

Field Test Analysis Information Report

Foreword—This SAE Information Report describes results of testing of the SAE J1746 ISP-Vehicle Standard. Testing was performed by the Oak Ridge National Laboratory (ORNL) and its contractors, resulting in a document from which this Information Report has been extracted. Location referencing Standards enable the location of an event or thing to be communicated from one component of an ITS system to another component in an unambiguous and mutually understandable way. The SAE J1746 ISP-Vehicle Standard was designed for the communication of spatial data references between central sites and mobile vehicles on roads. The ORNL testing program determined and characterized the ability of the J1746 standard to transfer information between various combinations of commonly available public and private sector databases, and makes conclusions and recommendations concerning error sources in unsuccessful transfers.

Testing was performed by the Oak Ridge National Laboratory and its contractors and reported in the document "Location Referencing Message Specification Test Report, Version 1.01, September 23, 1998, R. Haas, J. Lau, C. Goodwin, and S. Gordon," for a project sponsored by the U.S. Federal Highway Administration (FHWA). Tests were performed by computer analysis and corroborated by field tests with a mobile vehicle.

TABLE OF CONTENTS

1.	Scope	4
2.	References	4
2.1	Applicable Documents.....	4
2.2	Related Publications.....	4
3.	Definitions.....	6
4.	Executive Summary.....	8
5.	Overview	9
5.1	Overview of Test Results.....	9
6.	Introduction to the LRMS and LRMS Testing.....	10
6.1	LRMS Overview.....	11
6.2	Testing Overview.....	11
6.3	Factors that Influence the Effectiveness of an LRMS Profile	11
6.4	Test Data	19
6.5	LRMS Profile Testing.....	21

SAE Technical Standards Board Rules provide that: "This report is published by SAE to advance the state of technical and engineering sciences. The use of this report is entirely voluntary, and its applicability and suitability for any particular use, including any patent infringement arising therefrom, is the sole responsibility of the user."

SAE reviews each technical report at least every five years at which time it may be reaffirmed, revised, or cancelled. SAE invites your written comments and suggestions.

TO PLACE A DOCUMENT ORDER; (724) 776-4970 FAX: (724) 776-0790
SAE WEB ADDRESS <http://www.sae.org>

7.	The Cross-Streets Profile	22
7.1	Description of the Cross-Streets Profile	22
7.2	Using the Cross-Streets Profile	23
7.3	Failure Modes when Using the Cross-Streets Profile	24
7.4	The Cross-Streets Profile Testing Protocol	25
7.5	Preliminary Testing	28
7.6	Phase I Testing	31
7.7	Phase II Testing	34
7.8	Summary of Results	34
Appendix A	Cross-streets Profile Test Results	37
A.1	Preliminary Testing - The Data Used in the Cross-Streets Profile	37
A.2	Phase I Testing - Using Cross-Streets to Transmit Location References	49
A.3	Phase II Testing - Using Coordinates and Cross-Streets to Transmit Location References	58
A.4	Introduction	65
A.5	The Cross-Streets Profile - Description	65
A.6	XSP Problem Scenarios	69
A.7	Definitions and Evaluation Metrics	70
A.8	Experimental Procedure and Analytical Results	71
A.9	Discussion of Results	76
A.10	Conclusions and Recommendations	78
Appendix B	Cross-streets Profile Phase II Test Results - VITAL	79
B.1	Introduction	79
B.2	Test Dimensions	79
B.3	Experimental Design	81
B.4	Accuracy of Source Coordinate Transfer	85
B.5	Fan-Out	86
B.6	Results	86
B.7	Conclusions	88

TABLE OF FIGURES

Figure 1	Example Topological Discrepancies	12
Figure 2	Example Topological Ambiguities	13
Figure 3	Example Coordinate Discrepancies	14
Figure 4	Example Street Naming Discrepancies	14
Figure 5	Example Cross-Streets Ambiguities	15
Figure 6	Example Coverage Differences	17
Figure 7	Example of Differences in Level of Detail	18
Figure 8	The Cross Streets Profile	22
Figure 9	Typical Usage of the Cross-Streets Profile	27
Figure A1	The Test Data Set SB-A	38
Figure A2	The Test Data Set SF-D	38
Figure A3	Area Defined as Santa Barbara City Data	39
Figure A4	The Cross Streets Profile in a Nutshell	65
Figure A5	Topological Error Situation	69
Figure A6	Multiple-Intersection Ambiguities	70
Figure A7	Processing of Field Test Measurements	72
Figure B1	Downtown Santa Barbara	82
Figure B2	Name Processing and Coordinate Fallback	85

TABLE OF TABLES

Table 1	Failure Modes When Transmitting Location Reference Information	12
Table 2	Occurrence Frequency of Topological Ambiguities	13
Table 3	Occurrence Frequency of Cross-Streets Ambiguities	15
Table 4	Data Issues that Impact LRMS Profile Effectiveness	16
Table 5	Differences in Level of Completeness	18
Table 6	Other Factors that Influence Profile Effectiveness	19
Table 7	Data Set Code Names	19
Table 8	Types of Test Points for LRMS Profile Testing	20
Table 9	Possible Sampling and Weighting Techniques for Commercial Data Set Test Points	20
Table 10	Some Methodologies for Using the Cross-Streets Profile to Identify a Link	23
Table 11	Some Techniques for Modifying Link Matching Methodologies	23
Table 12	Alternate Uses of the Cross-streets Profile	24
Table 13	Failure Modes for Identifying Links Using the Cross-streets Profiles	24
Table 14	Failure Modes Using Offsets in the Cross-streets Profile	25
Table A1	Blank Street Names in the Santa Barbara County Data Sets	39
Table A2	Types of Nodes	41
Table A3	Types of Cross-Street References for Nodes	41
Table A4	Types of Cross-Street References for Links	42
Table A5	Links with Blank Street Names	43
Table A6	Methods of Street Name Normalization Used in the LRMS Testing	43
Table A7	Some Translations Used for Commonly Occurring Words	44
Table A8	Street Name Matching Using Different Methods of Street Name Normalization	45
Table A9	Reasons Why Method 3 Normalized Street Names Do Not Match	46
Table A10	Distance Between Nearest Nodes in Test Data Sets	47
Table A11	Distance Between Matching Nodes in Test Data Sets	48
Table A12	Comparison of Cross-Streets Matching Nodes and Closest Nodes	49
Table A13	Summary of Test Methods	50
Table A14	Cross-Street Node Matching, Test Method 1	50
Table A15	Cross-Street Node Matching, Test Method 2	51
Table A16	Cross-Street Link Matching, Test Method 1	52
Table A17	Cross-Street Link Matching for Links with Mixed Results; Test Method 1	53
Table A18	Links That Match to Multiple Links, Test Method 1	54
Table A19	Cross-Streets Link Matching, Test Method 1, Street Name Normalization 1	54
Table A20	Cross-Streets Link Matching, Test Method 1, Street Name Normalization 5	55
Table A21	Cross-Streets Link Matching, Test Method 1, Weighted by Link Length	56
Table A22	Cross-Streets Link Matching, Test Method 1, Restricted to Primary Roadways	57
Table A23	Cross-Streets Link Matching, Test Method 1, Restricted to Secondary Roadways	58
Table A24	Cross-Streets Link Matching, Test Method 1, Corrected for Valence 2 Nodes	60
Table A25	Cross-Streets Link Matching, Test Method 1, Corrected for Valence 2 Nodes and Dead-End Links	61
Table A26	Node Point Chart for Advanced Cross-Streets Profile Matching	62
Table A27	Link Point Chart for Advanced Cross-Streets Profile Matching	63
Table A28	Result Bins for Advanced Cross-Streets Profile Matching	63
Table A29	Cross-Streets Link Matching, "Best Methods"	64
Table A30	Incidence of Blank Street Names	67
Table A31	Street Name Components	67
Table A32	Database A Fields	67
Table A33	Database B/F Fields	68
Table A34	Database E Fields	68
Table A35	Name Matching Results for Goleta Test Area	73
Table A36	Name Matching Results for Entire County of Santa Barbara	74
Table A37	Transfer Results (Field Tests)	75
Table A38	Transfer Results (Lab Tests)	76

Table A39	Lab and Field Results Compared	76
Table B1	Incidence of Blank Triads	80
Table B2	Scoring of Transfers	85
Table B3	Result Summary Test Set B (Lab Generated Single Point Transfer).....	86
Table B4	Result Summary, Test Set C (Lab Generated Single Point Transfer)	87
Table B5	Generated Points (Single Point Transfer).....	90
Table B6	Generated Points (Two Point Independent Transfer)	90
Table B7	Generated Points (Two Point Simultaneous Transfer)	90
Table B8	Field Data (Single Point Transfer)	91
Table B9	Field Data (Two Point Independent Transfer).....	91
Table B10	Field Data (Two Point Simultaneous Transfer).....	91
Table B11	Generated Points (Single Point Transfer	92
Table B12	Generated Points (Two Point Transfer)	93
Table B13	Field Data (Single Point Transfer)	94
Table B14	Field Data (Two Point Transfer).....	95
Table B15	Generated Points (Single Point Transfer, Major Streets Only)	96
Table B16	Generated Points (Two Point Transfer, Major Streets Only)	96
Table B17	Santa Barbara County, All Streets.....	97
Table B18	Santa Barbara County, Major Streets Only	98

1. **Scope**—This SAE Information Report describes results of testing of the SAE J1746 ISP-Vehicle Standard for the communication of spatial data references between central sites and mobile vehicles on roads. Testing was performed by the Oak Ridge National Laboratory and its contractors, resulting in a document from which this Information Report has been extracted. Tests were performed by computer analysis and corroborated by field tests with a mobile vehicle.

2. References

2.1 **Applicable Publications**—The following standards contain provisions which, through references in this text, constitute provisions of this document. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this document are encouraged to investigate the possibility of applying the most recent editions of each standard listed below.

2.1.1 SAE PUBLICATIONS—Available from SAE, 400 Commonwealth Drive, Warrendale, PA 15096-0001.

SAE J1746—ISP-Vehicle Location Referencing Standard

SAE J2374—Vehicle Location Referencing Message Specification

2.2 **Related Publications**—The following publications are provided for information purposes only and are not a required part of this document.

2.2.1 AMERICAN ASSOCIATION OF STATE HIGHWAY AND TRANSPORTATION OFFICIALS (AASHTO) PUBLICATIONS—Available from AASHTO, Suite 225, 444 N. Capital St. NW, Washington, DC 2000.

GIS-T '95—Reinterpreting the Location Referencing Problem: A Protocol Approach—Proceedings, Geographic Information Systems for Transportation Symposium (GIS-T)—Reno, NV, April 3-5, 1995, Goodwin, C., Gordon, S., and Siegel, D.

2.2.2 FEDERAL HIGHWAY ADMINISTRATION DOCUMENTS—Available from U.S. DOT, FHWA, Office of Safety and Traffic Operations, ITS Research Division, HSR 10, 6300 Georgetown Pike, McLean, VA 22101.

- Task A: Spatial Data Interoperability Protocol For ITS Project—Location Reference Protocol Requirements, December 29, 1995, Goodwin, C., Siegel, D. and Gordon, S.
- Task B: Spatial Data Interoperability Protocol For ITS Project—Location Reference Message Protocol Preliminary Specification, February 29, 1996, Goodwin, C., Siegel, D. and Gordon, S.
- Task B: Spatial Data Interoperability Protocol For ITS Project—Location Reference Message Specification: Final Design, June 28, 1996, Goodwin, C., Siegel, D. and Gordon, S.
- Task B: Spatial Data Interoperability Protocol for ITS Project—Location Reference Message Specification: Revision A, September 18, 1996, Goodwin, C., Siegel, D. and Gordon, S.
- Task B: Spatial Data Interoperability Protocol for ITS Project—Location Reference Message Specification: Revision B (MDI), May 22, 1997, Goodwin, C., Siegel, D. and Gordon, S.
- Task B: Spatial Data Interoperability Protocol for ITS Project—Location Referencing Message Specification Test Report, Version 1.01, September 23, 1998, Haas, R., Lau, J., Goodwin, C. and Gordon, S.
- Task C: Spatial Data Interoperability Protocol For ITS Project—The ITS Datum Preliminary Data Structure and Content, February 25, 1996, Goodwin, C., Siegel, D. and Gordon, S.
- Task C: Spatial Data Interoperability Protocol For ITS Project—ITS Data Final Design Report, June 28, 1996, Goodwin, C., Siegel, D. and Gordon, S.
- Task A3: Nationwide Map Database and Location Referencing System Project—Functional Requirements for National Map Databases for ITS, December 31, 1994, Goodwin, C., Xiong, D. and Gordon, S.
- Task B3: Nationwide Map Database and Location Referencing System Project—Technical Requirements for National Map Databases for ITS, July 31, 1995, Goodwin, C., Gordon, S.
- Task F1: Nationwide Map Database and Location Referencing System Project—Recommendation for Location Referencing for ITS, December 31, 1994, Goodwin, C., Xiong, D. and Gordon, S.
- Task B: Spatial Data Interoperability Protocol for ITS Project—Location Reference Message Specification, Revision B (MDI), May 22, 1997, Goodwin, C., Siegel, D. and Gordon, S.
- “Location Referencing for ITS,” White Paper prepared for Oak Ridge National Laboratory and the National ITS Architecture Program, November, 1995, 18 pp., Goodwin, Cecil W. H.

2.2.3 NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY (NIST) DOCUMENT—Available from NIST, FIPS PUB 173, U.S. Government Printing Office, National Technical Information Service U.S. Department of Commerce, Springfield VA 22161

FIPS PUB 173—Spatial Data Transfer Standard (SDTS), 1992

2.2.4 OAK RIDGE NATIONAL LABORATORY DOCUMENTS—Related to spatial data interoperability are available on the World-Wide Web at: <http://itsdeployment.ed.ornl.gov/spatial>.

- “The Location Reference Message Protocol for ITS Applications,” White Paper prepared for Oak Ridge National Laboratory project ‘Nationwide Map Database and Location Referencing System’, December 1994, 19 pp., Goodwin, Cecil W. H. and Xiong, Demin
- “Reinterpreting the Location Referencing Problem: A Protocol Approach,” in press for Proceedings of GIS-T ’95, AASHTO, Reno, April 3-5, 1995, Goodwin, Cecil W. H., Gordon, Stephen R. and Siegel, Dave.
- “Recommendation for Location Referencing for ITS Needs,” Task F1: Nationwide Map Database and Location Referencing System Project, United States Department of Transportation, Federal Highway Administration Office of Safety and Traffic Operations, ITS Research Division, Contract DTFH61-91-Y-30066, Draft, June 30, 1995, 78 pp., Goodwin, Cecil W. H., Siegel, Dave, Gordon, Stephen R., and Xiong, Demin 1995
- “Location Reference Message Protocol Requirements,” Task A: Spatial Data Interoperability Protocol For ITS Project, United States Department of Transportation, Federal Highway Administration Office of Safety and Traffic Operations, ITS Research Division, Contract 61-94-Y-00001, Review Draft,

December 29, 1995, 64 pp., Goodwin, Cecil W. H., Siegel, Dave, and Gordon, Stephen R., 1995
 "Location Reference Message Protocol Preliminary Specification," Task B: Spatial Data Interoperability Protocol For ITS Project, United States Department of Transportation, Federal Highway Administration Office of Safety and Traffic Operations, ITS Research Division, Contract 61-94-Y-00001, Review Draft, February 29, 1996, 34 pp., Goodwin, Cecil W. H., Siegel, David and Gordon, Stephen R. 1996
 "Location Reference Message Specification: Final Design," Task B: Spatial Data Interoperability Protocol For ITS Project, United States Department of Transportation, Federal Highway Administration Office of Safety and Traffic Operations, ITS Research Division, Contract 61-94-Y-00001, Draft, June 28, 1996, 40 pp., Goodwin, Cecil W. H., Siegel, David and Gordon, Stephen R. 1996

2.2.5 VIGGEN CORPORATION DOCUMENTS—Related to location referencing and spatial data interoperability are available on the World-Wide Web at <http://www.locref.com>.

FHWA Location Referencing System for ITS User Services White Paper: "An Alternative Approach: Coordinate Snapping," November, 1994, Vigen Corporation, 1994
 FHWA Location Referencing System for ITS User Services White Paper: "Street Naming Conventions," November, 1994, Vigen Corporation, 1994
 FHWA Location Referencing System for ITS User Services White Paper: "An Alternative Approach: Cross-streets," November, 1994, Vigen Corporation, 1994

3. Definitions

- 3.1 **Altitude**—Elevation above or below a reference datum, as defined in FIPSPUB 70-1; the z-value in a spatial address.
- 3.2 **Coordinates**—Pairs of numbers expressing horizontal distances along orthogonal axes; alternatively, triplets of numbers measuring horizontal and vertical distances. (NIST, 1992)
- 3.3 **Datum**—A set of parameters and control points used to accurately define the three-dimensional shape of the Earth (e.g., as an ellipsoid). The corresponding datum is the basis for a planar coordinate system. (MnDOT 1993)
- 3.4 **Geodetic Datum**—A mathematical model of the Earth's shape. (Krzanowski et al, 1993) A geometric set of five quantities which serves as a locational reference or base for other quantities. The five quantities are the latitude and longitude of an initial point, the azimuth of a line from this point and two constants necessary to define the terrestrial spheroid. (MnDOT 1993)
- 3.5 **Information Service Provider (ISP)**—In the context of the LRMS, a public or private entity which may be responsible for gathering, fusing, analyzing, and/or reporting transportation related information to users, including vehicles and non-mobile users.
- 3.6 **Junction**—A collection of more than one node that represents a logical feature, such as a complex intersection. (Shuman, 1993)
- 3.7 **Link**—A topological connection between two nodes. A link may contain additional intermediate coordinates (shape points) to better represent the shape of curved features (Shuman, 1993). A link may be directed by ordering its nodes.
- 3.8 **Link Identifiers (or Link-ID)**—An identifier assigned to a link. Link-IDs may be arbitrary, or may be assigned by convention to assure that multiple occurrences of the same ID will not occur within one network or within the universe of similar networks or databases.
- 3.9 **Location Referencing System**—System of determining the position of an entity relative to other entities or to some external frame of reference.

- 3.10 Measured Distance**—Cumulative distance measured along a road centerline from a reference node.
- 3.11 NAD (North American Datum)**—The official reference ellipsoid used for the primary geodetic network in North America.
- 3.12 NAD27 (North American Datum 1927)**—Geodetic datum defined by the geographic position of triangulation station Meades Ranch and the azimuth from that station to station Waldo on the Clarke spheroid of 1866. (MnDOT 1993)
- 3.13 NAD83 (North American Datum 1983)**—Geodetic datum defining an Earth-fitting ellipsoid and for relating existing coordinate systems to the Earth's center of mass. It is based on both satellite and terrestrial data used in developing the Geodetic Reference System of 1980 and later the World Geodetic System of 1984 (WGS-84). This improved Earth-centered model slightly changes the latitude and longitude of almost every point in North America with respect to the earlier NAD27 datum. (MnDOT 1993)
- 3.14 Node (0-cell)**—A topologically significant point, such as a simple intersection of roadways or other linear features, or an endpoint of such a feature. (Shuman, 1993)
- 3.15 Node Identifier (or Node-ID)**—An identifier assigned to a node. Node-IDs may be arbitrary, or may be assigned by convention to assure that multiple occurrences of the same ID will not occur within one network or within the universe of similar networks or databases.
- 3.16 Reference Ellipsoid**—The mathematical model of the Earth used in geodetic computations. (Krzanowski et al, 1993)
- 3.17 Reference Node**—A node with respect to which distance measurements are made. In the ISP-Vehicle profile, Reference Nodes are simple intersection or endpoint nodes.
- 3.18 Topology**—The logical relationships among map features in a digital base map. It can be used to characterize spatial relationships such as connectivity and adjacency.
- 3.19 World Geodetic System of 1984 (WGS-84)**—An earth-centered global reference frame, including an earth model, based on satellite and terrestrial data. It contains primary parameters that define the shape, angular velocity, and the earth mass of an earth ellipsoid, and secondary parameters that define a gravity model of the earth. Primary parameters are used to derive latitude-longitude coordinates (horizontal datum). Secondary parameters are used for determining the orbits of GPS navigation satellites and define a coarse vertical datum; therefore WGS-84 is a complete geodetic system. EGM-96 is a refined model established in 1996.

3.20 Acronyms

FHWA	Federal Highway Administration
IR	Information report
ISP	Information Service Provider
ITS	Intelligent Transportation Systems
LRM	Linear Referencing Model
LRMS	Location Referencing Message Specification
MDI	Model Deployment Initiative
ORNL	Oak Ridge National Laboratory
SAE	Society of Automotive Engineers, Inc.
VITAL	Vehicle Intelligence & Transportation Analysis Laboratory (formerly Vehicle Intelligence Testing and Analysis Laboratory)

4. **Executive Summary**—This document is a report on testing being conducted on a specification for location referencing for ITS. This work was performed as part of a project being carried out by US Oak Ridge National Laboratory (ORNL) and Viggen Corporation, its primary subcontractor, sponsored by the US Federal Highway Administration (FHWA). The LRMS was developed to support the transfer of location information as part of standard messages being passed between ITS applications.

Location referencing has emerged as a critical component of the ITS National Architecture and is part of virtually all ITS message sets. The variation in the types of spatial information transferred in ITS messages creates the demand for a flexible referencing system. These include point locations of accidents or emergencies, link locations of congestion, routes of transit or emergency service vehicles, and area locations of hazardous materials impact. Another spatial data concern for ITS deployers is the variation between spatial data sets used in ITS applications. Differences in accuracy, coverage, completeness, and currency currently make unambiguous transfer of location difficult. Although some ITS applications in local areas may be satisfied by having one common data set -- for which location references may be implemented in any number of ways -- many ITS applications will have broad interoperability requirements within the nation or a region. The LRMS was developed to address these needs.

One of the most important design considerations for the LRMS, and the primary focus of testing, is supporting the transfer of information between applications employing different spatial data sets. In addition, because different applications have different location referencing needs, including point locations, links, routes, and areas, the LRMS concept acknowledges that no single specification can simultaneously meet all user requirements. Instead, the LRMS is implemented in a series of profiles, designed to meet special application needs.

The technical evaluation of the LRMS has been extensive and is ongoing. The basic concepts have been thoroughly reviewed and widely supported within public and private sectors of the ITS community through an active outreach. However, further laboratory and field testing is needed to advance the LRMS from the laboratory to deployment as an ITS standard. To be sure that the LRMS will work in the real world, it is necessary to determine how well location information is transferred between real vehicles, supporting real applications, across real communication links, degraded by real physical environments. While some operational experience with deploying LRMS profiles is already available, current deployments represent special cases that do not fully evaluate the LRMS in the expected range of situations. Therefore, a controlled test environment is needed to test variations in applications, message sets, and infrastructure elements, both existing and future developments in accordance with an experimental design in a controlled environment. Testing is being conducted at two sites. The Vehicle Intelligence & Transportation Analysis Laboratory (VITAL) located in Santa Barbara, California, and by Viggen Corporation in Knoxville, Tennessee. Field and computer tests will be conducted at both sites.

The following list summarizes the results of the Cross-streets Profile testing:

- a. Because of (1) the differences in street naming conventions, (2) the presence of street name aliases, and (3) the number of missing street names in the test data sets, using cross-streets alone (i.e., without coordinate data) is not effective at transmitting location references.
- b. By adding coordinate data to the Cross-streets Profile, higher values of profile effectiveness (up to 90%) were achieved.
- c. The effectiveness of the profile depends critically upon (1) the data sets (e.g., the profile was much more effective in the urban test areas than in the rural test areas) and (2) the matching algorithms employed (e.g., sophisticated matching algorithms resulted in significantly higher values of profile effectiveness than simple matching algorithms).
- d. The degree of effectiveness required of a profile depends upon the application. For example, an effectiveness of 90% is probably more than adequate for broadcasting traffic flow information to the public, but would not be adequate for use by emergency response personnel.

5. Overview—One of the elements required for implementation of interoperable intelligent transportation systems (ITS) is the ability to transmit location information between ITS system components. Viggen Corporation, on contract to the U.S. Oak Ridge national Laboratory (ORNL) for a project sponsored by the U.S. Federal Highway Administration (FHWA), has proposed the following seven Location Referencing Message Specification (LRMS) profiles:

- a. The Geometry Profile, which contains referencing formats based on spatial primitives
- b. The Geographic Coordinate Profile, which contains referencing formats based on latitude, longitude, and altitude
- c. The Grid Profile, which contains referencing formats intended for limited bandwidth applications
- d. The Linear Referencing Profile, which contains referencing formats based on known positions of spatial objects
- e. The Cross-streets Profile, which contains referencing formats based on intersecting streets
- f. The Address Profile, which contains referencing formats based on street addresses
- g. The Model Deployment Initiative (MDI) Profile, which contains referencing formats for use in MDI programs

The LRMS is being developed by the Society of Automotive Engineers (SAE) Map Data set Committee as SAE J2374. One of the impediments to improvement, finalization, and usage of the LRMS is an understanding of the effectiveness of each profile under real-world conditions. LRMS testing is designed to evaluate the effectiveness of these profiles by using them to transmit location references between some currently available map databases. This document describes the LRMS test plan and describes the test results to date. The first chapter is an introduction to LRMS testing. The remaining chapters, one for each LRMS profile, describe the test plan and summarize the test results for the associated profile. The appendices provide detailed test results.

Although the LRMS testing does evaluate profile effectiveness, the effectiveness in a particular application will depend upon a number of factors that cannot be considered during the LRMS testing. For example, the data sets used in a particular application may be either more or less consistent than the test data sets used in the LRMS testing. Therefore, LRMS testing is not designed to predict how a profile will perform in a particular application, but to understand the general strengths and weaknesses of each profile in order to aid readers in selecting a profile appropriate for their application.

This document is a work-in-progress and will be revised as LRMS testing proceeds. This first version of the document (Version 1.0) reports the results of testing of the Cross-streets LRMS Profile at the Vehicle Intelligence and Transportation Analysis Laboratory (VITAL) at Santa Barbara, California, and at Viggen Corporation offices at Knoxville, Tennessee. Later versions of the document will report testing of other LRMS profiles at either or both of these locations, or at additional sites.

The document is structured as follows: Test Plan for all profiles; Results of the Viggen Cross-streets profile tests and analysis; Results of the VITAL Phase I tests (without coordinates); Results of the VITAL Phase II tests (with coordinates). The Viggen tests and VITAL tests follow different methodologies and, in some cases, were run with the same input data in order to establish credibility of the methodologies used and consistency of results between the two testing locations. Having established this, remaining profiles will be tested at only one location.

5.1 Overview of Test Results—The tests for the Cross-streets Profile are complete, and the test results are described in Section 7 and Appendices A and B. In general, the differences in the following factors help create a tremendous amount of variability that can be expected when using the Cross-streets Profile to transmit location references between ITS applications based upon currently available map data sets:

- a. The geographic area. Highly urbanized areas seem to have more complete, and therefore more compatible, map data sets, resulting in higher profile effectiveness.
- b. The map data set. Some map data sets were less compatible with others due to differences of completeness, currency, detail, and spatial accuracy of coordinates. This is not necessarily a reflection of quality of a particular data set, but may reflect its being developed for a particular purpose with particular requirements. Nonetheless, map data sets produced at different levels of detail, accuracy, currency, etc., will result in lower profile effectiveness for these data sets.
- c. The sophistication of the cross-streets matching algorithm. The initial tests used simple name-, node-, and link-matching algorithms, did not use coordinates to enhance cross-streets matching, and did not correct for degenerate nodes (e.g., valence 2 nodes, end-nodes of dead-end streets). The results of the initial tests were much lower than the results of later tests using more sophisticated algorithms.
- d. The application itself. Different ITS applications may have different thresholds of acceptance for numerical Cross-streets results.
The best results (nearly 90% effectiveness) were obtained using sophisticated matching algorithms with cross-streets and coordinates for the most compatible data sets for the city of San Francisco. The worst results (around 10% effectiveness) were obtained using simpler matching algorithms without coordinates for the least compatible data sets for Santa Barbara County.
The test results also imply the following general principals of using cross-streets and coordinates for location referencing:
 - e. Using cross-streets for location referencing gives a high degree of confidence when the cross-streets match. However, using cross-streets also results in a large number of cross-street references that cannot be located, in which case the location reference fails completely.
 - f. Using coordinates gives little confidence in identifying exact locations. However, coordinates always imply a general location reference and, therefore, never fail completely in identifying a location reference.
 - g. Using cross-streets and coordinates together can result in significantly improved location references. The primary weakness of coordinates (difficulty in identifying an exact location) is offset by the strength of cross-streets (exact location referencing, when it works).

6. Introduction to the LRMS and LRMS Testing—SAE J2374 defines the following seven location referencing message specification (LRMS) profiles for transmitting location reference information:

- a. The Geometry Profile, which contains referencing formats based on spatial primitives
- b. The Geographic Coordinate Profile, which contains referencing formats based on latitude, longitude, and altitude
- c. The Grid Profile, which contains referencing formats intended for limited bandwidth applications
- d. The Linear Referencing Profile, which contains referencing formats based on known positions of spatial objects
- e. The Cross-streets Profile, which contains referencing formats based on intersecting streets
- f. The Address Profile, which contains referencing formats based on street addresses
- g. The Model Deployment Initiative (MDI) Profile, which contains referencing formats for use in MDI programs

This document is a report on (1) the tests designed to evaluate the ability of these LRMS profiles to effectively transmit location reference information within an ITS environment and (2) the results of performing these tests. This first version of the document (Version 1.0) reports the results of testing of the Cross-streets LRMS Profile at the Vehicle Intelligence and Transportation Analysis Laboratory (VITAL) at Santa Barbara, California, and at Viggen Corporation offices at Knoxville, Tennessee. Later versions of the document will report testing of other LRMS profiles at either or both of these locations, or at additional sites. The Viggen tests and VITAL tests follow different methodologies and, in some cases, were run with the same input data in order to establish credibility of the methodologies used and consistency of results between the two testing locations. Having established this, remaining profiles will be tested at only one location.

6.1 LRMS Overview—The LRMS profiles include profile formats that enable users to create and interpret data formatted according to the profile descriptions. Some profile descriptions (e.g., the Cross-streets Profile) include only the data necessary to transmit a point location reference within a linear street network. Other profiles include additional information (e.g., the Geometry Profile includes sufficient information to transmit location references for points, lines, and areas, and could be used to transmit an entire map data set). In other words, each LRMS profile was developed to address the needs of a particular group of potential users, and when these needs suggested including additional data over-and-above that required for pure location referencing, the profiles were enhanced to include this additional data. See SAE J2374 for more information on the LRMS profiles.

6.2 Testing Overview—Each of the LRMS profiles is capable of being used effectively under ideal conditions. However, the effectiveness of each profile under real-world conditions and the factors that influence this effectiveness are not known. The LRMS testing is designed to evaluate this effectiveness by using the profiles in a simulated ITS environment. For profiles whose primary purpose is transmission of location references among heterogeneous ITS applications, this ITS environment is simulated by compiling a collection of both publicly and commercially available map data sets (the test data sets), and using the profile to transmit location references between these test data sets. For profiles with additional purposes, a more advanced simulation of an ITS environment is required. For example, because one of the likely uses of the Geometry Profile is the transmission of entire map data sets between users, the testing of this profile includes a “fitness-for-use” evaluation of whether or not the profile includes sufficient detail to support this activity. A complete description of the testing for each profile is described in the chapters associated with each profile.

In addition to the primary tests that help evaluate the overall effectiveness of each profile, other tests will help (1) identify individual factors that might impact the profile effectiveness, (2) quantify this impact, and (3) evaluate methods of reducing negative impacts of these factors. For example, testing of the Cross-streets Profile identified street naming-conventions as a factor that impacted the profile effectiveness, quantified the prevalence of different naming conventions, and evaluated six different methods of reducing the impact of different naming conventions on the profile effectiveness.

After completing the design and implementation of the various test, the test results are compiled, and these results are interpreted in order to help achieve the following objectives:

- a. Estimate the over-all effectiveness of the profile.
- b. Understand the issues that impact the effectiveness of each profile in order to provide (1) guidelines for using the profiles and (2) suggestions for how changes to the profile might improve this effectiveness.
- c. Identify data issues that might impact the effectiveness of each profile in order to provide suggestions for how changes to ITS data sets might improve the effectiveness of the LRMS profiles.

6.3 Factors that Influence the Effectiveness of an LRMS Profile—In order to achieve the objectives listed in 6.2, it is necessary that the LRMS testing evaluate not only the effectiveness of the profile, but also the reasons why a profile might fail (i.e., the failure modes). Section 6.3.1 describes the failure modes associated with use of the LRMS profiles, and the LRMS testing considers and, when possible, attempts to quantify how these failure modes influence the effectiveness of the profiles.

Some of these failure modes are associated with more general data issues, which are described in 6.3.2. Because these data issues could bias the test results, the data sets were chosen to minimize the impact of most of these general issues on the testing results. The potential impacts of these issues should be considered both when interpreting the test results and when designing ITS applications that might use the profiles.

Section 6.3.3 describes some additional factors that might influence the effectiveness of using an LRMS profile. These factors are also addressed during the LRMS testing.

6.3.1 FAILURE MODES WHEN USING AN LRMS PROFILE—An understanding of the failure modes can help to (1) indicate ways to improve the profiles, (2) indicate ways to modify the data sets to enhance the effectiveness of the protocols, and (3) provide guidelines on limitations that might be expected when using an LRMS profile. Table 1 lists some of the failure modes that might occur when transmitting location reference information and lists the sections of this document that further describe each failure mode. When possible, the LRMS tests are designed to identify the impact of each failure mode on the effectiveness of each LRMS profile.

TABLE 1—FAILURE MODES WHEN TRANSMITTING LOCATION REFERENCE INFORMATION

Failure Mode	Description
Failure Modes Associated with Location Reference	
Topological discrepancies	See 6.3.1.1 and Figure 1
Topological ambiguities	See 6.3.1.2 and Figure 2
Coordinate discrepancies	See 6.3.1.3 and Figure 3
Street naming discrepancies	See 6.3.1.4 and Figure 4
Cross-streets ambiguities	See 6.3.1.5 and Figure 5
Address discrepancies	See 6.3.1.6
Failure Modes Associated with Other Uses	
Fitness-for-use deficiencies	See 6.3.1.7
User input errors	See 6.3.1.8

6.3.1.1 *Topological Discrepancies*—A topological discrepancy is a difference in the nodes and links present in two test data sets (see Figure 1). Because topological discrepancies represent differences in the location information present in the two data sets, these discrepancies limit the effectiveness of transmitting location reference information, regardless of the profile used.

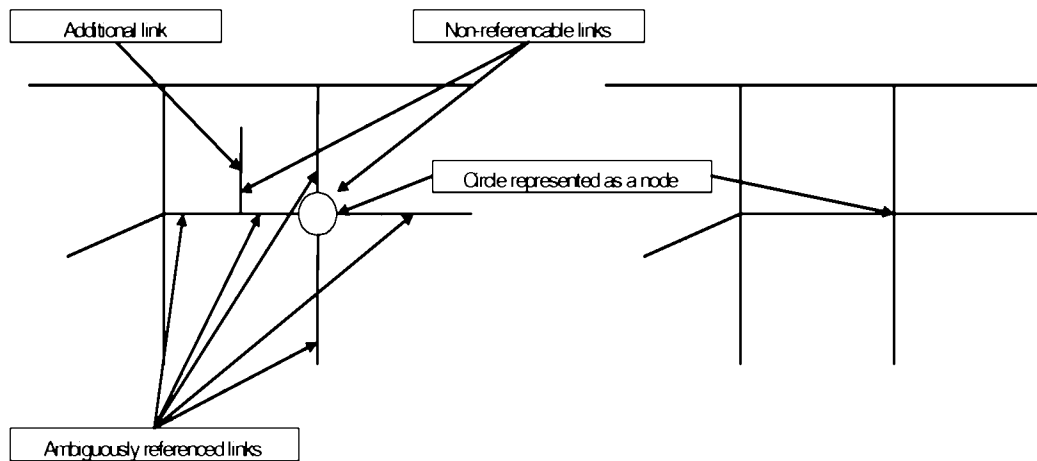


FIGURE 1—EXAMPLE TOPOLOGICAL DISCREPANCIES

To generate quantitative results on the prevalence of topological discrepancies between two test data sets requires a relationship between the nodes of the two data sets. This type of testing is not within the scope of the current LRMS testing activities. However, visual inspection of the maps associated with the test data sets verifies that topological discrepancies are one of the primary reasons for failure of an LRMS profile. Thus, the over-all failure rate for the Cross-streets Profile does provide an indirect measure of the prevalence of topological discrepancies between the test data sets and, hence, between likely ITS system components.

Although topological discrepancies do result in a considerable loss of effectiveness in the test data sets, it is expected that the impact of these discrepancies will lessen as ITS data sets mature. In near term, ITS applications will probably be limited to primary roadways where topological discrepancies are expected to occur less frequently.

- 6.3.1.2 *Topological Ambiguities*—Topological ambiguities are real-world street topologies (see Figure 2) that prevent unambiguous link references in terms of the from- and to-nodes. In some cases (e.g., a loop), the from- and to-nodes indicate the correct link, but the correct direction along the link (e.g., counter-clockwise or clockwise) is ambiguous. In other cases (e.g., a figure eight), specifying the from- and to-nodes does not unambiguously identify the correct link. Topological ambiguities are a real-world fact and limit profile effectiveness.

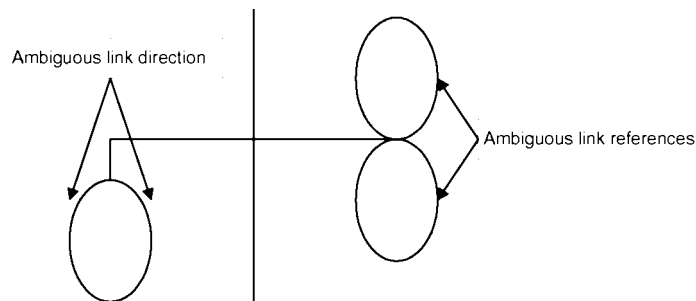


FIGURE 2—EXAMPLE TOPOLOGICAL AMBIGUITIES

The frequency of occurrence of topological ambiguous links in the test data sets listed in Table 2. These results indicate that topological ambiguities are not a significant concern when transmitting location references.

**TABLE 2—OCCURRENCE FREQUENCY OF TOPOLOGICAL AMBIGUITIES
(LINKS WITH TOPOLOGICAL AMBIGUITIES)**

Data Set	Frequency	Data Set	Frequency	Data Set	Frequency
SF-A	0.10%	SB-A	0.17%	SBC-A	0.96%
SF-B	0.00%	SB-B	0.00%	SBC-B	0.00%
SF-D	0.21%	SB-D	0.00%	SBC-D	0.00%
SF-F	0.05%	SB-E	0.39%	SBC-E	0.39%
		SB-F	0.15%	SBC-F	0.90%

The data sets listed in Table 2 are described in 6.4.

- 6.3.1.3 *Coordinate Discrepancies*—Coordinate discrepancies are differences in the coordinate data for map features in different ITS system components (see Figure 3). These discrepancies limit the effectiveness of LRMS profiles that rely on coordinates to transmit location references.

To directly test for coordinate discrepancies between test data sets requires a relationship between the nodes two of test data sets so that the distance between the matching nodes can be computed. This type of testing is not within the scope of LRMS testing. However, the LRMS testing does provide some limited evaluations of the coordinate discrepancies present in the test data sets by evaluating the distance between (1) “nearest nodes” and (2) nodes identified as matching nodes using one of the LRMS profiles (see 7.5.4).

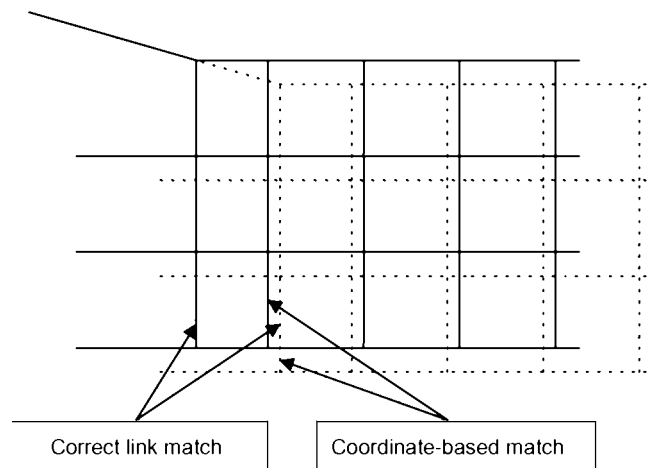


FIGURE 3—EXAMPLE COORDINATE DISCREPANCIES

These discrepancies should lessen as ITS data sets mature and evolving technology facilitates creation of more accurate map data sets.

- 6.3.1.4 *Street Naming Discrepancies*—Street naming discrepancies are differences in the street names for a specific link in two ITS system components (see Figure 4). Three basic types of street naming discrepancies exist: (1) missing street names, (2) different representations of the same street names (e.g., “Main St” and “Main Street”), and (3) streets with multiple names. These discrepancies limit the effectiveness of using street names to transmit location reference information. The LRMS testing directly evaluates the various types of street naming discrepancies and their impact on the effectiveness of LRMS profiles.

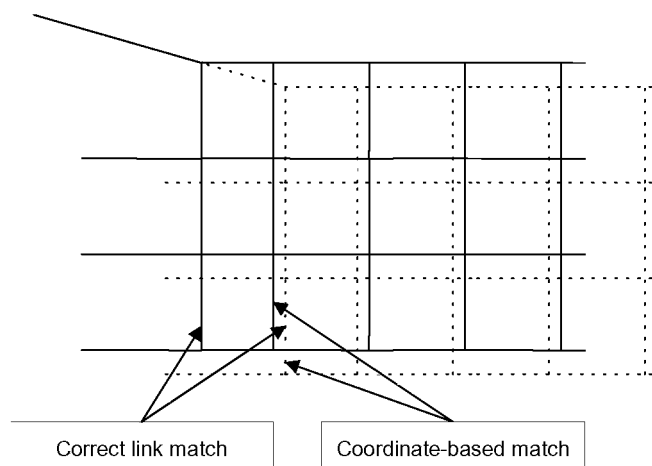


FIGURE 4—EXAMPLE STREET NAMING DISCREPANCIES

It is expected that the problem of missing street names will lessen as ITS data sets mature. The problem of street name representations will persist until street naming standards are established and followed. Designing ITS data sets to accommodate street name aliases can circumvent difficulties with multiple street names. See Section 7 for a more complete description of street naming discrepancies.

6.3.1.5 Cross-Streets Ambiguities—Cross-streets ambiguities (see Figure 5) are real-world street topologies such as dead-end links and multiple intersections between streets that limit the effectiveness of using cross-streets to transmit location reference information by preventing the formation of a valid cross-street reference that unambiguously identifies a link. This ambiguity can sometimes be removed by enhancing the street name data with other information (e.g., coordinate data, extending the definition of cross-streets to describe such links). The LRMS testing will estimate the prevalence of this type of ambiguity.

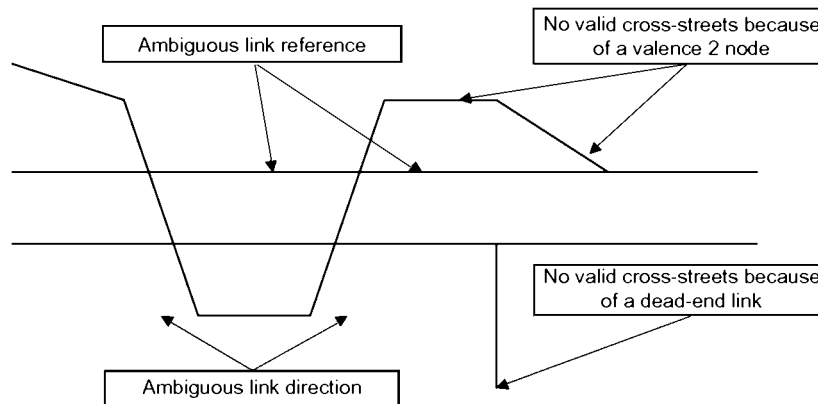


FIGURE 5—EXAMPLE CROSS-STREETS AMBIGUITIES

The frequencies of occurrence of cross-streets ambiguities are listed in Table 3. The columns labeled “Dead End” and “Valence 2” refer to dead-end links and links with a valence 2 end-node that have the same street name for both intersecting links. The columns labeled “Ambiguous Link” refer to links for which another link in the data set has the same cross-streets reference, either without blank street names in the reference (“No Blanks”) or with at least one blank street name (“Blanks”). The columns labeled “Ambiguous Direction” refer to links for which the from- and to-street are the same, with the “No Blanks” and “Blanks” values as described previously.

The results in this table indicates that ambiguous links and link directions because of cross-streets ambiguities is not a significant concern, and many of the situations where ambiguous references occur will be resolved as the number of blank street names is reduced. However, a significant percentage of the links in the test data sets are either dead-end links or links with a valence 2 end-node.

TABLE 3—OCCURRENCE FREQUENCY OF CROSS-STREETS AMBIGUITIES

Data Set	No Valid Profile Dead End	No Valid Profile Valence 2	Ambiguous Link No Blanks	Ambiguous link With Blanks	Ambiguous Direction No Blanks	Ambiguous Direction With Blanks
SF-A	6.1%	15.8%	1.7%	2.4%	1.7%	0.0%
SF-B	5.4%	12.0%	1.4%	0.2%	1.0%	0.1%
SF-C	6.6%	5.8%	1.7%	0.6%	2.3%	0.6%
SF-D	5.7%	8.2%	0.7%	1.3%	1.0%	1.3%
SB-A	14.0%	30.9%	0.6%	4.4%	1.5%	0.2%
SB-B	15.1%	23.0%	2.0%	0.4%	1.5%	0.1%
SB-D	16.2%	8.7%	2.3%	3.3%	2.4%	0.1%
SB-E	17.0%	5.5%	1.0%	3.9%	2.3%	0.0%
SB-F	15.8%	18.4%	0.9%	1.4%	2.1%	0.1%
SBC-A	15.4%	127.6%	0.3%	6.6%	0.7%	0.1%
SBC-B	12.5%	38.3%	1.7%	2.8%	0.7%	0.1%
SBC-D	16.4%	11.2%	2.3%	4.5%	2.2%	0.1%
SBC-E	15.7%	11.0%	0.8%	4.7%	1.7%	0.1%
SBC-F	15.7%	26.3%	0.4%	5.4%	0.8%	0.1%

See Section 7 for more information on cross-streets ambiguities, including recommendations for eliminating the ambiguities associated with dead-end links and links with a valence 2 end-node.

- 6.3.1.6 *Address Discrepancies*—Address discrepancies are differences in the address data in different ITS system components (e.g., missing address information on one ITS system component).
- 6.3.1.7 *Fitness-for-Use Deficiencies*—Fitness-for-use deficiencies refer to the inability of a profile to meet the needs of a potential user group. For those profiles whose primary usage is location referencing, the fitness-for-use deficiencies are addressed by the specific location referencing failure modes described in the previous sections. For other profiles, the test plan lists specific “other uses” of the profile and describes the methods used to evaluate whether or not the profile is sufficient for these other uses. Because these deficiencies are related to specific profiles, additional descriptions of fitness-for-use deficiencies are included in the individual chapters associated with each profile.
- 6.3.1.8 *User Input Errors*—User input errors refer to errors in the location reference or other information supplied by a user of an ITS system component. Although user input of a location reference is a location reference transmission that could be covered under an LRMS profile, the types of errors likely to occur are different than those evaluated by the LRMS testing. For example, a user-input error might be an imprecise GPS position rather than a difference between coordinates in different map data sets. Testing of user input errors are not within the scope of the LRMS testing.
- 6.3.2 DATA ISSUES THAT INFLUENCE PROFILE EFFECTIVENESS—The effectiveness of an LRMS profile to transmit a location reference depends strongly on the data sets associated with the two ITS system components that are using the profile. Highly effective transfer rates cannot result if the two data sets differ in any of the four factors listed in Table 4.

TABLE 4—DATA ISSUES THAT IMPACT LRMS PROFILE EFFECTIVENESS

Data Issue	Effect on LRMS Profile Effectiveness
Differences in coverage	See 6.3.2.1 and Figure 6
Differences in level of detail	See 6.3.2.2 and Figure 7
Differences in level of completeness	See 6.3.2.3
Differences in currency	See 6.3.2.4

In non-interoperable systems, the system designers can select compatible data sets in order to control the effect these issues might have on transmitting location references. However, in fully interoperable systems, the system designers cannot control the data sets that might interact with a system and, therefore, cannot fully control the effect these issues might have on transmitting location references. The following sections describe these data issues, relate these issues to the failure modes discussed in 6.3.1, and describe the effect these issues might have on transmitting location references.

- 6.3.2.1 *Differences in Coverage*—Coverage differences refer to the differences in the areas covered by two map data sets (see Figure 6). If two data sets have different coverage areas, then transfer of location references for those areas not present in both data sets is not possible.

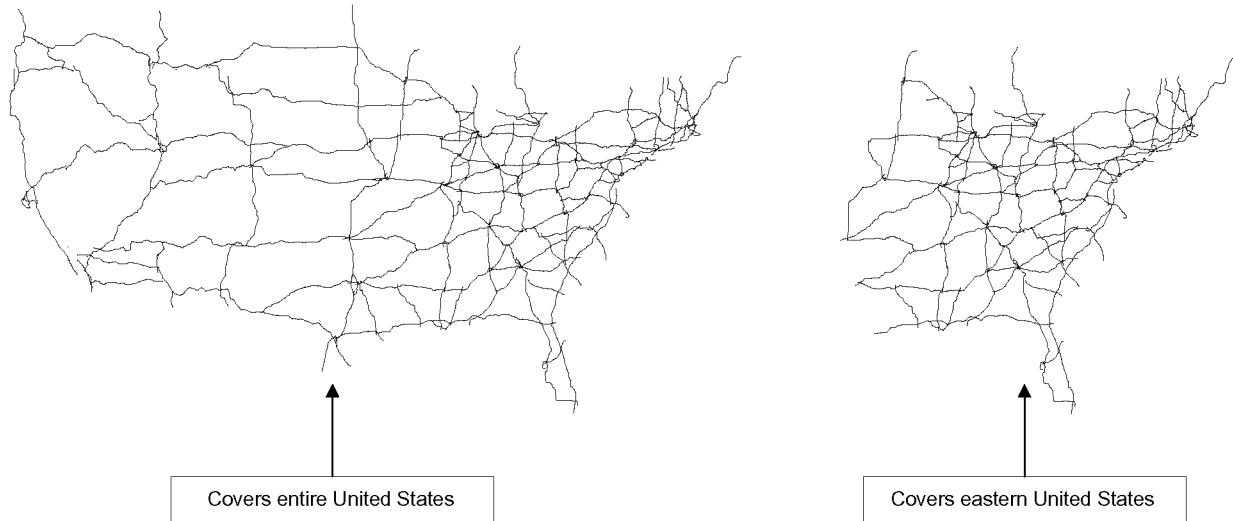


FIGURE 6—EXAMPLE COVERAGE DIFFERENCES

Because transmission of location references between ITS system components cannot occur if the components have different coverage areas, the potential effect of coverage differences is not directly considered in the LRMS testing. In fact, the LRMS testing reduces the potential impact of coverage differences by using data sets with the same general coverage (e.g., the city of San Francisco, Santa Barbara County).

6.3.2.2 Differences in Level of Detail—Two types of difference in the level of detail between two map data sets exist: (a) differences in the types of roadways included in a map data set, and (b) differences in the detail with which a roadway is represented. Figure 7 exemplifies these differences. For example, one map data set may include only interstate highways while another also includes secondary roads. As another example, one map data set may treat all roadways as single links and all intersections as single nodes, while another uses separate links for each side of a divided highway and represents each entrance and exit ramp as a separate link. If two data sets have differences as described in (a) above, then some locations included in one data set are purposefully omitted from the other data set, and transfer of these location references is not possible. If two data sets have differences as described in (b), then transfer of detailed location information from the more detailed data set is not possible.

Because transmission of detailed location references between ITS system components cannot occur if the receiving component does not include detailed data, the potential effect of differences in the level of detail is not directly considered in the LRMS testing. In fact, the LRMS testing reduces the potential impact of these differences by using test data sets with the same general level of detail.



FIGURE 7—EXAMPLE OF DIFFERENCES IN LEVEL OF DETAIL

6.3.2.3 Differences in the Level of Completeness—The level of completeness of a map data set refers to the extent of errors of omission in the data. If two data sets have different levels of completeness, then some locations will be mistakenly omitted from the data sets, and transfer of these location references is not possible. Differences in completeness will usually appear as topological discrepancies (see 6.3.1.1 for more information), and the impact of topological discrepancies is considered by the LRMS testing. Because currently available commercial databases were used as test data sets in the LRMS testing, the LRMS testing does reflect the differences in level of completeness that is likely to be present in any soon-to-be-developed ITS system components.

One crude measure of the relative level of completeness in the test data sets is the number of links in each data set. (Because different vendors may represent street networks differently, the number of links in a data set is not an exact measure of this data issue.) Table 5 lists the number of links in each data set used in the LRMS testing.

TABLE 5—DIFFERENCES IN LEVEL OF COMPLETENESS

Data Set	Number of Links ⁽¹⁾	Data Set	Number of Links ⁽¹⁾	Data Set	Number of Links ⁽¹⁾
SF-A	15411	SB-A	7043	SBC-A	30142
SF-B	15072	SB-B	5772	SBC-B	26409
SF-D	15740	SB-D	6083	SBC-D	8082
SF-F	14830	SB-E	5969	SBC-E	18963
		SB-F	5620	SBC-F	25649

1. Combines adjacent links connected by a valence 2 node into a single link.

The results in this table indicate that the test data sets for San Francisco have similar levels of completeness, the data sets for Santa Barbara County have very different levels of completeness. In part because of this difference, some of the tests for Santa Barbara were performed after restricting the data to those areas near the city center. The column labeled "City Only" lists the number of links in this subset of the data, and indicates that the level of completeness is much more similar when restricted to this subset.

6.3.2.4 Differences in Currency—Differences in currency between two map data sets refer to how recently each data set has been updated to reflect changes in the real-world data. For example, if two data sets have been last updated at different times, then any changes to the street topology that occurred between the dates the data sets were updated will result in differences between the data sets. Transfer of location references affected by these changes may not be possible. Differences in currency will usually appear as topological discrepancies (see 6.3.1.1), and the impact of topological discrepancies is considered by the LRMS testing. Because currently available commercial databases were used as test data sets in the LRMS testing, the LRMS testing does reflect the differences in currency that is likely to be present in soon-to-be-developed ITS system components.

6.3.3 OTHER FACTORS THAT INFLUENCE PROFILE EFFECTIVENESS—From the perspective of the potential user, some additional factors exist that might impact the effectiveness of an LRMS profile. Table 6 lists some of these factors that are considered during the LRMS profile testing.

TABLE 6—OTHER FACTORS THAT INFLUENCE PROFILE EFFECTIVENESS

Factor	Description
Types of streets (e.g., primary roadways, secondary roadways)	Some near-terms ITS applications will likely be limited to primary roadways, and the effectiveness of LRMS profiles may be different for primary roadways than for other types of streets. The LRMS testing evaluates some of the differences in effectiveness between primary and other roadways.
Type of geographical area (e.g., urban area, rural area)	Differences in LRMS profile effectiveness may exist between geographical areas. For example, an urban area may have more consistent data because of a higher concentration of ITS activities, but could have more currency problems (because of more frequent road building) and topological problems (because of a more dense road network resulting in more data errors). The LRMS testing includes data from both a highly urbanized area and a less urbanized area, and differences between the test results for the two areas provide some indication of the differences in profile effectiveness that might be expected in different geographical areas.

6.4 Test Data—In order to evaluate the range of results that might occur in using the LRMS profiles, data from several commercial data sets for two geographical areas are used for the LRMS profile testing. This data was used to (1) generate test points for transmitting location reference information and (2) serve as an ITS system component receiving location reference information. In addition, field data and incident data were collected to provide additional test points. Table 7 lists the code names for the data that was used for LRMS profile testing.

TABLE 7—DATA SET CODE NAMES

Test Site	Data Set Code Names
City of San Francisco	SF-A, SF-B, SF-D, and SF-F
Santa Barbara City ⁽¹⁾	SB-A, SB-B, SB-D, SB-E, and SB-F
Santa Barbara County	SBC-A, SBC-B, SBC-D, SBC-E, and SBC-F

1. The Santa Barbara City data was extracted from the Santa Barbara County data by retaining all links within a rectangular grid centered over Santa Barbara. See Figure A3.

6.4.1 DATA NORMALIZATION—All of the data for the LRMS testing were translated into a common data format for testing. Because normalizing the data could bias the test results, other data normalization steps were avoided for most of the LRMS tests. However, the following tests did require additional normalization steps:

In order to compare two sets of coordinates, both sets of coordinates must be expressed relative to a standard coordinate system. The coordinates were normalized to NAD83 for LRMS testing.

Because of the variability in street naming conventions, effective street name matching requires normalized street names. The LRMS testing considers a variety of possible techniques for normalizing street names

(see Section 7 for more information).

Evaluating the portion of location reference transmissions that fail because of topological discrepancies between data sets requires topological normalization (i.e., identifying the relationship between the nodes and links in one data set with those of another). Topological normalization is not within the scope of the current LRMS testing activities. However, visual inspection of the maps associated with the test data sets verifies that topological discrepancies are a significant cause of failure when transmitting location references.

- 6.4.2 DATA SAMPLING AND WEIGHTING—Three types of test point sources (i.e., commercial data sets, field data, and incident data) are used to generate test points for the LRMS testing. Table 8 describes the advantages and disadvantages of using each of these sources to generate test points.

TABLE 8—TYPES OF TEST POINTS FOR LRMS PROFILE TESTING

Test Point Source	Description
Commercial data sets	Data from one commercial data set was used to generate test points to transmit to a second data set. Advantages: Inexpensively generates a large number of test points. Disadvantages: The distribution of test points generated is not likely to agree with the points associated with a real-world application, making the results difficult to interpret. Comments: To mitigate the disadvantages, results based on different samples and weightings are reported (see Table 9).
Field data	Field data was gathered to use as test points. Advantages: Data gathered in a manner more consistent with how data would be gathered in a real-world application. Disadvantages: Test points are expensive to generate.
Incident data	Incident data from police and emergency medical systems was used to generate test points to transmit to a data set. Advantages: Data reflects the distribution of data points from a real-world emergency management system. Disadvantages: Limited number of test points. Comments: Results may not be appropriate for uses other than emergency response.

Because of the potential difficulties when interpreting the results obtained using test points generated from a commercial data set, several sampling and weighting techniques are used in conjunction with these data sets. Table 9 describes these techniques and how the results of these techniques might be used.

TABLE 9—POSSIBLE SAMPLING AND WEIGHTING TECHNIQUES FOR COMMERCIAL DATA SET TEST POINTS

Technique	Description
All nodes or links	All nodes or links are used as test points. This sample is appropriate if an application will involve all types of roadways, and is the primary sampling technique used in the testing.
Primary roadway sample	All nodes or links that are part of a primary roadway are used as test points. This sample is appropriate if an application will be limited to primary roadways. This sample is also used to evaluate whether LRMS profile effectiveness varies with the roadway type.
Secondary roadway sample	All nodes or links that are part of a secondary roadway are used as test points. This sample is used to evaluate whether LRMS profile effectiveness varies with the roadway type.
Local roadway sample	All nodes or links that are part of a local roadway are used as test points. This sample is used to evaluate whether LRMS profile effectiveness varies with the roadway type.
Special features sample	All nodes or links that are part of a special roadway feature are used as test points. This sample is used to evaluate whether LRMS profile effectiveness varies with the roadway type.
Regional samples	All nodes or links that are part of a region of the coverage area are used as test points. This sample is used to evaluate whether the LRMS profile effectiveness varies within regions in the coverage area (e.g., the effectiveness is different near the city center than on the city periphery).
Weight equally	All nodes or links are used as test points, and each node is equally weighted. This weighting is appropriate if transmission of each link is equally likely, and is the primary weighting used in the profile testing.
Weight by link length	Weight each link by the link length. Within a roadway category, weighting each link by link length may more accurately reflect the distribution of required location reference transmission in a real-world system.

The "link class" field that was available in each test data set was used to define the roadway classifications. The link classes associated principally with each roadway category were collected into the sample for that roadway category. Regional samples were generated by considering only those objects that occur in a geographical "rectangle" bounded by specific latitude and longitudes. The straight-line distance between from- and to-nodes for the links were used to weight by link length. The following observations will help when interpreting the results using the different sampling and weighting methods:

- a. A single street may have different roadway classifications in different test data sets. Because no standard exists for classifying roadways, the roadway classifications differ between the data sets. Therefore, the results sampled by roadway classification should be compared only to other results for the same data set in order to identify difference in effectiveness by roadway classification.
- b. The link classifications differed considerably between test data sets. For example, one test data set might classify a ramp as part of a primary roadway, and another data set would classify a ramp as a special feature.
- c. For each data set, each link is included in a single roadway class. However, a single street may be in several roadway classes, if different links that comprise the street are in different classes. A single node is probably in several roadway classes because the links that intersect that node are likely in different roadway classes.
- d. The regional samples are drawn very broadly and are intended primarily to (1) eliminate differences in coverage (see 6.3.2.1 and (2) contrast the effectiveness for roads near to the city center with those at the periphery where private roads and other anomalies might occur.

6.5 LRMS Profile Testing—Testing of each LRMS profile occurs more or less independently, and a separate chapter of this document is devoted to the testing of each profile. Because this testing is exploratory in nature, testing of a profile is performed in a series of phases with each phase refining the results of the previous tests. The testing begins with a conceptual review of the profile and its uses, followed by preliminary testing. The test results are reviewed, and further testing is performed to refine these results, as necessary. The process of test, review, and refine is repeated until the objectives of the testing are met.

For example, the preliminary tests might address the types of data used in the profile (e.g., look at the frequency with which blank street names occur). The first phase might evaluate the overall effectiveness of the LRMS profile in transmitting link information and, when applicable, explore the data set discrepancies that cause low transmission rates. Additional phases might (1) explore the improve

.ments obtained by revising the profile and (2) use sampling, weighting, and normalization techniques to further evaluate the causes of low transmission rates.

The specifics of each testing phase for an LRMS profile is described by a separate test plan and summary of results included in the chapter for that profile. Detailed tables generated by the testing are provided in appendices to each chapter. The following sections describe the types of information that is included in the test plans.

6.5.1 WAYS OF USING AN LRMS PROFILE—The LRMS profiles are designed to transmit location reference information between components of an ITS system, and each profile is designed to transmit this information in a different way. In some cases, the profile data can be used in different ways to identify a location (e.g., for the Cross-streets Profile, (1) use cross-streets to identify links and coordinates to resolve ambiguities or (2) use coordinates to identify links and cross-streets to resolve ambiguities). In other cases, the profile data can be used to transmit information other than point locations. Each test plan describes the different ways the profile might be used, and identifies those ways that are considered during the profile testing.

- 6.5.2 **FAILURE MODES ASSOCIATED WITH USING AN LRMS PROFILE**—Only some of the failure modes described in 6.3.1 apply to the use of each profile to transmit location reference information, and other failure modes apply to other uses of the profile. Each chapter includes a section describing those failure modes that apply to the given profile. One of the objectives of the testing is to estimate the impact of each of these failure modes on the overall effectiveness of the profile and to provide suggestions on how to reduce the impact of each failure mode and increase the profile effectiveness.
- 6.5.3 **LRMS PROFILE TEST PROTOCOL**—Each chapter includes a general test protocol that applies to most of the testing for that chapter. The general test protocol will describe how, for most of the testing, (1) the profile is used, (2) the sample is selected, and (3) the results are weighted. When the profile is used differently, or a different sampling or weighting technique is used for specific tests, those differences are included in the description of the test. Otherwise, the test description refers back to this general protocol.
- 6.5.4 **LRMS PROFILE TESTS**—Specific tests must be designed to help evaluate the profile effectiveness and the impact of each failure mode on this effectiveness. A series of sections describe these tests and their results, one section for each testing phase. Detailed results for each testing phase are contained in an appendix associated with each chapter.
- 6.5.5 **LRMS PROFILE TEST RESULTS SUMMARY**—Each chapter concludes with a section that summarizes the test results for that chapter and lists any recommendations for improving the profile effectiveness. These recommendations may be in the form of suggested profile modifications or enhancements, suggestions for future testing, or ITS system implementation suggestions.
7. **The Cross-streets Profile**—The Cross-streets Profile is designed to use coordinates and street names as the primary means of transmitting location reference information. Because the Cross-streets Profile was identified as a likely candidate for early implementation, this profile was the first of the LRMS profiles tested.
- 7.1 **Description of the Cross-streets Profile**—The cross-streets profile was originally designed to use street names to identify streets, street intersections to identify nodes in a linear referencing model (LRM), and uses two nodes (identified by street intersections) to identify a link. The original profile also included offset data to specify a position on a link, and side data to specify a side of a link. Coordinate data was added to the Cross-streets Profile after Phase I testing in order to help resolve street name ambiguities. Figure 8 provides a graphical depiction of the Cross-streets Profile.

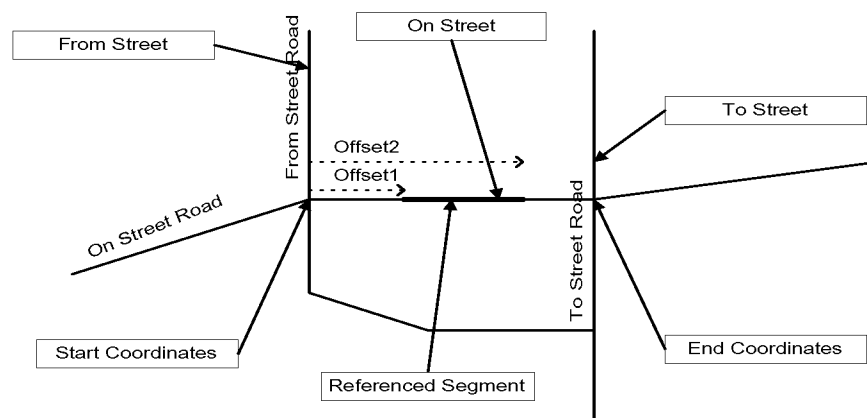


FIGURE 8—THE CROSS STREETS PROFILE

7.2 Using the Cross-Streets Profile—The cross-streets profile is designed primarily to use the following methodology to identify locations:

1. Use coordinates to identify links in a linear referencing network
2. Use cross-streets to resolve ambiguities in the links determined by coordinates
3. Use offsets to identify a segment (of the link)
4. Use side data to identify a side of the segment

However, **a specific methodology for translating profile data into a location reference is not part of the Cross-streets Profile**, and methodologies other than that describe above can be used. Table 10 lists several methodologies that could be used for identifying a link using the data in the cross-streets profile. After determining a link, the offset and side data are used to identify the location segment and side. Although other uses of the Cross-streets Profile are possible, the profile testing concentrates on uses consistent with the primary usage describe above. However, the tests are designed so that many intermediate test results could be applied to other uses of this profile.

TABLE 10—SOME METHODOLOGIES FOR USING THE CROSS-STREETS PROFILE TO IDENTIFY A LINK

Methodology	Description
Cross-streets	Use cross-streets to identify a link
Cross-streets, clarified with coordinates	Use cross-streets to identify a link, then use coordinates to clarify any ambiguities
Cross-streets, selected by coordinates	Select a set of possible links based upon coordinates, then choose the link from that set of cross-streets to best match those in the profile
Coordinates	Use coordinates to identify a link
Coordinates, clarified with cross-streets	Use coordinates to identify a link, then use cross-streets to clarify any ambiguities
Statistical	Rate each link according to the goodness of fit of both the cross-streets and coordinate matching, retain all links with ratings above some minimal setting, and select the link with the highest rating
Statistical, selected by cross-streets	Select a set of links based on cross-streets, rate each link according to the goodness of fit of both the cross-streets and coordinate matching, retain all links with ratings above some minimal setting, and select the link with the highest rating
Statistical, selected by coordinates	Select a set of links based on coordinates, rate each link according to the goodness of fit of both the cross-streets and coordinate matching, retain all links with ratings above some minimal setting, and select the link with the highest rating

In addition to the different methodologies listed in Table 10, different techniques can be used within each methodology. For example, street names can be matched on an “as provided” basis or can be normalized to correct for different naming conventions in the various data sets before matching. Table 11 lists several different techniques that might be used within each methodology.

TABLE 11—SOME TECHNIQUES FOR MODIFYING LINK MATCHING METHODOLOGIES

Technique	Description
Street name normalization	Perform one or more of the following activities: (1) remove prefixes and suffixes (e.g., “N. Main” becomes “Main”, “Birch St” becomes “Birch”), (2) modify numeric names (e.g., “3 rd ” becomes “3”, “Third” becomes “3”), and (3) modify highway names (e.g., “I-24” becomes “Interstate 24”)
Street name comparison	Allow matching of “similar” street names (e.g., “Andrew Jackson Highway” matches “Jackson Highway”, “Barch” matches “Birch”)
Coordinates snap to nodes	Match coordinates to the nearest node and identify a link by the two “matched nodes”
Coordinates snap to links	Match coordinates to the nearest link and identify a link if both coordinates “match” the same link
Coordinate auto-correction	Assume most of the coordinate error is auto-correlated, and determine a single correction vector that will match both coordinates with nearby nodes or links.
Statistical coordinate selection	Use estimates for the coordinate and mapping errors and consider all nodes within the 95% confidence interval for the coordinate. Select nodes from those considered that form a link.

The Cross-streets profile may also be used to identify information other than a position on a linear reference model. For example, a from-street and an on-street can be used to specify a node, or a coordinate can be used to specify a position either on or off the linear referencing network. Table 12 lists some alternate uses of the Cross-streets Profile.

TABLE 12—ALTERNATE USES OF THE CROSS-STREETS PROFILE

Usage	Description
Identify nodes	Any of the methods described in Table 10 can be modified to identify a node rather than a position. For example, transmission of a from-street and an on-street can specify a node.
Identify links	Any of the methods described in Table 10 can be modified to identify a link rather than a position on the link.
Identify off-network points using coordinates	Coordinates can identify points not on the network, either with or without cross-streets to help verify the location
Identify off-network points using offsets	Cross-streets, coordinates, and Offset 1 can be used to identify a network point with Side and Offset2 identifying an off-network point (i.e., 50 m off the left side of the road)

Because of the large number of ways in which these methodologies, techniques, and uses could be used, it is not feasible to test all possible combinations. Instead, a typical usage is defined in 7.4, and results are presented for this typical usage. In order to facilitate the use of the testing results by a wide audience, independent testing is performed for many of the steps that make up the typical usage. If the reader has an application that only involves a few of the steps in the Cross-streets Profile typical usage, then the reader can use those part of the test results that are pertinent to their application.

7.3 Failure Modes When Using the Cross-streets Profile—Understanding the ways in which the cross-streets profile can fail to correctly identify a location reference will help in both testing and improving the cross-streets profile. In general, a Cross-streets Profile location reference can fail in one of two ways: (1) the data can fail to identify a link in the receiving ITS system component or (2) the data can fail to correctly identify the correct position on the specified link. Table 13 lists some of the ways the profile may fail to identify a link, and Table 14 lists some of the ways the profile may fail to correctly identify a position on a specified link. The Cross-streets Profile testing will evaluate both of these types of failure.

TABLE 13—FAILURE MODES FOR IDENTIFYING LINKS USING THE CROSS-STREETS PROFILES

Failure Mode	Example
	Topological Discrepancies
Different nodes	The link cannot be identified because one or both nodes defining the link is missing in the receiving test data set
Different links	The link cannot be identified because the link is either missing or equivalent to concatenated links in the receiving test data set
	Topological Ambiguities
Self-intersections	The link cannot be specified because it contains self-intersections
	Geographical Discrepancies
Imprecise node coordinates	Coordinates specify the wrong node because the coordinates for a node in the transmitting and receiving test data sets do not match
	Street Naming Discrepancies
Blank names	The link cannot be identified because street names are missing in either the transmitting or receiving test data set
Differing names (aliases)	The link cannot be identified because the different street names are used in the transmitting and receiving test data set
Differing names (spelling)	The link cannot be identified because the street name spelling differs in the transmitting and receiving test data set
Differing names (prefixes, types, and suffixes)	The link cannot be identified because the prefixes and suffixes differ in the transmitting and receiving test data set
Duplicate names	Either the link cannot be identified or an incorrect link is identified because duplicate street names occur in the receiving test data set

TABLE 13—FAILURE MODES FOR IDENTIFYING LINKS USING THE CROSS-STREETS PROFILES (CONTINUED)

Failure Mode	Example
	Street Naming Ambiguities
Multiple intersections	The link cannot be identified because multiple intersections between roads creates ambiguity in the cross-streets reference
Differing cross-streets	The link cannot be identified because the transmitted cross-streets specify the endpoints of separate links in the receiving test data set (i.e., an intervening cross-street exists in the receiving test data set)
	Input Discrepancies
Coordinate errors	GPS coordinate errors cause imprecise location in the transmitting test data set
Improper cross-street specification	The cross-streets are incorrectly identified in the transmitting test data set

TABLE 14—FAILURE MODES USING OFFSETS IN THE CROSS-STREETS PROFILE

Failure Mode	Example
	Geographical Discrepancies
Link lengths	An event is improperly located on a link because the link lengths differ in the transmitting and receiving test data set s
Shape points	An event is improperly located on a link because the shape points differ in the transmitting and receiving test data set s

7.4 The Cross-streets Profile Testing Protocol—Defining the Cross-streets Profile testing protocol consists of describing the following factors:

- How the profile was used during the testing
- How sample points were generated for testing
- How results were weighted when producing test results

The following topics describe these factors.

7.4.1 USING THE CROSS-STREETS PROFILE DURING TESTING—The typical usage of the Cross-streets Profile for the LRMS tests is the transmission of a position on a link between two different map data sets. This process could represent the situation where an Information Service Provider (ISP) is transmitting location reference information based on a proprietary map data set to a user application based on a different map data set. The individual steps involved in this usage are described below, and are illustrated in Figure 9.

- Steps A through F**—Select a location on a linear network and create the Cross-streets Profile that describes this location. For most of the testing, these locations were identified by selecting a sample of locations from a map data set, and the Cross-streets Profile data was assembled based on the information in that data set. These test locations were supplemented by a small number of test locations generated in the field. Because steps A through F concern creation of test points rather than use of the Cross-streets Profile, there are few test results concerning these steps. However, the Cross-streets Profile cannot represent some locations (e.g., locations on dead-end streets because the end-node does not have cross-streets), and some results are presented concerning locations which cannot be represented by the Cross-streets Profile.
- Steps G through I**—Assemble, transmit, and parse the Cross-streets Profile information. For the LRMS testing, the Cross-streets Profile information was transmitted to routines that interpreted this information (steps J through O). However, this information was not typically transformed into the bit-level Cross-streets Profile data structure as described in SAE J2374. Instead, different data structures with the same information content were used.

- c. **Step J**—Use the coordinates to identify a location area (i.e., all points within a specified radius of the coordinates specified in the profile) in which the resulting location should lie. Some preliminary tests were performed to help (1) understand the discrepancies in the coordinate data between the map data sets used in the testing and (2) choose the radius that defines the location area. When coordinate data was used, this data helped to (1) expedite the testing by reducing the total number of nodes that were checked and (2) reduce the occurrence of false cross-street matches by requiring that the match have coordinates near to those in the cross-streets reference. For some tests, coordinate data was not used, and this step was omitted. In addition, a limited number of tests were performed using coordinate data as the primary means of identifying locations.
- d. **Step K**—Use the street names to identify streets. Some preliminary tests were performed to help understand the discrepancies in the street name data between the map data sets used in the testing. Because a large number of techniques are possible for matching street names (e.g., omit/maintain street types), the results of these preliminary tests were used to identify a simple, effective technique that could be used for the remainder of the testing. This technique, described in 7.5.3, was used for much of the remainder of the Cross-streets Profile testing.
- e. **Step L**—Consider the nodes within the location area identified in Step J, and use the cross-streets to identify the from- and to-node. Because several techniques are available for using cross-streets to identify nodes (e.g., how to interpret blank street names), some preliminary tests were performed to identify a simple technique that could be used for the remainder of the testing. This technique, described in 7.5.2, was used throughout the remainder of the Cross-streets Profile testing. Because the Cross-streets Profile may be used to transmit node information, independent results are reported on using the profile for nodes.
- f. **Step M**—Use the coordinate data to resolve any ambiguities. In some cases, a Cross-streets Profile will not uniquely identify an object in the receiving test data set. Some preliminary testing was performed to identify the frequency with which such cross-street ambiguities occur. When coordinate data is included in the Cross-streets Profile, most ambiguities are resolved by restricting attention to those nodes contained within the location area identified in Step J. If cross-street ambiguities exist within this location area, then the object whose coordinates most closely match those in the profile is selected.
- g. **Step N**—Use the from- and to-nodes to identify a link. The primary purpose of the Cross-streets Profile testing is to report on the effectiveness of using the Cross-streets Profile to correctly transmit a link reference. The effectiveness of the profile for transmitting link information is reported in the test results. In some cases, from- and to-nodes from one map data set will not identify a link in another map data set (e.g., the matching “link” contains an intervening node). Because some users of the Cross-streets Profile may not be able to resolve such “degenerate” links, results are reported on the frequency with which degenerate links are found during cross-streets testing.
- h. **Step O**—Use the offsets to identify a location on the link. The final step of using the Cross-streets Profile is using the offset to identify a position on the link. However, this step is independent of the process of identifying the correct link, and separate tests results are reported for this step.

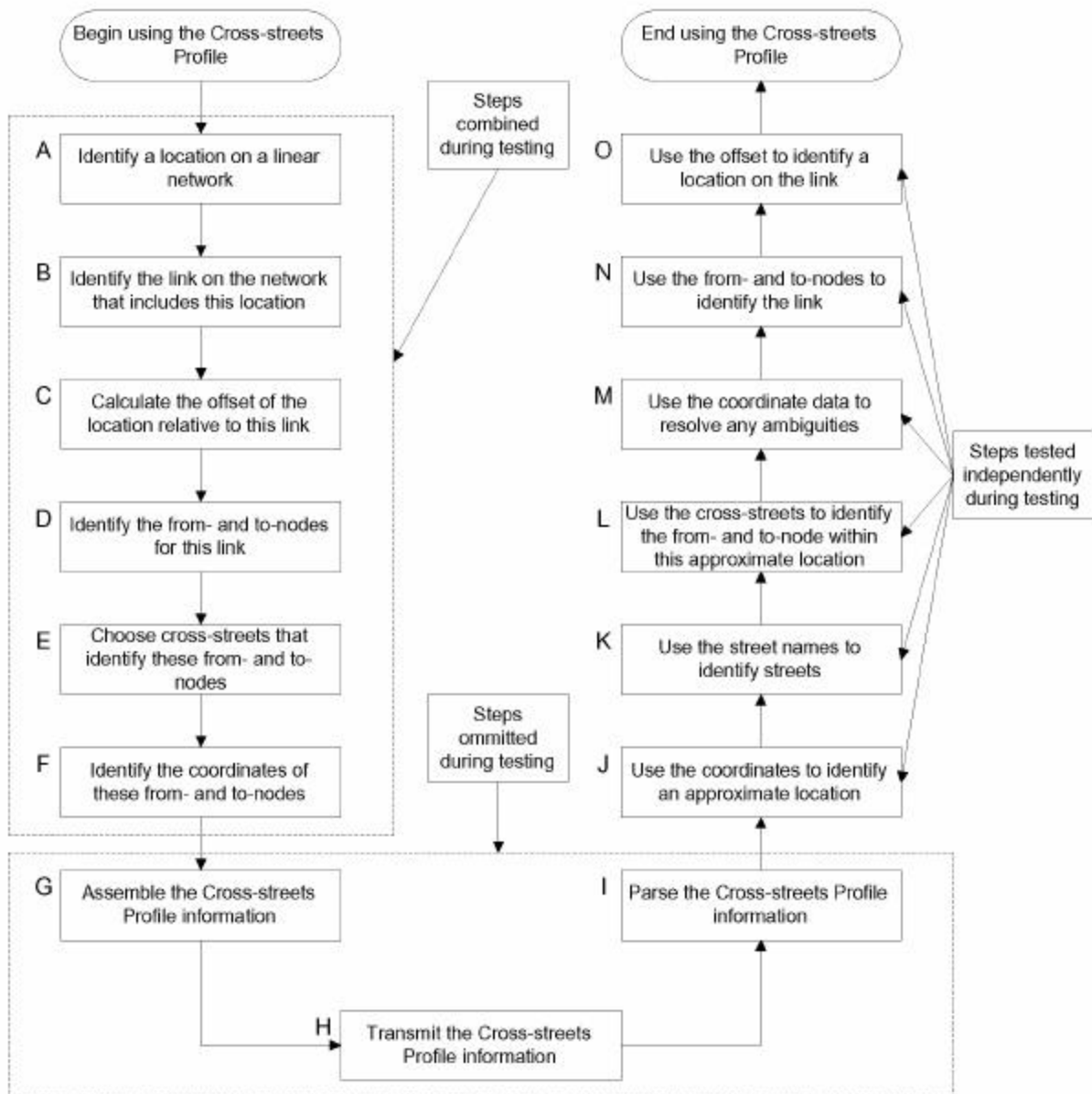


FIGURE 9—TYPICAL USAGE OF THE CROSS-STREETS PROFILE

7.4.2 SAMPLE POINTS FOR CROSS-STREETS PROFILE TESTING—The following list describes the types of samples used for most of the Cross-streets Profile testing:

- a. When a test involved only a single data set, the sample consists of every instance in the data set. The following situations apply to the testing:
 1. When testing nodes (e.g., the number of nodes with a blank street name in the cross-streets reference), each node is tested.
 2. When testing links (e.g., the number of links with a blank street name), each link is tested.
- b. When a test involves two data sets, the sample consists of every instance in the first data set (i.e., the transmitting data set). In the tables, the data sets are identified using the terminology “[transmitting data set] to [receiving data set].” The following situations apply to the testing:
 1. When testing streets (e.g., the number of street names in the transmitting data set that match a street name in the receiving data set), each street is tested.
 2. When testing nodes (e.g., the number of nodes in the transmitting data set whose cross-street reference matches that of a node in the receiving data set), each node is tested.
 3. When testing links (e.g., the number of links in the transmitting data set whose cross-street reference matches that of a link in the receiving data set), each link is tested.

In order to address the following specific issues, the different types of samples are used:

- a. The effectiveness of the Cross-streets Profile might differ by roadway classification. In particular, the effectiveness of the Cross-streets Profile might be different for primary roadways, and initial ITS applications are likely to be limited to primary roadways. Some results are generated for each class of roadways described in Table 9.
- b. The effectiveness of the Cross-streets Profile might differ by geographic region. In particular, the effectiveness of the Cross-streets Profile might be different near city centers. Because the roadways near city centers typically have more traffic, the effectiveness of the profile for these roadways may be more important for some ITS applications.

7.4.3 WEIGHTING RESULTS FOR CROSS-STREETS PROFILE TESTING—Most of the results of the Cross-streets Profile testing are presented in terms of their “natural” weighting: results concerning streets weight each street equally, results concerning nodes weight each node equally, and results concerning links weight each link equally.

7.5 Preliminary Testing—Before testing the effectiveness of the Cross-streets Profile, some preliminary tests on the City of San Francisco test data were performed to help better understand the data and to optimize the test procedures. The following types of preliminary tests were performed.

- a. **Data Set Verification**—Section 6.3.2 identified data issues that could strongly influence location referencing and suggested that, if data sets differ greatly in either coverage or level of detail, then those data sets are not appropriate for testing. Section 7.5.1 describes the test results concerning the coverage and level of detail of the test data sets.
- b. **Map Topology Testing**—The effectiveness of the Cross-streets Profile can be affected by the map topology, and this testing investigated the presence of topological factors that could influence this effectiveness. The preliminary map topology testing is described in 7.5.2.
- c. **Street Name Testing**—Tests were performed to investigate the prevalence of blank street names and to identify a simple, effective means of “normalizing” street names to improve street name matching. The preliminary street name testing is described in 7.5.3.

- d. **Coordinate Testing**—Tests were performed to investigate the consistency of the coordinate data in the test data sets. The preliminary coordinate testing is described in 7.5.4.
 - 1. **Generation of field data.** Field data was generated for Santa Barbara County by using a differential global positioning system (DGPS) receiver to determine the coordinates of test points and recording the street locations. See Appendices A and B for more information on the field data.

Detailed results from the preliminary testing are presented in Section A.1, and summaries of these results are provided in the following sections.

7.5.1 DATA SET VERIFICATION—The following results summarize the data issues discovered during data set verification:

- a. **San Francisco Data**—Each of the data sets for San Francisco had similar coverage areas and similar levels of detail. Street names were generally present throughout the data sets.
- b. **Santa Barbara Data**—Four of the five data sets for Santa Barbara had similar coverage areas, though the coverage area of the fifth data set was significantly smaller than that of the other four. The level of detail of the four data sets showed some variation in the outlying areas (one data set had 30% fewer links than the others), and was markedly less for the fourth data set. Street name coverage followed this pattern to an even stronger degree; street names were often missing in the outlying areas, but were commonly available near the city center. Because of these differences, any test results that encompass the entire county must be interpreted carefully.
- c. **Santa Barbara City Data**—Because of the differences in the test data sets for the County of Santa Barbara, a geographic region near the city center was defined and identified as Santa Barbara City Data. The data sets, when restricted to this region, were much more uniform in coverage, level of detail, and presence of street names.

7.5.2 PRELIMINARY NODE AND LINK TOPOLOGY TESTING—The following results summarize the results on the features of the node and link topology that were investigated during the preliminary testing:

- a. **Node Valences**—The valences (i.e., the number of links that terminate at a node) of the nodes in a map data set is an indirect indicator of problems that might be encountered when applying the Cross-streets Profile. A node with a valence of 1 represents the end of a dead-end street, and a link that terminates at such a node cannot be represented by the Cross-streets Profile. A node with a valence of 2, if both links that terminate at the node have the same street name, also has no Cross-streets Profile representation. A node with valence greater than 4 is likely to have several valid cross-street references.
- b. The majority of the nodes in the test data sets have valence 3 or 4, nodes for which the Cross-streets Profile is likely to create a unique cross-streets reference. However, 20% to 25% of the nodes in the San Francisco data sets and up to 45% of the nodes in Santa Barbara City have valence 1 or 2 nodes, and no cross-streets references will exist for many of these nodes. The next topic confirms this supposition.

Valid cross-streets for nodes and links. As mentioned above, some nodes do not have valid cross-streets references, and other nodes have multiple valid cross-street references. The same applies to the links in the test data sets. In particular, up to a quarter of the nodes and links in the San Francisco test data sets and up to half in the Santa Barbara City test data sets have no valid cross-streets reference. Most of these “bad” nodes are caused by dead-end or valence 2 nodes, as described in the previous item. These “bad” nodes are the primary cause of “bad” links; the remaining “bad” links are caused by blank on-street names. Section 7.8 includes suggestions for reducing the number of nodes and links that have no valid cross-streets reference.

- a. A smaller number of nodes and links have multiple valid cross-streets references. While this fact does not degrade the effectiveness of the Cross-streets profile, it does present a problem in interpreting the test results. If the success of the location transfer for a node or link depends upon which cross-street reference was selected, the location transfer for that node or link cannot be unambiguously classified as either "successful" or "unsuccessful."

7.5.3 PRELIMINARY STREET NAME TESTING—The following features of the street names were investigated during preliminary testing:

- a. **Blank Street Names**—The presence of blank street names limits the effectiveness of the cross-streets profile by creating the potential for matching ambiguities because blank street names carry less location reference information than non-blank street names. Blank street names also prevent the formation of a valid cross-streets reference for some nodes because all of the street names at the node are blank. In the test data sets, between 1% and 5% of the street names are not specified in the San Francisco test data sets. For Santa Barbara City, up to 10% of the street names are blank, and over 40% of the street names are blank in two of the Santa Barbara County test data sets. A similar percentage of the nodes and links have a cross-streets reference that includes a blank street name, and have links with no valid cross-streets reference because the on-street name is blank. In the end analysis, blank street names were a significant, but not dominant, cause of degradation in the effectiveness of the Cross-streets Profile in San Francisco, but were dominant in Santa Barbara County.

On the other hand, blank street names can be a significant source of confusion when implementing the Cross-streets profile. In particular, one must decide when to allow blank street names in cross-streets references and how to interpret blank street names when trying to identify a cross-streets reference that includes a blank name. Section 7.8 includes suggestions on how to handle blank street names.

- **Street name normalization.** Because street naming standards are not universally applied across map data sets, the street names in the test data sets differ. Consequentially, preliminary testing of the street names concentrated on identifying and investigating simple methods (i.e., street name normalization techniques) that could be used to improve street name matching. Six methods of street name normalization were investigated in some detail, ranging from method 1, which did little besides convert the names to upper case, to method 6, which implemented extensive normalization.

The testing indicated that the street names for three of the four test data sets were fairly consistent (about 80% to 90% overlap) without any normalization. For these data sets, the most extensive street name normalization added another 3% to 4% overlap at the cost of nearly 10% of the street names matching multiple street names. Later testing indicated that, because of the built-in redundancy in the cross-streets profile, the extraneous street name matches caused by the street name normalization techniques did not often result in extraneous node and link matches. The fourth data set achieved street name matching nearly on par with the other data sets when street name normalization was applied.

In addition to the tests on how street name normalization might impact street name matching, a manual analysis was performed to identify the reasons that street names did not match. This analysis identified the following features of the unmatched street names:

- a. The majority of the unmatched street names were caused by apparent errors in the test data sets
- b. A significant number of the unmatched street names would have matched using the most extensive normalization method
- c. Naming conventions for ramps and inclusion of street name aliases could significantly increase the effectiveness of street name matching