value of Ranking Position Change

Absolute value of ranking position change was used to quantify the shift in rankings of the top 50, 100, and 200 locations again using the 150 foot buffer as a baseline. The number of locations in each absolute value rank shift category is shown in table 4.2. The categories are defined as absolute value of shift in rank as 0, (1-25), (26-100), (101-200), and (>200).

For 200 feet spatial proximity, almost all the shifts were 25 or less in absolute value in rank shift. At 200 foot proximity 100 percent of the top 50 locations, 97 percent of the top 100 locations, and 96 percent of the top 200 locations had an absolute value shift in rank of 25 or less. Also for the 100 foot spatial proximity almost all of the shifts were shifts of 1 to 25. At the top 50 locations 94 percent , at top 100 locations 92 percent, and at top 200 locations 87 percent of locations had an absolute value shift in rank of 1 to 25. No locations using 200 foot proximity had a shift in rank of greater than 100. Only the 100 foot proximity had absolute value shifts in rank greater than 100.

Table 4.2. Spatial proximity absolute value change in ranking position

Absolute Value Rank Shift	Top 50 Locations		Top 100 Locations		Top 200 Locations	
	100 Feet	200 Feet	100 Feet	200 Feet	100 Feet	200 Feet
0	0 (0%)	24 (48%)	0 (0%)	24 (24%)	0 (0%)	24 (12%)
1 - 25	47 (94%)	26 (52%)	92 (92%)	73 (73%)	173 (87%)	167 (84%)
26 - 100	2 (4%)	0 (0%)	5 (5%)	3 (3%)	10 (5%)	9 (5%)
101 - 200	0 (0%)	0 (0%)	0 (0%)	0 (0%)	6 (3%)	0 (0%)
> 200	1 (2%)	0 (0%)	3 (3%)	0 (0%)	11 (6%)	0 (0%)

Maximum Rankings

Lowest rank and maximum shift in rank were also calculated as measures of shift for the top 50, 100, and 200 locations. Table 4.3 lists the lowest rank for 100 foot and 200 foot spatial proximities. For 100 foot spatial proximity the lowest rank for the top 50 locations is nearly 6 times the original rank of 50. Also for the 100 foot proximity, the lowest rank was over 6.5 times the original value of 100 for the top 100 locations and almost 3.5 times the original value of 200 for the top 200 locations. The lowest ranks of the 200 foot proximity were all near to their original values.

Table 4.3. Spatial proximity lowest rank

Top Locations	100 Feet	200 Feet
50 Locations	292	52
100 Locations	668	106
200 Locations	668	225

The maximum shift in ranking for the 100 foot proximity was 275 for top 50 locations and 608 for the top 100 and 200 locations as seen in table 4.4. The shifts in rank were significant but only a few locations had that magnitude of shift change. The shifts in rank were minimal for 200 foot proximity with the largest shift of 81 for the top 200 locations.

Table 4.4. Spatial proximity maximum shift in ranking

Top Locations	100 Feet	200 Feet
50 Locations	275	4
100 Locations	608	45
200 Locations	608	81

Quartile Rankings

Shifts in quartile rankings were also calculated. The quartiles were calculated for the top 50, 100, and 200 locations for both the 100 foot and 200 foot spatial proximities (see table 4.5). The first and second quartile rank for all categories was equal to the original ranks. The third quartile rank for the top 50 and 100 locations were almost equal to the original quartile rank. The only quartile rank that was slightly different from the original quartile rank was the third quartile of the top 200 locations at a spatial proximity of 200 feet.

Table 4.5. Spatial proximity quartile rankings

Absolute Value Rank Shift	Top 50 Locations		Top 100 Locations		Top 200 Locations	
	100 Feet	200 Feet	100 Feet	200 Feet	100 Feet	200 Feet
1st Quartile	13	13	26	26	51	51
2nd Quartile	26	26	51	51	101	103
3rd Quartile	37	39	75	78	150	162

Spatial Proximity and Attributes

Crash attributes may also be used to help identify intersection related crashes. To test the effect of choice of these attributes, five intersections were selected from the study area to be included in a before and after study of the number of crashes at each intersection. The before period was chosen as 2000 to 2002 and the after period was from 2003 to 2005. Three categories of crash attributes were used for assigning crashes. Category A crashes were defined as all crashes the spatial proximity. Category B crashes were defined as possible intersection related and category C crashes were defined as highly likely intersection related. Listed below are the attributes used to define the categories B and C. A crash is defined as a category B crash if either of the first two criteria is satisfied while both must be met for category C. The categories were used in conjunction with spatial proximities.

2000 Crash data possible intersection related

- Intersection classification ≠ 9
  - Intersection class1 thru 8 are intersecting road classifications
- Roadway character > 11 and roadway character ≠ 99
  - Roadway character ≤ 11 are non-intersection crashes
  - Roadway character 11 thru 88 are intersection/interchange crashes
  - Roadway character = 99 are unknown crashes
- Major cause ≠ 1
  - Major cause of 1 is animal crash

or

- Accident type $\geq$ 10 and accident type $\neq$ 99
    - Accident type < 10 are non-collision crashes
    - Accident type 10 thru 88 are collision with something
    - Accident type = 99 are unknown crashes
- Major cause $\neq$ 1
    - Major cause of 1 is animal crash
- Vehicle action $\leq$ 13 and vehicle action $\neq$ 10
    - Vehicle action = 10 is backing
    - Vehicle action = 14 is properly parked
    - Vehicle action = 15 is improperly parked
    - Vehicle action = 16 is unattended moving vehicle
    - Vehicle action = 77 is unreported
    - Vehicle action = 88 is other
    - Vehicle action = 99 is unknown
- Type of trafficway < 6
    - Type of traffic way = 6 is alley
    - Type of traffic way = 7 is driveway
- Driver/vehicle contributing circumstance $\neq$ 12 and $\neq$ 13
    - Driver/vehicle contributing circumstance = 12 is failure to yield from driveway
    - Driver/vehicle contributing circumstance = 13 is failure to yield from parked position

2001 to 2005 Crash data possible intersection related

- Roadway junction/feature $\geq$ 11 and roadway junction/feature $\leq$ 22
    - Roadway junction/feature < 11 are non-intersection crashes
    - Roadway junction/feature 11 thru 22 are intersection crashes
    - Roadway junction/feature > 22 are not reported and unknown crashes
- Major cause $\neq$ 1
    - Major cause of 1 is animal crash

or

- Manner of collision ≤ 7 and manner of collision $\neq 1$
    - Manner of collision = 1 is non-collision
    - Manner of collision 2 thru 7 are collisions
    - Manner of collision > 7 are unreported and unknown collisions
- Major cause $\neq 1$
    - Major cause of 1 is animal crash
- Vehicle action ≤ 11 and vehicle action $\neq 9$
    - Vehicle action = 9 is backing
    - Vehicle action = 12 is properly parked
    - Vehicle action = 13 is improperly parked / unattended
    - Vehicle action = 14 is other
    - Vehicle action > 14 are unreported and unknown
- Trafficway type $\neq 6$ and trafficway type $\neq 7$
    - Trafficway type = 6 is alley
    - Trafficway type = 7 is driveway
- Driver/vehicle contributing circumstance $\neq 17$ and $\neq 18$
    - Driver/vehicle contributing circumstance = 12 is failure to yield from driveway
    - Driver/vehicle contributing circumstance = 13 is failure to yield from parked position

For each intersection the number of crashes in the before and after period were identified for each category for each spatial proximity as listed in Table 4.6. The difference in the number of crashes from the before to the after period is also listed in table 4.6.

Table 4.6. Summation of crashes using spatial proximity and crash attributes

Spatial Proximity	Intersection #1			Intersection #2			Intersection #3			Intersection #4			Intersection #5			Grand Total		
& Attribute Data	Before	After	Δ	Before	After	Δ	Before	After	Δ	Before	After	Δ	Before	After	Δ	Before	After	Δ
100' A	11	17	-6	8	15	-7	7	13	-6	1	11	-10	19	16	3	46	72	-26
100' B	9	9	0	7	13	-6	6	13	-7	1	11	-10	17	16	1	40	62	-22
100' C	3	5	-2	2	11	-9	3	12	-9	0	9	-9	10	15	-5	18	52	-34
150' A	11	17	-6	8	18	-10	7	13	-6	2	11	-9	19	16	3	47	75	-28
150' B	9	9	0	7	16	-9	6	13	-7	1	11	-10	17	16	1	40	65	-25
150' C	3	5	-2	2	12	-10	3	12	-9	0	9	-9	10	15	-5	18	53	-35
200' A	12	18	-6	17	18	-1	11	13	-2	4	11	-7	18	16	2	62	76	-14
200' B	8	9	-1	14	16	-2	10	13	-3	4	11	-7	17	16	1	53	65	-12
200' C	4	5	-1	9	12	-3	9	12	-3	2	9	-7	13	15	-2	37	53	-16

Segmentation

While different states use various segment lengths for static segmentation, no rationale for this could be identified in the literature. To quantify the affects of using different segment lengths, three analyses were performed using predefined fixed length segments. Segments are then compared to concurrent segments of different lengths. For example, 1-mile segment was compared to its concurrent 2-mile and two one-half mile segments.

Two-Mile Segments

Sensitivity analysis was performed using two-mile segments as a baseline. The analysis compared the ranks of two-mile segments to average, high, and low ranks of their concurrent one-mile and one-half mile segments. Figure 4.2 shows the rank of the two-mile segments with the concurrent one-mile and one-half mile segments. To allow a higher level of detail, the graph scale is limited to segments ranks higher than 100. See appendix for the original graph showing all concurrent segment ranks.

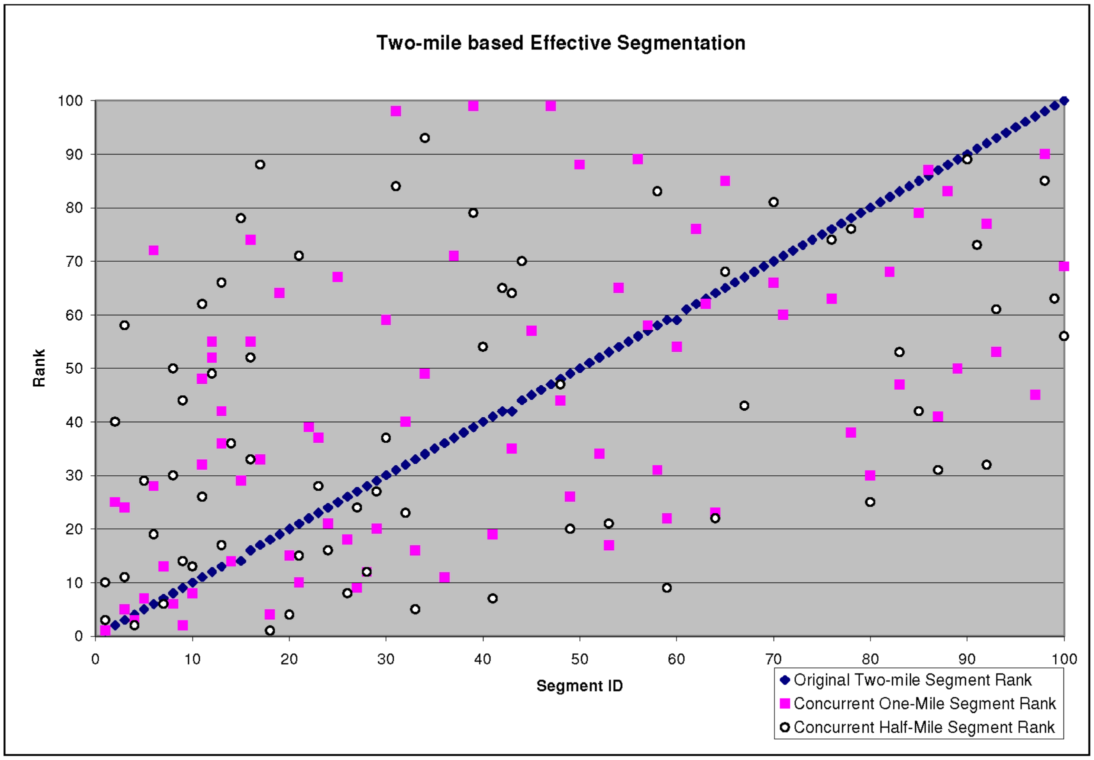


Figure 4.2. Two-mile based effective segmentation

Magnitude of Ranking Shifts

The magnitude of ranking shift was calculated for the number of locations that shifted out of the original 1 to 50, 1 to 100, and 1 to 200 rankings. The original rankings were identified using the ranking of the two-mile segments. The number and percentage of locations that shifted out of the top 50, 100, and 200 locations of the concurrent segments are listed in table 4.7. The concurrent segments include the average, high, and low of one-mile and one-half mile segments.

Of the concurrent one-mile and one-half mile average and low rank segments for all three top locations, nearly all locations shifted out of their respective lists. The concurrent high rank one-mile segment had 34 percent of locations shift out of the top 50 locations to 22 percent of locations that shifted out of the top 200 locations. For the concurrent one-half mile segments, the high ranked segments had the lowest number of locations that shifted out of the top locations with 40 percent of locations shifting out of the top 50 locations to 29 percent of locations shifting out of the top 200 locations.

Table 4.7. Two-mile segment shifts in rank

Concurrent Segment	Top 50 Locations		Top 100 Locations		Top 200 Locations	
	Shift out	Percentage Shift	Shift out	Percentage Shift	Shift out	Percentage Shift
Average 1-mile	46	92%	92	92%	176	88%
High Rank 1-mile	17	34%	27	27%	44	22%
Low Rank 1-mile	47	94%	94	94%	188	94%
Average ½ -mile	50	100%	100	100%	199	99.5%
High Rank ½ -mile	20	40%	41	41%	58	29%
Low Rank ½ -mile	50	100%	100	100%	200	100%

Absolute Value of Ranking

For the two-mile segments, the absolute value of ranking position change was calculated for the top 50, 100, and 200 locations. The top 50, 100, and 200 locations were identified using the original two-mile segment rankings. Listed in table 4.8 is the number of locations in each category of absolute value of shift in rank. The absolute value rank shift categories are groups of locations that experienced 0, (1-25), (26-100), (101-200), and (>200) change in ranking position.

The average and low rank one-mile segments had most locations for all three top location category have shifts in rank greater than 100. The high rank one-mile segments had most locations with an absolute value of shift in rank between 1 and 25. The concurrent average and low rank one-half mile segments had completely all locations having a shift in rank greater than 200. The percentage of locations that experienced a shift in rank greater than 200 ranges from 98 percent for the top 50 locations to 100 percent for the top 50, 100, and 200 locations. The high rank one-half mile segments had most locations either experiencing a shift in rank of 1 to 100.

Table 4.8. Two-mile segments absolute value change in ranks

Concurrent Segment	Absolute Value Rank Shift	Top 50 Locations		Top 100 Locations		Top 200 Locations	
Average 1-mile	0	0	(0%)	0	(0%)	0	(0%)
	1 - 25	1	(2%)	1	(1%)	1	(1%)
	26 - 100	10	(20%)	11	(11%)	13	(7%)
	101 - 200	15	(30%)	23	(23%)	35	(18%)
	> 200	24	(48%)	65	(65%)	151	(76%)
High Rank 1-mile	0	2	(4%)	2	(2%)	3	(2%)
	1 - 25	32	(64%)	50	(50%)	73	(37%)
	26 - 100	14	(28%)	34	(34%)	85	(43%)
	101 - 200	2	(4%)	13	(13%)	28	(14%)
	> 200	0	(0%)	1	(1%)	11	(6%)
Low Rank 1-mile	0	0	(0%)	0	(0%)	0	(0%)
	1 - 25	1	(2%)	1	(1%)	1	(1%)
	26 - 100	7	(14%)	8	(8%)	8	(4%)
	101 - 200	4	(8%)	5	(5%)	11	(6%)
	> 200	38	(76%)	86	(86%)	180	(90%)
Average ½ -mile	0	0	(0%)	0	(0%)	0	(0%)
	1 - 25	0	(0%)	0	(0%)	0	(0%)
	26 - 100	0	(0%)	0	(0%)	0	(0%)
	101 - 200	1	(2%)	1	(1%)	1	(1%)
	> 200	49	(98%)	99	(99%)	199	(100%)
High Rank ½ -mile	0	0	(0%)	0	(0%)	0	(0%)
	1 - 25	28	(56%)	37	(37%)	51	(26%)
	26 - 100	18	(36%)	42	(42%)	90	(45%)
	101 - 200	2	(4%)	14	(14%)	32	(16%)
	> 200	2	(4%)	7	(7%)	27	(14%)
Low Rank ½ -mile	0	0	(0%)	0	(0%)	0	(0%)
	1 - 25	0	(0%)	0	(0%)	0	(0%)
	26 - 100	0	(0%)	0	(0%)	0	(0%)
	101 - 200	0	(0%)	0	(0%)	0	(0%)
	> 200	50	(100%)	100	(100%)	200	(100%)

Maximum Ranking

For the original top 50, 100, and 200 locations, the lowest rank of comparing two-mile segment to its concurrent segments is listed in table 4.9 The lowest rank of the concurrent one-mile segment is slightly over 21 times the original for top 50 locations, top 50 locations is almost 3 times the original and the lowest shift in rank for top 50 locations is 84. For the top 100 locations the lowest rank is almost 2.5 times the original and the lowest shift in rank is 145. The top 200 locations' lowest rank is nearly twice of the original with a maximum shift in rank of 182.

Table 4.9. Two-mile segment lowest rank

Top Locations	Average 1-mile Segment	High Rank 1-mile Segment	Low Rank 1-mile Segment	Average ½ - mile Segment	High Rank ½ -mile Segment	Low Rank ½ -mile Segment
50 Locations	1072.5	163	1982	2096	278	2793
100 Locations	1072.5	371	1982	2116	487	2793
200 Locations	1073.5	821	1982	2120	913	2793

One-Mile Segments

A sensitivity analysis was next performed on one-mile segments. The analysis compared the ranks of one-mile segments to the rank of the concurrent two-mile segment and the average, high, and low ranks of the concurrent one-half mile segments. Figure 4.3 shows the rank of the one-mile segments with the concurrent two-mile segment and one-half mile segments. For illustration purposes that graph does not show any concurrent segments ranked higher than 100. See appendix for the original graph showing all concurrent segment ranks.

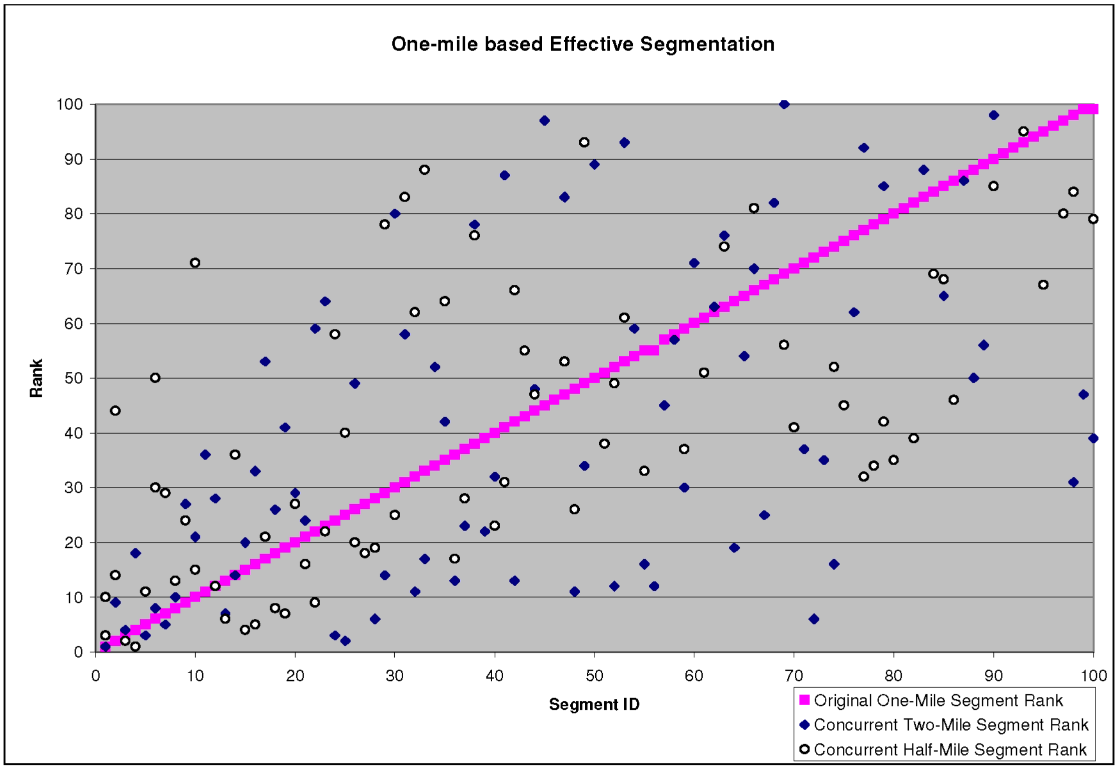


Figure 4.3. One-mile effective segmentation

Magnitude of Ranking Shifts

For the top 50, 100, and 200 locations the magnitude of ranking shift was calculated. One-mile segments were used to determine the original ranking. Table 4.10 lists the number and percentage of locations that shifted out of the top 50, 100, and 200 locations of the concurrent segments. The concurrent segments include the two-mile segment and the average, high, and low of the one-half mile segments. The portion of the concurrent 2-mile segments shifting out of the top locations ranges from 15.5 to 28 percent. The concurrent average and low rank one-half mile segments had a high percentage of locations shifting out of the top ranked sites with a range from 92 to 97 percent. The concurrent high rank one-half mile segments range of shifting locations is from 21 to 34 percent.

Table 4.10. One-mile segment shifts in rank

Concurrent Segment	Top 50 Locations		Top 100 Locations		Top 200 Locations	
	Shift out	Percentage Shift	Shift out	Percentage Shift	Shift out	Percentage Shift
2-mile	14	28%	20	20%	31	15.5%
Average ½ -mile	46	92%	95	95%	188	94%
High Rank ½ -mile	17	34%	30	30%	42	21%
Low Rank ½ -mile	47	94%	96	96%	194	97%

Absolute Value of Ranking

The absolute value of ranking position change was calculated for the top 50, 100, and 200 locations for one-mile segments is shown in table 4.11. The categories of the absolute value rank in shift are based from the absolute value of shift in ranking positions and are the same as the previous sections absolute value ranking change.

The concurrent 2-mile segments had mostly shifts in rank of 1 to 100. The concurrent average and low rank one-half mile segments had completely all locations having a shift in rank greater than 200. The high rank one-half mile segments had most locations either experiencing an absolute value shift in rank from 1 to 100.

Table 4.11. One-mile segments absolute value change in rank

Concurrent Segment	Absolute Value Rank Shift	Top 50 Locations		Top 100 Locations		Top 200 Locations	
2-mile	0	2	(4%)	2	(2%)	3	(2%)
	1 - 25	33	(66%)	53	(53%)	73	(37%)
	26 - 100	15	(30%)	45	(45%)	101	(51%)
	101 - 200	0	(0%)	0	(0%)	23	(12%)
	> 200	0	(0%)	0	(0%)	0	(0%)
Average ½ -mile	0	0	(0%)	0	(0%)	0	(0%)
	1 - 25	1	(2%)	1	(1%)	1	(1%)
	26 - 100	4	(8%)	4	(4%)	4	(2%)
	101 - 200	9	(18%)	11	(11%)	17	(9%)
	> 200	36	(72%)	84	(84%)	178	(89%)
High Rank ½ -mile	0	1	(2%)	1	(1%)	1	(1%)
	1 - 25	35	(70%)	53	(53%)	74	(37%)
	26 - 100	11	(22%)	34	(34%)	95	(48%)
	101 - 200	2	(4%)	9	(9%)	17	(9%)
	> 200	1	(2%)	3	(3%)	13	(7%)
Low Rank ½ -mile	0	0	(0%)	0	(0%)	0	(0%)
	1 - 25	1	(2%)	1	(1%)	1	(1%)
	26 - 100	3	(6%)	3	(3%)	3	(2%)
	101 - 200	2	(4%)	2	(2%)	3	(2%)
	> 200	44	(88%)	94	(94%)	193	(97%)

Maximum Rankings

The lowest rank of the concurrent segments to the original one-mile segments for the top 50, 100, and 200 locations is listed in table 4.12. The lowest rank of the concurrent two-mile segment is slightly over 2 times the original for top 50 locations and almost 1.5 times the original max rank for the top 200 locations. The lowest rank of the concurrent low rank for all top locations is 2,793 which is the lowest rank of all one-half mile segments. The lowest rank for the high rank one-half mile segment ranges from approximately 5.5 times of the original for the top 50 sites to 3.5 times of the original for the top 200 sites. The range of lowest ranks for the average one-half mile is from 1,410 for top 50 sites to 1,520 for the top 200 locations.

Table 4.12. One-mile segment lowest rank

Top Locations	2-mile Segment	Average One-half mile Segment	High Rank One-half mile Segment	Low Rank One-half mile Segment
50 Locations	120	1410	278	2793
100 Locations	175	1439	448	2793
200 Locations	296	1520	720	2793

One-half mile Segments

A sensitivity analysis was performed on one-half mile segments. The analysis compared the ranks of one-half mile segments to the rank of the concurrent two-mile and one-mile segments. Figure 4.4 shows the rank of the one-half mile segments with the concurrent two-mile and one-mile segments. For illustration purposes that graph does not show any concurrent segments ranked higher than 100. See appendix for the original graph showing all concurrent segment ranks.

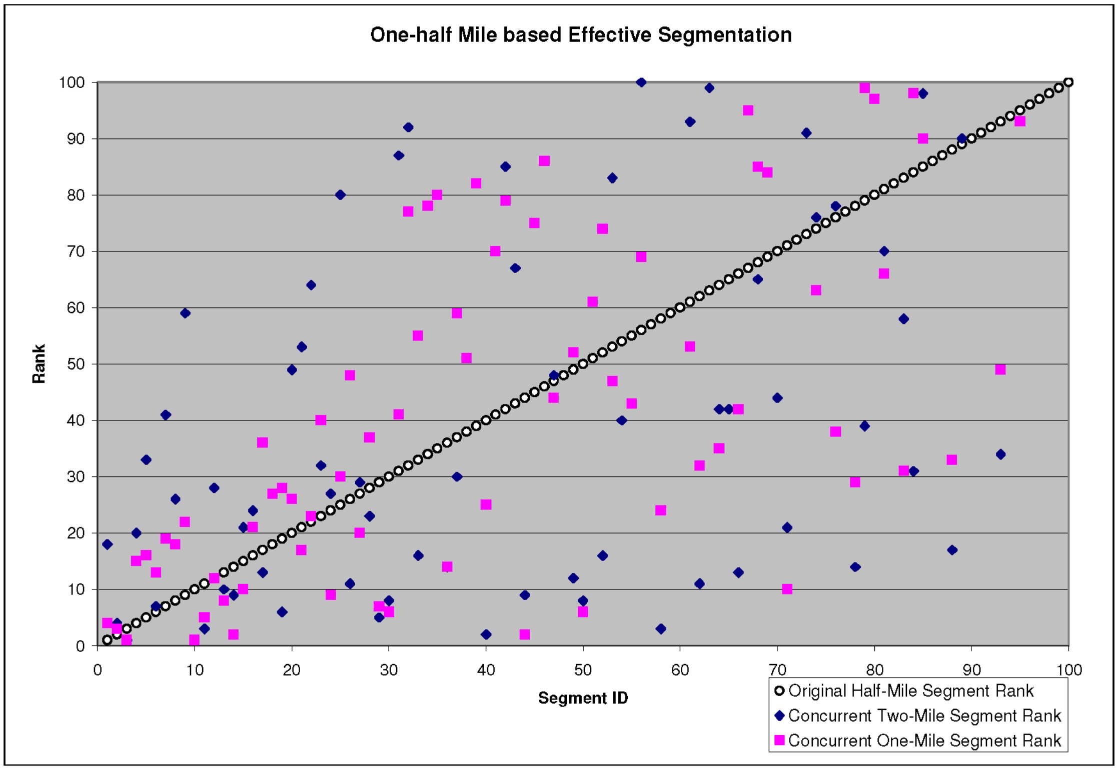


Figure 4.4. One-half mile effective segmentation

Magnitude of Ranking Shifts

The number and percentage of locations shifting out of the top 50, 100, and 200 sites determined by one-half mile segments is shown in table 4.13. The concurrent segments include the one-mile and two-mile segments. The portion of the concurrent 2-mile segments shifting out of the top locations ranges from 18 to 28 percent. The concurrent one-mile segments had a high percentage of locations shifting out of the top ranked sites with a range from 20 to 34 percent.

Table 4.13. One-half mile segments shifts in rank

Concurrent Segment	Top 50 Locations		Top 100 Locations		Top 200 Locations	
	Shift out	Percentage Shift	Shift out	Percentage Shift	Shift out	Percentage Shift
2-mile	14	28%	26	26%	36	18%
1-mile	17	34%	33	33%	40	20%

Absolute Value of Ranking

The absolute value of ranking position change was calculated for the top 50, 100, and 200 locations for 1-half mile segments is shown in table 4.14. The categories of the absolute value rank in shift are based from the absolute value of shift in ranking positions and are the same as the previous sections absolute value ranking change.

The concurrent 2-mile segments had mostly shifts in rank of 1 to 100. The concurrent 1-mile segments most locations either experiencing a shift in rank of 1 to 100.

Table 4.14. One-half mile segments absolute value change in rank

Concurrent Segment	Absolute Value Rank Shift	Top 50 Locations		Top 100 Locations		Top 200 Locations	
2-mile	0	0	(0%)	0	(0%)	0	(0%)
	1 - 25	27	(54%)	38	(38%)	53	(27%)
	26 - 100	19	(38%)	49	(49%)	105	(53%)
	101 - 200	4	(8%)	13	(13%)	40	(20%)
	> 200	0	(0%)	0	(0%)	2	(1%)
1-mile	0	1	(2%)	1	(1%)	1	(1%)
	1 - 25	37	(74%)	54	(54%)	77	(39%)
	26 - 100	12	(24%)	44	(44%)	108	(54%)
	101 - 200	0	(0%)	1	(1%)	13	(7%)
	> 200	0	(0%)	0	(0%)	1	(1%)

Maximum Rankings

Listed in table 4.15 is the lowest rank of the concurrent segments to the original one-half mile segments for the top 50, 100, and 200 locations. The lowest rank of the concurrent two-mile segment is slightly over 3 times the original for top 50 locations and almost 9 times the original max rank for the top 200 locations. The concurrent low rank of the one-mile segment is almost 3 time the original for the top 50 locations and slightly over 7.5 time the original for the top 200 locations.

Table 4.15. One-half mile segment lowest rank

Top Locations	2-mile Segment	1-mile Segment
50 Locations	164	143
100 Locations	284	230
200 Locations	464	381

Crash Rate versus Vehicle Rate

To test the effects of using different exposure rates, descriptive statistics were calculated for the two analyses. In both analyses, crash rate was used as a baseline to compare vehicle rate. The first analysis performed was using vehicle instead of crash rate in the Iowa prioritization procedure. The second analysis performed was using crash rate and vehicle rate as the only identifiers of high crash locations.

Iowa Prioritization Procedure

A sensitivity analysis was performed to identify the change in rank of the intersections using different exposure rates in the Iowa prioritization procedure. As seen in figure 4.5, using vehicle rate instead of crash rate in the Iowa prioritization procedure does not cause large shifts in rank. This was due to the fact that crash rate or vehicle rate only comprises 20 percent of the final composite score which intersections are ranked from.

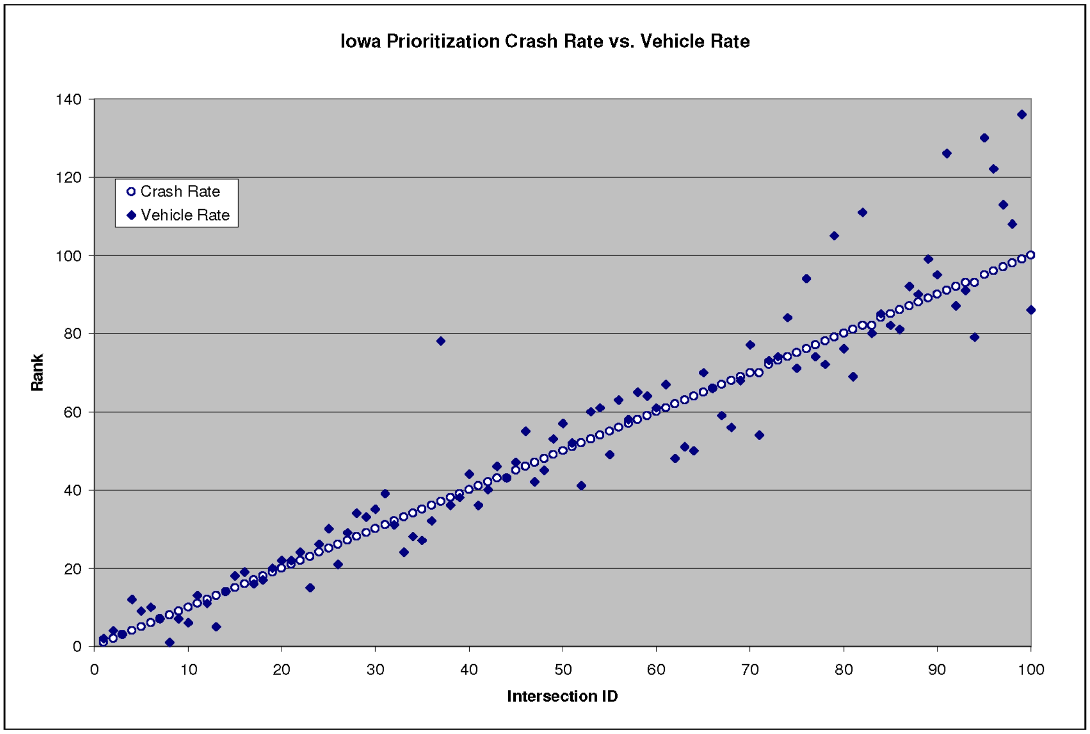


Figure 4.5. Iowa prioritization crash rate vs. vehicle rate

Magnitude of Ranking Shifts

Table 4.16 lists the number and percentage of locations that fell out of the top 50, 100, and 200 locations. The range of shift in ranking for the three top high crash locations is from 8 percent for the top 50 and 100 locations to 8.5 percent for the top 200 locations.

Table 4.16. Iowa prioritization procedure shifts in rank

Top 50 Locations		Top 100 Locations		Top 200 Locations	
Shift out	Percentage Shift	Shift out	Percentage Shift	Shift out	Percentage Shift
4	8%	8	8%	17	8.5%

Absolute Value of Ranking Positions Change

Absolute value of ranking position change was used to quantify the shift in rankings of the top 50, 100, and 200 locations. The final ranking of the Iowa prioritization procedure using crash rate was compared to the final ranking of the Iowa prioritization procedure using vehicle rate. The absolute value of rank shift was calculated by the difference in change of rank. The number of locations in each absolute value rank shift category is shown in table 4.17. The absolute value rank shift categories are groups of locations that experienced 0, (1-25), (26-100), (101-200), and (>200) shifts in ranking position.

For the top 50 locations 8 percent of locations did not shift in rank while 90 percent of locations had a shift of 1 to 25. The top 100 locations had 5 percent of locations that did not shift and 88 percent of locations that had a shift in rank between 1 and 25. At the top 200 locations 4 percent of the locations did not shift and 73 percent of the locations shifted between 1 and 25 ranking positions. Locations that either did not change rank or had an absolute value shift in ranking between 1 to 25 comprised of 98, 93, and 77 percent of the top 50, 100, and 200 locations respectively.

Table 4.17. Iowa prioritization rate absolute value change in ranking position

Absolute Value Rank Shift	Top 50 Locations Composite Score		Top 100 Locations Composite Score		Top 200 Locations Composite Score	
0	4	(8%)	5	(5%)	7	(4%)
1 - 25	45	(90%)	88	(88%)	146	(73%)
26 - 100	1	(2%)	7	(7%)	47	(24%)
101 - 200	0	(0%)	0	(0%)	0	(0%)
> 200	0	(0%)	0	(0%)	0	(0%)

Maximum Rankings

The lowest rank and the maximum shift in rankings were calculated. Again the top high crash locations were identified using crash rate and compared vehicle rate in the prioritization procedure. The lowest rank for the top 50 locations is 78, 136 for the top 100 locations, and 284 for the top 200 locations as seen in table 4.18. The maximum shift in rank is 41 for both the top 50 and100 locations and 90 for the top 200 locations.

Table 4.18. Iowa prioritization rate lowest rank

Top Locations	Composite Score
50 Locations	78
100 Locations	136
200 Locations	284

Quartile Ranking

The first, second, and third quartile rankings using vehicle rate instead of crash rate for the top 50, 100, and 200 locations were calculated as listed in table 4.19. As one may expect the quartiles are almost exactly equal to the original quartile rankings. The maximum difference in rankings is one and only occurred in four categories.

Table 4.19. Iowa prioritization rate quartile rankings

Absolute Value Rank Shift	Top 50 Locations	Top 100 Locations	Top 200 Locations
	Composite Score	Composite Score	Composite Score
1st Quartile	13	26	51
2nd Quartile	25	51	101
3rd Quartile	38	75	151

Rate as an Identifier

The second sensitivity analysis of exposure rate was comparing crash rate to vehicle rate as an identifier of high crash locations. Crashes assigned to an intersection at a spatial proximity of 150 feet were used for this analysis. The analysis compared the shift in ranking of locations comparing vehicle rate to crash rate as a baseline. The same descriptive statistics as in the previous analysis were calculated to quantify the shift of rankings. In this analysis, unlike the previous analysis of exposure rate, there was a wide range of shifts in rank as seen in figure 4.6.

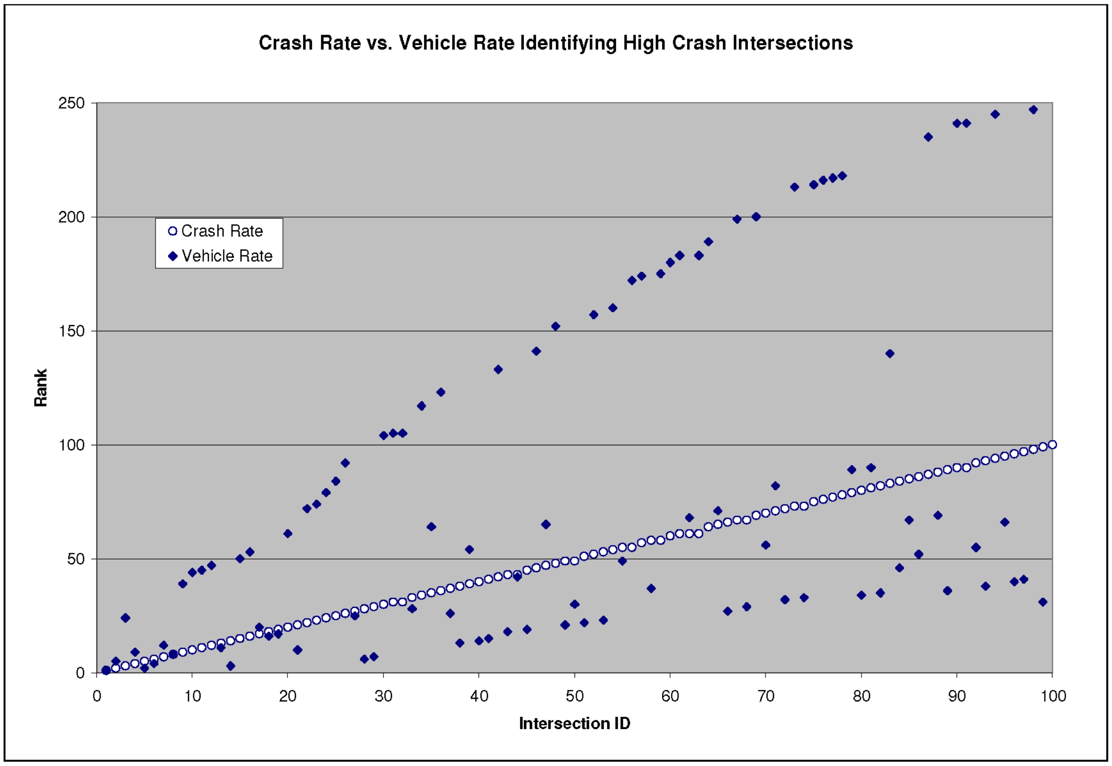


Figure 4.6. Exposure rate identification of high crash intersections

The rank of intersections using vehicle rate in figure 4.6 seem to follow two asymptotes, the top asymptote which was roughly three times the rank of intersection using crash rate and the lower which was roughly one-half the rank of the intersection using crash rate. In investigating the two trends, a graph was produced using crash rate and vehicle rate as before but for the three different segmentation lengths (figure 4.7). A similar trend was identified for all three segments lengths but the top asymptote was roughly twice the rank and the lower portion of the graph had more randomness in the ranks. Although this was thoroughly investigated, no explanation could be identified to explain the two trends.

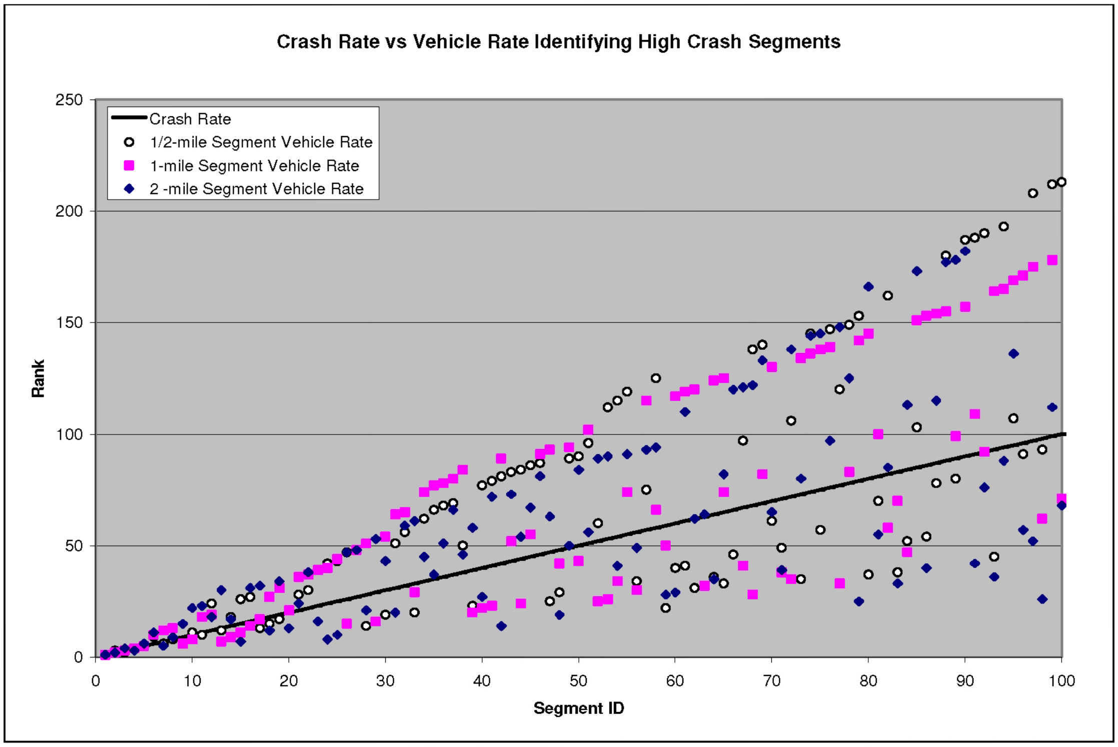


Figure 4.7. Exposure rate identification of high crash segments

Magnitude of ranking shifts

As in the previous analysis the shift in ranking was calculated. The number and percentage of locations that fell out of the top 50, 100, and 200 locations with the new ranking applied by using vehicle rate is listed below in table 4.20. When vehicle rate was used the number of locations that drop out of each category ranges from 36 percent for top 50 locations to 25 percent for the top 100 locations. Nearly one-third of top 50 and 100 locations shifted out of the original listings compared to one-fourth of the top 200 locations.

Table 4.20. Exposure rate shifts in rank

Top 50 Locations		Top 100 Locations		Top 200 Locations	
Shift out	Percentage Shift	Shift out	Percentage Shift	Shift out	Percentage Shift
18	36%	31	31%	50	25%

Absolute Value of Ranking Position Change

The absolute value of rank shift between crash rate and vehicle rate was also calculated for the top 50, 100, and 200 locations. The same categories were used as in the previous absolute value of rank shift analysis and the number of locations in each category is listed in table 4.21.

Locations that did not shift rank ranged from 4 percent for the top 50 locations to 1 percent for the top 200 locations. A shift in rank of 1 to 25 occurred at 46 percent of the top 50 locations, 33 percent of the top 100 locations, and 26 percent of the top 200 locations. The range of locations that experienced a shift in rank between 26 and 100 was from 48 percent of the top 50 locations to 44 percent of the top 200 locations. Only one location had a shift in rank of 101 to 200 for the top 50 locations while the top100 and 200 locations had shifts in rank of 23 percent and 30 percent respectfully. No locations had an absolute value of shift in rank greater than 200.

Table 4.21. Exposure rate absolute value change in ranking position

Absolute Value Rank Shift	Top 50 Locations		Top 100 Locations		Top 200 Locations	
	Rate Only		Rate Only		Rate Only	
0	2	(4%)	2	(2%)	1	(1%)
1 - 25	23	(46%)	33	(33%)	51	(26%)
26 - 100	24	(48%)	42	(42%)	87	(44%)
101 - 200	1	(2%)	23	(23%)	60	(30%)
> 200	0	(0%)	0	(0%)	0	(0%)

Maximum Rankings

The lowest rank of comparing vehicle rate to crash rate in identifying high crash locations is listed in table 4.22 for the top 50, 100, and 200 locations. The lowest rank of the top 50 locations is almost 3 times the original and the maximum shift in rank for top 50 locations is 104. For the top 100 locations the lowest rank is almost 2.5 times the original and the maximum shift in rank is 151. The top 200 locations' lowest rank is nearly twice of the original with a maximum shift in rank of 187.

Table 4.22. Exposure rate lowest rank

Top Locations	Rate Only
50 Locations	152
100 Locations	251
200 Locations	386

Quartile Ranking

Table 4.23 lists the first, second, and third quartile rankings. All top locations first and second quartile ranks were approximately equal to the original ranks. For all locations the third quartile was higher than the original as shown by 84 percent change for top 50 locations, 80 percent change for top 100 locations, and 35 percent change for top 200 locations.

Table 4.23. Exposure rate quartile rankings

Absolute Value Rank Shift	Top 50 Locations	Top 100 Locations	Top 200 Locations
	Rate Only	Rate Only	Rate Only
1st Quartile	13	26	50
2nd Quartile	29	54	102
3rd Quartile	70	135	202

HRRR Case Study

Changing the process of segmentation can greatly change the results of an analysis. An example of this is the high risk rural roads (HRRR) project in Iowa. The HRRR project identified eligible rural paved collectors and local road segments that had an above statewide average crash rate or crash density of fatal and major injury crashes from 2001 to 2005. The statewide average crash rate and crash density that were used was the averages of the rural paved collectors and local roads. The first segmentation used was without county constraint but since the HRRR project is used by county engineers, segmentation with county constraints was added. County constrained segmentation divided segments at county boundaries. Applying county constraints to the segmentation increases the number of eligible segments from 1,673 to 1,706 but reduced the total length of segments from 7,063 center line miles to 6,697 miles.

The effect of both segmentations is shown in figure 4.8. In figure 4.8, the green lines are county boundaries and the gray lines are paved roads. The yellow and red segments are eligible using segmentation without county constraints. Using county constraints segmentation reduces eligible segments to only the yellow segments and the red segments are now ineligible.

Three types of change in eligible segments are illustrated below. One type of change is if a long segment with a major portion in one county and a short portion in a different county has crashes located along the longer portion. The long portion is still eligible but the short portion is ineligible. Another change is the same segment as previous but all the crashes are located on the shorter portion. Now the short portion is eligible but the longer portion is ineligible. Still another type is a segment divided in half and has the crashes uniformly distributed along its length but has one more crash on one side of the county boundary then the other. This will in turn cause one side to remain eligible but the other side is ineligible.

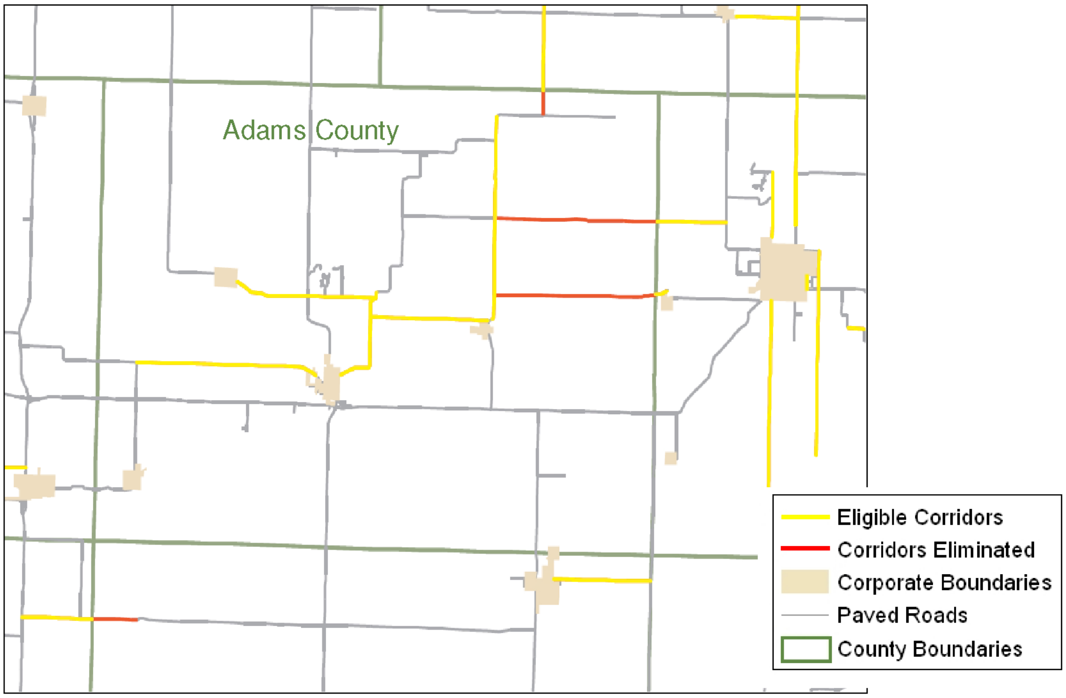


Figure 4.8. Eligible corridors with county constraints

The HRRR project also identified the top 15 percent of mileage of the eligible segments for both crash rate and crash density. The top 15 percent of mileage was identified using both segmentations. Using county constraint segmentation, more segments were included in the top 15 percent of mileage than segments without county constraints. This was because the segments were shorter in length using county constrained segmentation.

Figure 4.9 shows the top 15 percent of mileage using crash density and county constrained segmentation. The black segments are the top 5 percent of mileage and the red segments are the top 6 to 15 percent of mileage. Two segments on the left side of figure 4.9 show segments with very short length compared to their original length. The crashes for the original segment are located on the short portion. Having a short length and crashes associated to it, the short segment has a very high crash density and crash rate compared to longer segments. So these short segments are included in the top 15 percent of mileage.

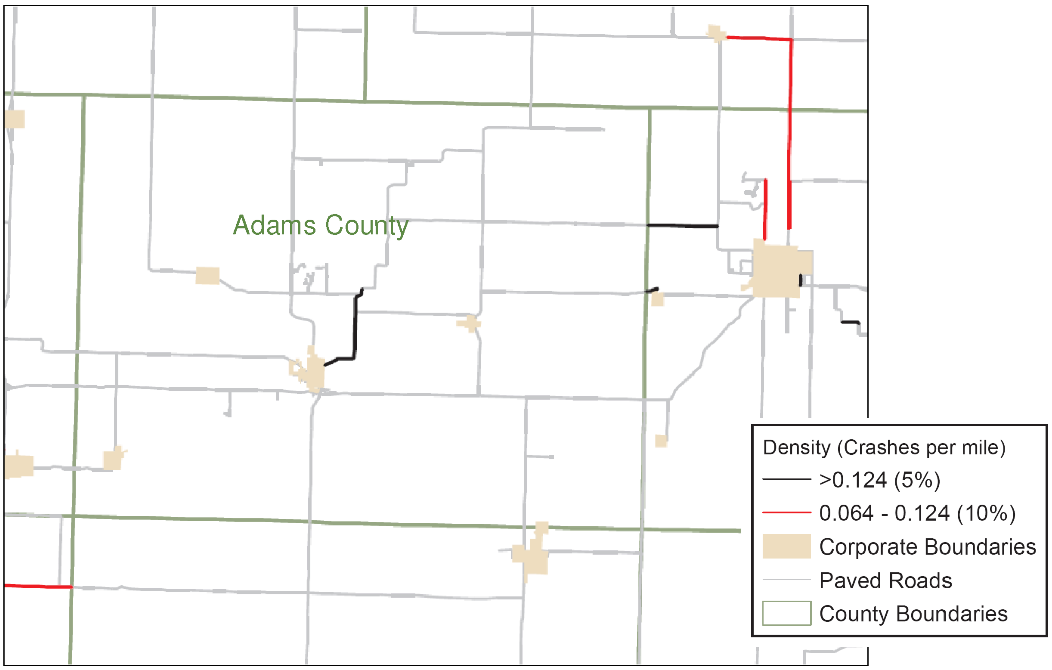


Figure 4.9. Top 15% of mileage using crash density with county constraints

When segmentation without county constraints is used as in figure 4.10, those segments are no longer included in the top 15 percent of mileage again identified by red or black segments. The resulting longer length of both the segments lowered the crash density.

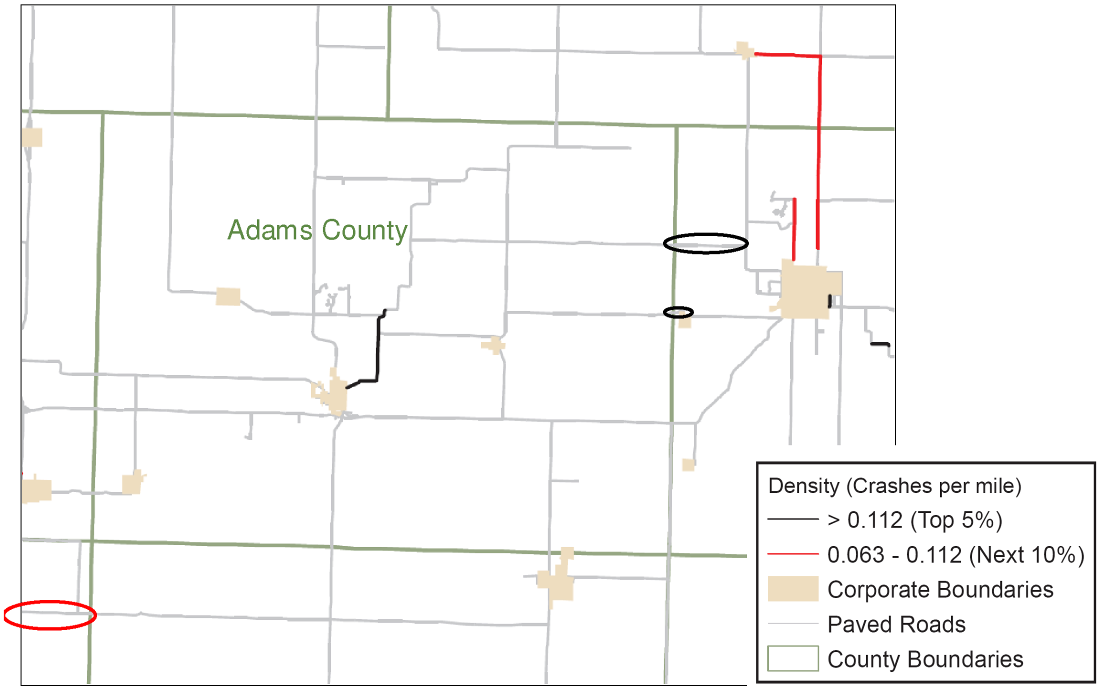


Figure 4.10. Top 15% of mileage using crash density no county constraints

Ultimately, segmentation without county constraints was used to identify the top 15 percent of mileage using crash rate and crash density. Although the project was indented to be used by county engineers and eligible segments are assessed within each county, it was decided to use longer segments as a true representative of the segment.

CHAPTER 5 CONCLUSIONS, LIMITATIONS, AND RECOMMENDATIONS

Conclusions

Identifying high crash locations is an important step in improving the safety of the highway network. This study has investigated the effects of various data preprocessing steps on identification of these locations as well as their impact on other highway safety analyses and procedures. The sensitivity of common crash rating schemes to data aggregation methodologies was tested for topics such as intersection crash assignment, segmentation, and exposure rate.

First, crashes were assigned to intersections at three different spatial proximities and then ranked the intersections. Descriptive statistics revealed only small differences in rank between the three methods. The greatest change in rank was observed when comparing crashes obtained fusing 100 foot buffers to those identified using 150 feet. The spatial proximity of 150 foot used as the baseline comparisons, was thought to be somewhat limiting in size in a project level analysis as opposed to a system wide analysis. In the limitations section, a rationale was developed for further assessment of proximity distances used in intersection crash assignment.

Although shifts in rank were minimal, distances use to assign crashes to intersections to may have a much larger effect on site studies, such as benefit cost analysis. Consider a before and after study of an intersection where an estimate of the number of crashes reduced (or to be reduced) by a mitigation is desired. As buffer size increases, so does the number of crashes that are assigned to the intersection, in both the before and after case. While increasing crashes in both periods may not affect a measure such as the ratio between before and after crashes it has a higher the potential to affect absolute change in the number of crashes. This change is fundamental to the calculation of benefits (reduced costs).

The effect of segmentation was tested using three different static, predefined lengths: two-miles, one-mile, and one-half mile. Locations were ranked using each of the three lengths. Using two-mile segments as a baseline, significant shifts in rank (average and lowest ranked segment) were observed as compared to the use of one-mile and one-half mile segmentation. Limited shifting was observed in the highest ranks of segments. When using one-mile segments as a baseline, a similar effect was observed although low rank one-half mile segments experienced largest shifts in rank. The smallest effects were observed where one-half mile segmentation was used as the baseline.

As expected, varying the crash rate exposure metric (denominator) between crash and number of vehicles involved had little effect on site crash ranking results, as Iowa composite scores are based only twenty percent on rate (the only composite input effected by exposure assumptions). If rate alone is used to rank sites, there is clearly a larger effect, but only below the top 20 locations, as most of these involve only single vehicle crashes. From rank 40-100, the use of crashes versus number of vehicles involved makes a significant difference.

Limitations

Complicating Factors

In this study, only three distances were used to buffer an intersection. In each case, the same buffer distance was used independent of road type. It would be more appealing to consider the characteristics of approach roads in determining the buffer distance for a particular intersection. Of course, additional data would be required along with the development of an automated process if thousands of intersections are to be processed. Assignment could also be based on physical intersection area or, more appropriately, its functional area. Roadway characteristics potentially useful in specifying this distance may include approach traffic volume, speed, and geometries.

Traffic volumes may be used in conjunction with geometry (capacities) to estimate congestion and queuing at an intersection. A crash may be considered intersection related if it occurs near or within the queue, which of course, varies over time. Figure 5.1 illustrates a functional intersection area that may be defined by end of queue. To begin to estimate queue lengths, at a minimum, hourly volumes and intersection geometries must be known. Statewide, it is unlikely that this information would be available. In the case of the illustration, had a spatial distance of 75 feet been used to assign crashes, a crash at back-of-queue would have not been included, though clearly intersection related. However, increasing the buffer distance to a degree that would catch all such crashes is likely to include some non-intersection crashes occurring in non-peak periods.

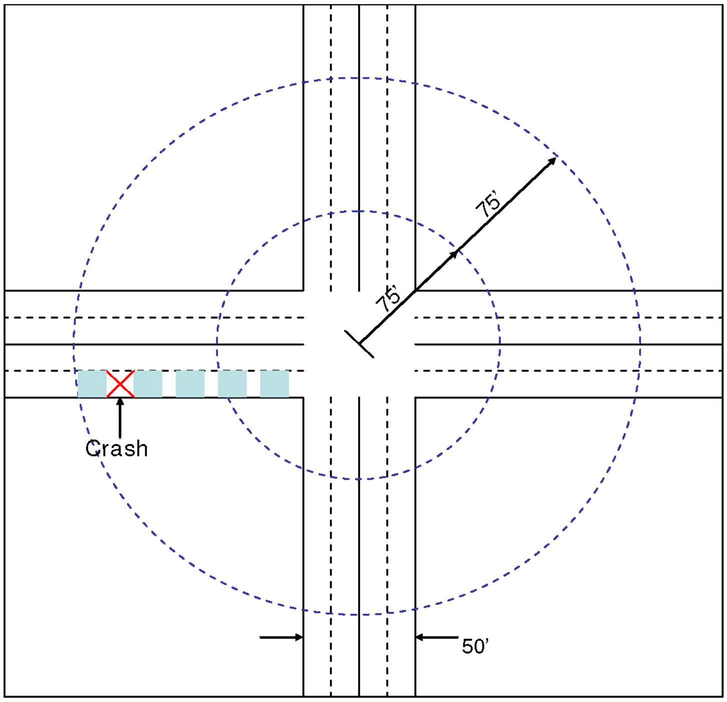


Figure 5.1. Crash at the end of queue

Proximity to driveways is also an important determinant of intersection relation. Had a 150 foot buffer been used in the case illustrated in Figure 5.2, the crash which is related to the side-road would have been assigned to the intersection. Further complicating this situation would be a queue extending from the intersection beyond the driveway which would require additional information beyond physical location for decision.

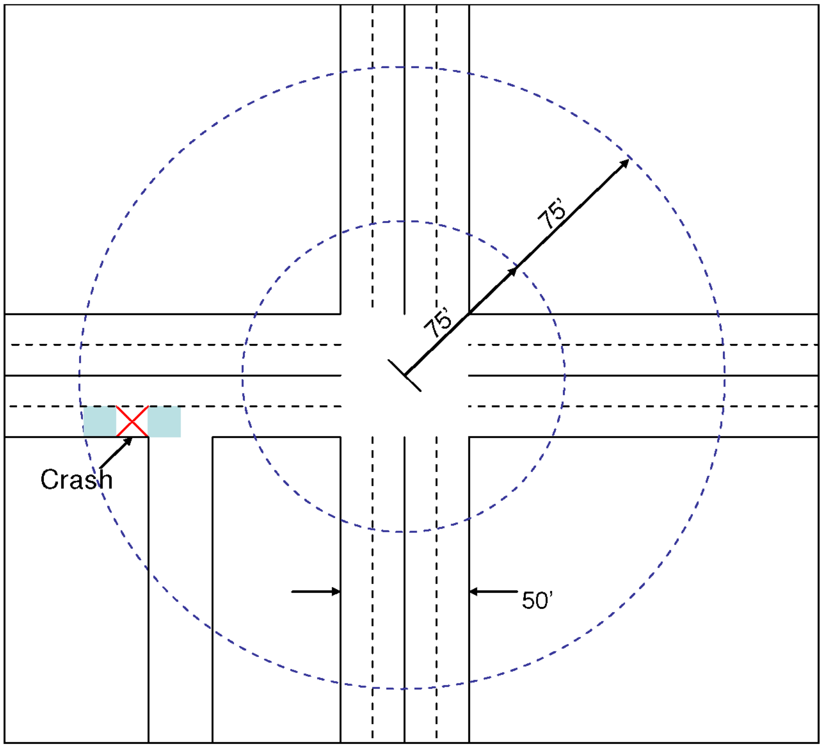


Figure 5.2. Driveway proximity to intersection

The location of proximate intersections should also affect intersection assignment. If one is simply interested in the question of a particular crash being associated with any intersection, it is a moot point. However, if the number of crashes associated with a particular intersection is desired, buffers may have to be designed to fall midway between the intersections. All of the complicating factors listed above then also come into play. See figure 5.3.

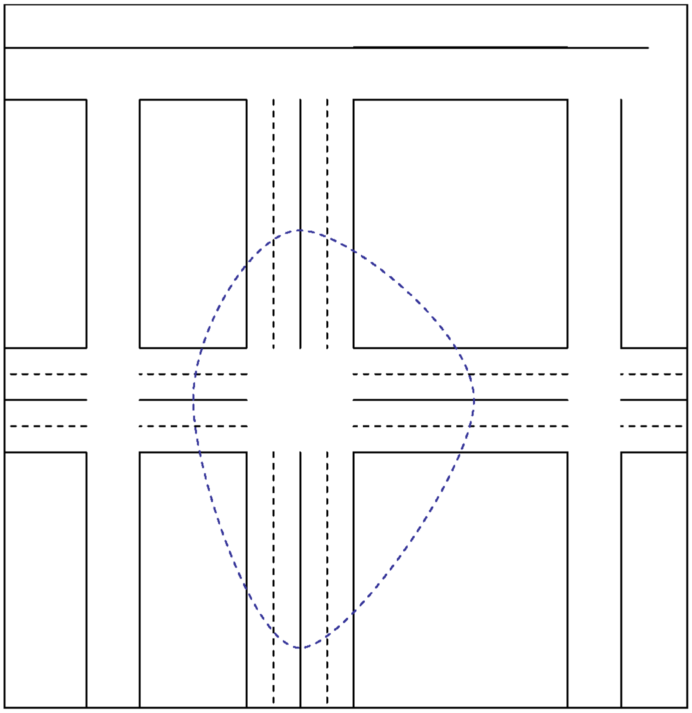


Figure 5.3 Intersection legs with different buffer distances

Finally, Figure 5.4 illustrates yet another complication of intersection assignment. Direction of travel is clearly important to the question. Short of highly precise GPS coordinates, inbound or outbound direction of travel is required to determine whether the crash is intersection related. Spatial proximity alone cannot address this complication.

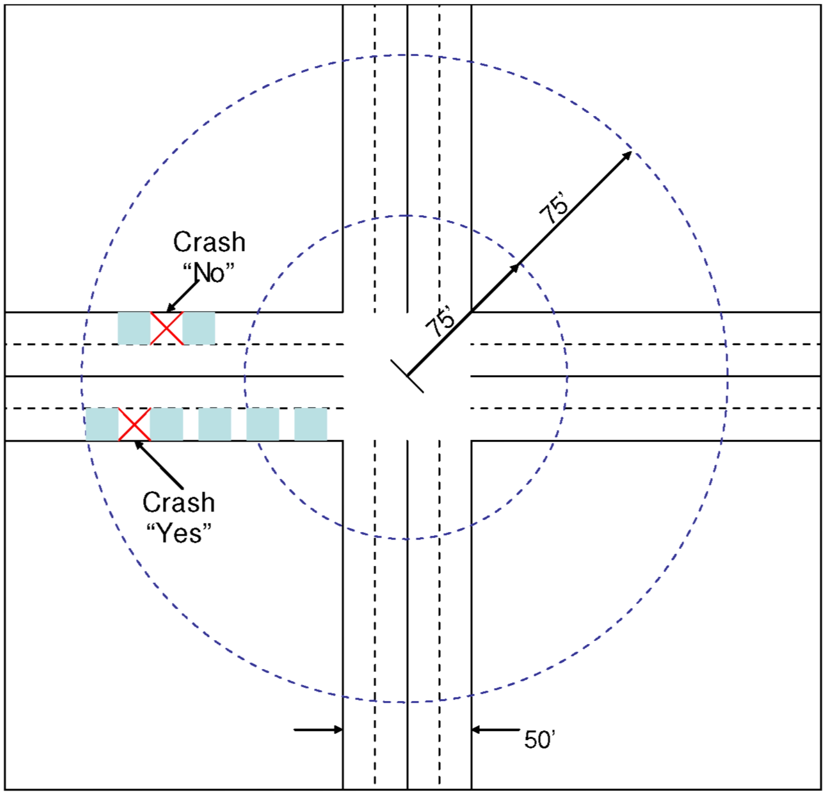


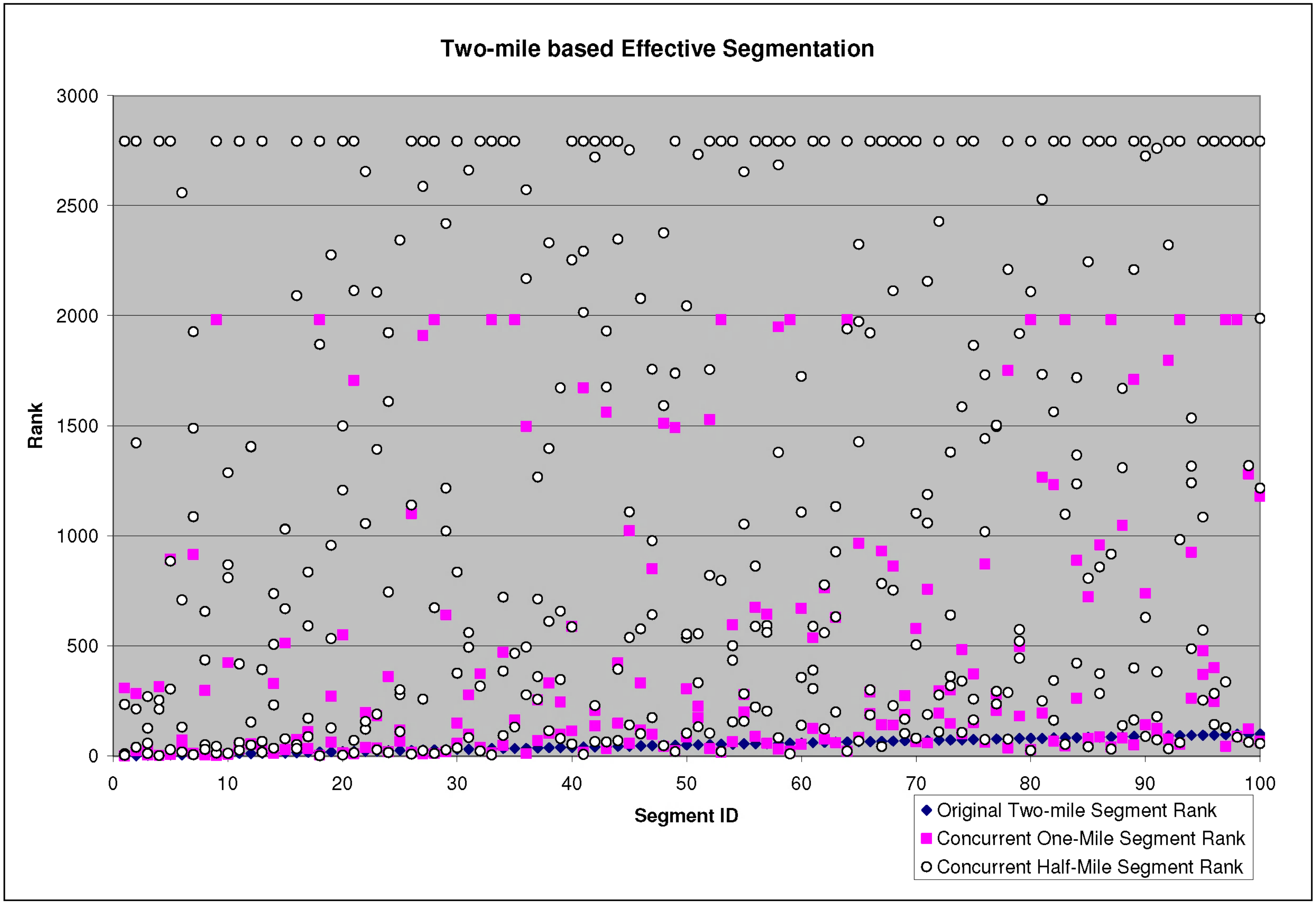
Figure 5.4. Direction of travel

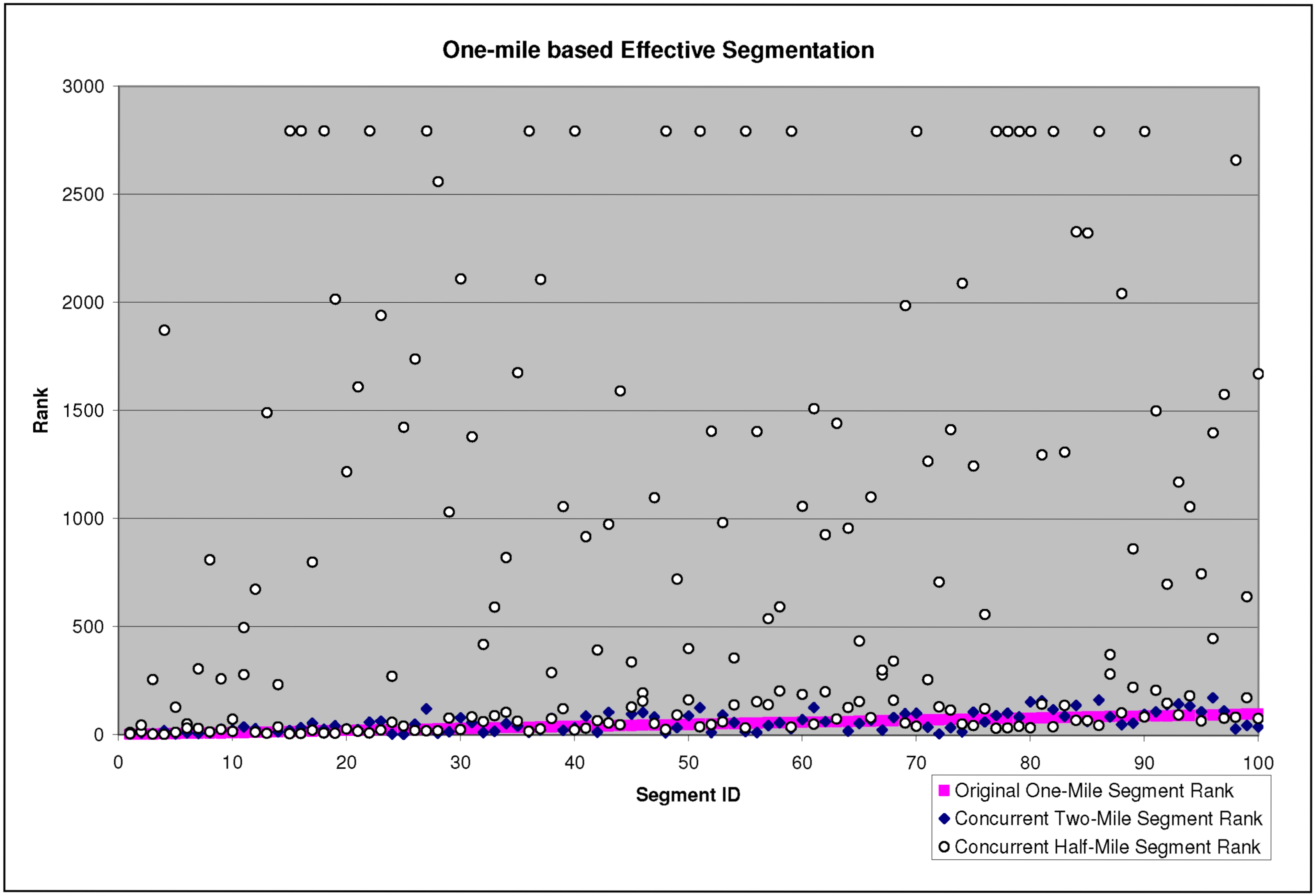
Recommendations

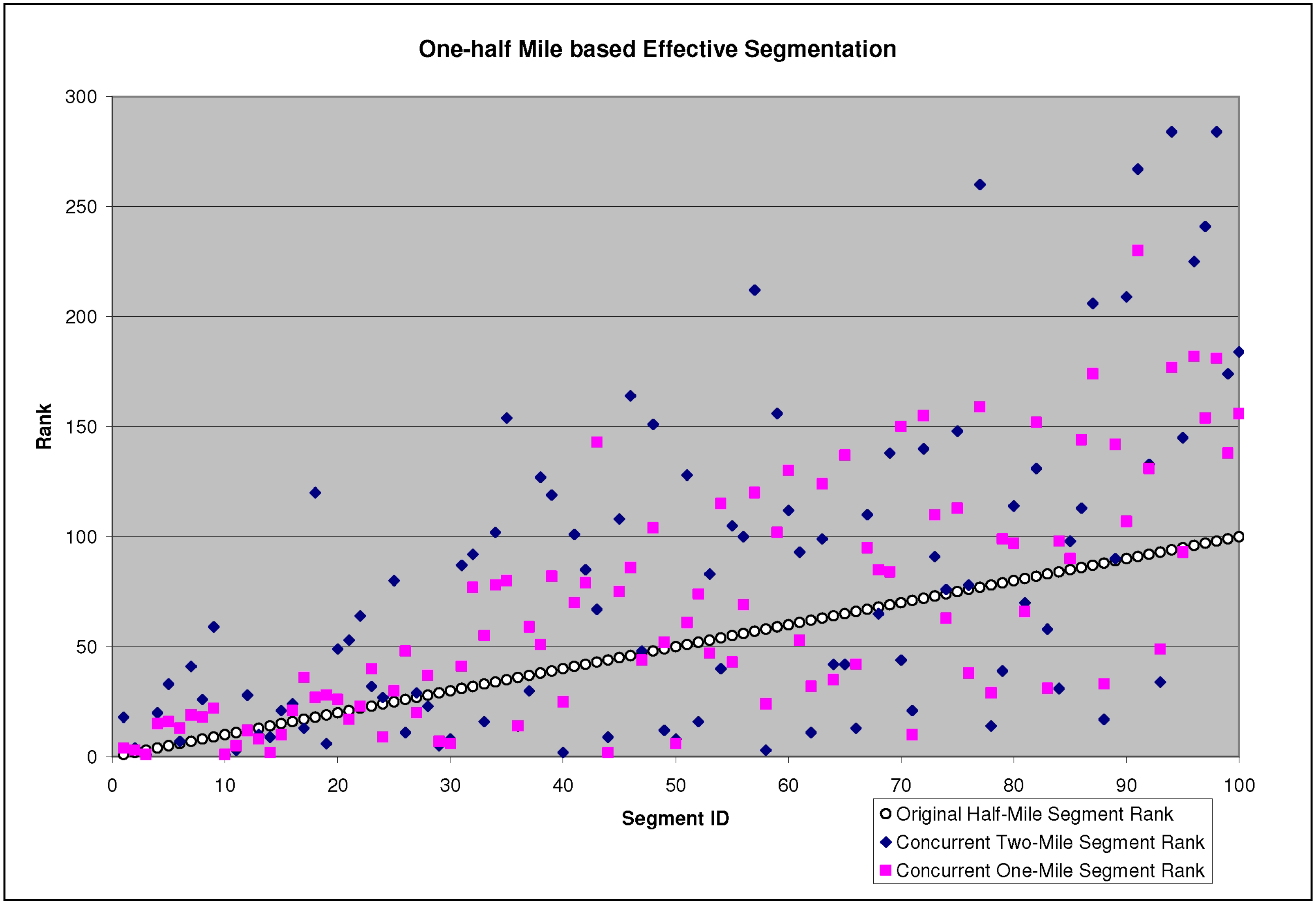
Based on the results of this study it is recommended that further research be conducted in both intersection assignment and segmentation for identification of high crash locations. Such research was limited in this study by the format of the available Iowa data.

More detailed study of the assignment of crashes to intersections using crash attributes is also recommended. This may require careful examination of original crash reports and narratives. Given adequate data from intersection related crashes, an attribute matrix could be compiled, enabling the use of attributes and spatial proximity for intersection crash assignment on a system level.

For segmentation, it is recommended that shorter (one-half mile in this study) segments be used in analysis. However, segments that are too short may lead to difficulties in developing statistically robust models of crash location and analysis (the small sample size problem). This phenomenon would form the basis for an interesting and useful study. Studies of the effect variable segment lengths ad of fixed and variable length sliding scale are also recommended.


Two-mile based Effective Segmentation
Rank
3000
2500
2000
1500
1000
500
0
0
10
20
30
40
50
60
70
80
90
100
Segment ID
Original Two-mile Segment Rank
Concurrent One-Mile Segment Rank
Concurrent Half-Mile Segment Rank


One-mile based Effective Segmentation
Rank
Segment ID
0
500
1000
1500
2000
2500
3000
0
10
20
30
40
50
60
70
80
90
100
Original One-Mile Segment Rank
Concurrent Two-Mile Segment Rank
Concurrent Half-Mile Segment Rank


One-half Mile based Effective Segmentation
Rank
Segment ID
0
50
100
150
200
250
300
0
10
20
30
40
50
60
70
80
90
100
Original Half-Mile Segment Rank
Concurrent Two-Mile Segment Rank
Concurrent One-Mile Segment Rank

REFERENCES

Agent, K., L. O'Connell, E. Green, D. Kreis, J. Pigman, and N. Tollner, *Development of Procedures for Identifying High-Crash Locations and Prioritizing Safety Improvements, Research Report KTC-03-15/SPR250-02-1F*. Kentucky Transportation Center, University of Kentucky, Lexington, KY. 2003.

Black, William R. *Highway Accidents: A Spatial and Temporal Analysis.* Transportation Research Record 1318. Transportation Research Board. 1992.

City of Mankato, Minnesota. *Mankato/North Mankato Area Transportation and Planning*. SRF Consulting Group, Inc. June 2003.

Drakopoulos, Ertan. *Addressing Non-Intersection Crashes Statewide: Identification and Ranking of Highway Segments in Need of Safety Upgrades Using a Floating Highway Segment Tool Integrated with GIS.* Proceeding of 2005 Mid-Continent Transportation Research Symposium, Ames, Iowa. 2005

Fatality Analysis Reporting System (FARS), accessible online at http://www-fars.nhtsa.dot.gov/ Access date: 10/15/2006.

Federal Highway Administration (FHWA). *AIRS Approach to Analyzing Intersection Crashes*. Public Roads. March/April 2004.

Federal Highway Administration (FHWA). *SAFETEA-LU Guidance Highway Safety Improvement Program 23 U.S.C 148(c)(1)(D) "5 Percent Report"*, accessible online at http://safety.fhwa.dot.gov/safetealu/fiveguidance.htm. Access date: 6/1/2006. (a)

Federal Highway Administration (FHWA). *SAFETEA-LU High Risk Rural Roads Interim Guidance,* accessible online at http://safety.fhwa.dot.gov/safetealu/hrrrpattachment.htm. Access date: 5/31/2006. (b)

Geyer, J., E. Lankina, D. Ragland, C. Chan, and A. Sharafsaleh. *Literature Survey of Methods for Identifying High Concentration Collision Locations (HCCL).* Traffic Safety Center and California PATH University of California, Berkeley. 2005.

GIS/Trans, LTD. *GIS Safety Analysis Tools Version 3.0 User Guide.* Federal Highway Administration. 2000.

Green, E., and K. Agent. *Crash Rates at Intersection*. Kentucky Transportation Center College of Engineering. August 2003.

Hallmark, Shauna L. *Evaluation of the Iowa DOT's Safety Improvement Candidate List Process*. Center for Transportation Research and Education Iowa State University. June 2002.

Hauer, E., J. Kononov, B.Allery, and M. Griffith. *Screening the Road Network for Sites with Promise*. Transportation Research Record, TRB, No. 1784, National Research Council, Washington, D.C., 2002.

Iowa Dot. *Accident and Related Data for Rural and Municipal Intersections in Iowa*. Bureau of Transportation Safety. 1989.
Iowa DOT. *Highway Safety Strategies for Iowa Executive Summary of the Iowa SMS Toolbox*. Iowa Highway Safety Management System. 2003.

Iowa DOT. *Iowa Safety Improvement Candidate Location (SICL) Method, accessible online at* http://www.dot.state.ia.us/crashanalysis/pdfs/iowa_safetyimprovementcandidatelocation_method_20060228.pdf. Access date 6/21/2006.

Knox, Todd. *Safety of High Speed Expressway Signals: a Comparison of Classical and Empirical Bayes Methods*. Iowa State University. Ames, Iowa. 2005.

Lankina, E., C. Chan, D. Ragland, T. Pham, and A. Sharafsaleh. *Technical Approaches and Data Review for Identifying High-Concentration Collision Locations*. Traffic Safety Center and California PATH University of California, Berkley. 2005.

Markos Papageorgiou, *Fellow,* IEEE, and Apostolos Kotsialos. *Freeway Ramp Metering: an Overview*. IEEE Transactions on Intelligent Transportation Systems, Vol. 3, No. 4. 2002.

National Cooperative Highway Research Project (NCHRP). *NCHRP Report 17-18(3) Signalized Intersections,* accessible online at http://safety.transportation.org/htmlguides/sgn_int/exec_sum.htm Access date 5/24/2006.

National Highway Traffic Safety Administration (NHTSA), DOT. *State Traffic Safety Information System Improvement Grants*, accessible online at http://www.dottrcc.gov/pages/TrafSafetyGrants.html. Access date 9/20/2006**.**

Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) August 2005. http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=109_cong_public_laws&docid=f:publ059.109.pdf

Sando, T., R. Mussa, H. Wu, J. Sobanjo, and L. Spainhour. *A Cost-Effective GIS Safety Analysis Tool for Improving Highway Safety*, accessible online at http://gis.esri.com/library/userconf/proc04/docs/pap1148.pdf. Access date 5/24/2006.

Strathman, J., K. Dueker, J. Zhang, and T. Williams. *Analysis of Design Attributes and Crashes on the Oregon Highway System.* Center for Urban Studies, College of Urban and Public Affairs. Portland State University. 2001.

Thomas, Isabelle. *Spatial Data Aggregation: Exploratory Analysis of Road Accidents.* National Fund for Scientific Research, Department of Geography, Universite Catholique de Louvain. 1995.

Vogt, Andrew. *Crash Model for Rural Intersections: Four-Lane by Two-Lane Stop-Controlled and Two-Lane by Two-Lane Signalized.* FHWA 1999.