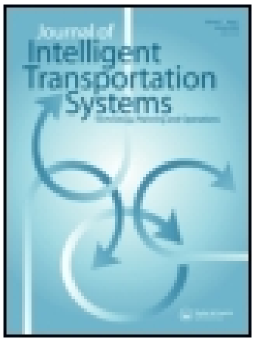

Journal of Intelligent Transportation Systems

ISSN: 1024-8072 (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gits19

A NEW LOCATION CODING SCHEME FOR INTELLIGENT TRANSPORTATION SYSTEMS

Zhihui Huang & Kang G. Shin

ITS Journal, 1996, Vol. 3(2), pp. 99–109
Reprints available directly from the publisher
Photocopying permitted by license only.

A NEW LOCATION CODING SCHEME FOR INTELLIGENT TRANSPORTATION SYSTEMS

Zhihui Huang and **Kang G. Shin**
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, Michigan

Many Intelligent Transportation Systems (ITS) applications require some form of location information. For example, traveler information services and traffic management are both critically dependent on the specification of location(s). Some ITS applications may need this information in finer resolution than others as demonstrated by the navigation and route guidance systems vs. severe weather warnings for a local area. A location can be a street address, a highway segment, an intersection, a stretch of roadway or a small geographical area. These types of locations can be coded in various ways. In this paper, we first examine several coding schemes reported in the literature and/or used in real applications. Then, we propose a new location coding scheme based on the first three digits of the Zipcode plus local information. This scheme uniquely represents all useful types of locations and is versatile and efficient. Moreover, it fits well the location reference message protocol (LRMP) described in [Goodwin et al., 1995; Goodwin and Xiong, 1994]. To demonstrate its utility and power, we compare the proposed scheme with the Road Name ID scheme [Ramakrishnan et al., 1994].

Key words: Intelligent Transportation Systems (ITS), Advanced Traveler Information System (ATIS), location coding, location naming, Zipcode-based coding, Zip+Local.

The ability to uniquely encode and identify a location or a collection of locations easily and efficiently is critical to many ITS applications (e.g., ATMS and ATIS). An ITS location can be a street address, a highway segment, an intersection, a stretch of roadway or a small geographical area. Some applications require a finer level of details in location identification than others. For example, a map database for navigation and route guidance represents a level that requires more accurate coding (in a few meters), while traffic and weather reports represent another level that requires less accuracy. A traveler may inquire about traffic and/or weather conditions of a location, or s/he may ask about lodging information in a particular area. When a low bandwidth wireless channel is used for these inquiries, an inquiry from the client and the response from the information server should be as compact as possible. Only recently there have been published reports on the location naming convention [Ramakrishnan et al., 1994].

In general, any satisfactory ITS location coding scheme should be:

1. Universal for all intended geographical areas which normally include metropolitan areas and rural roadways. Location finding in off-road areas can be dealt with using GPS-capable devices.
2. Not tied directly to any specific vendor or enabling technology. This is important because of the wide variety of delivery technologies that will be used in practice. One example that does not satisfy this condition is the Alert-C protocol (see Alert-C).
3. Able to distinguish the two directions of a divided highway where an incident occurred on one side of the road does not normally interfere with the traffic flow on the other side. It should also be able to distinguish locations on stacked highways.
4. Able to depict small geographical areas. This is useful for describing weather or general road conditions.
5. Able to let a user inquire with insufficient information about locations and/or services, such as electronic yellow books.
6. Compact. Because bandwidth limitations of many enabling technologies (e.g., FM subcarrier), the coding should be compact in order to be useful by all delivery technologies.

RELATED WORK describes several existing and proposed location naming schemes including the LRMP protocol. The PROPOSED APPROACH outlines ZIP+Local scheme. The naming of various location types using the proposed scheme and the mapping to/from the user-provided information are discussed next, followed by CONCLUSIONS.

RELATED WORK

One single standard location referencing (SLR) must be able to "glue" all of the ITS functionalities. The authors of [Goodwin et al., 1995] presented an exhaustive analysis of functional requirements of ITS map database applications. They concluded emphatically that a single, standardized, across-the-board ITS system in not achievable or advantageous. Rather, they suggested the strategy of "appropriate(ing) IVHS." In the context of location referencing schemes, the authors proposed an umbrella message protocol the Location Reference Message Protocol (LRMP) to provide a conduit for different location schemes, while maintaining a standard application interface at the system level (see also [Goodwin and Xiong, 1994]).

In the remainder of this section, we will evaluate several proposed location naming schemes with respect to the required features listed above. These schemes, along with our own proposal, can best be viewed as individual application protocols that are to be encapsulated by LRMP in implementation.

Alert-C

The European Traffic Message Coding Protocol (Alert-C) [RDS ALERT Consortium, 1990] was developed to unify location and message coding for RDS (Radio Data System), utilizing one of FM radio subcarriers to transmit digital data. It provides unique location naming throughout Europe.

The usable bandwidth of an FM subcarrier is small (with a channel net data rate of 1187.5 bps). RDS contains sixteen message groups which compete for bandwidth. Of these sixteen groups, only the Traffic Message Channel (TMC) can be used to transmit traffic messages that conform to the Alert-C protocol. As a result, only a 256 bps data rate can be attained for TMC under peak operational conditions (a maximum of eight TMC groups, with 64 information bits each, can be transmitted in any two-second interval). Alert-C is designed to be compact in order to conserve the bandwidth. It is a protocol that relies on implicit location information transmitted through periodic "system information messages," so that only a 16 bit location field needs to be explicitly specified in a message. The locations covered by the European Broadcasting Union (EBU) stations and defined in the Alert-C protocol are divided into hierarchical location databases in the form of (*Country Code, Database Number, Location*), where *Country Code* is four bits long, *Database Number* is six bits long, and *Location* is sixteen bits long. Of these three elements, only *Location* is transmitted in the actual TMC messages. Short messages (as defined in the protocol), containing 8-bit location codes, utilize an additional attribute called "location subset number"—also transmitted in system information messages—to cross-reference 256 locations.

The difficulty of Alert-C lies in forming user inquiries and its reliance on one particular delivery technology. It is designed for a one-way communication environment (FM broadcasting). It is trivial to map a location code to an actual location while the reverse map from an actual location of interest to the complete location code in nontrivial. This is also because that the naming is based on implicit coding of countries and FM station coverage areas.

(Latitude, Longitude) Pair

Since the Global Positioning System (GPS [Getting, 1993]) has been made available for civilian use, one can expect to obtain GPS location information with errors as small as one meter, with differential position correction.

One natural proposal is to use the (latitude, longitude) pair to represent locations. While this scheme provides universal location naming and it is compact, there is one factor against this proposal: the (latitude, longitude) pair requires and provides a level of detail in naming that is an overkill for many intended applications. In order to name a segment of highway, for example, a series of (latitude, longitude) pairs with a certain density are needed.

Another weakness of the (latitude, longitude) scheme is its inability to identify road segments such as overlapping segments of stacking highways, where the (latitude, longitude) information for points on different highways can be the same.

Road Name ID Scheme

Ramakrishnan et al. [Ramakrishnan et al., 1994] proposed a scheme that uses road names as the base for universal location naming. *All* street/road names used in the country are first collected and duplicate names removed. Each unique street name is given an identifying number. The scheme then uses indexed road names as the basis, combined with directional information, street types and regional information (in the form of city numbers) to form unique identification of locations throughout the country. This scheme has many merits such as low volatility of the data set since new roads constructed are likely to use names that have already been used, possibly in other cities. However, it takes four location identifiers (IDs) in the scheme to identify a small problematic area. It is difficult to use the scheme to describe larger areas that are not bounded by named roads, such as in weather reports.

Metropolitan Area Dependent Coding

A location coding scheme [Kady and Ristenbatt, 1993] that is based on segment IDs and designed for RBDS (Radio Broadcast Data Systems, the US counterpart of the European RDS, see [NRSC, 1992]). The scheme calls for a local authority to identify the most traveled road segments and small strategic areas in a particular metropolitan area's road network, map each such location to a uinque 14-bit integer. While this scheme may work for one metropolitan area, it may not work when a cross-metropolitan area (global) inquiry is desired.

THE PROPOSED APPROACH

The United States Postal ZIP+4 coding can be used to systematically and uniquely identify locations down to a city block or blocks, as stated in the ZIP+4 State Directory (United States Postal Service, 1986):

> The first two digits of the "+4" denote a delivery "sector," which may be several blocks, a group of streets, several office buildings or a small geographic area. The last two digits denote a delivery "segment," which might be one floor of an office building, one side of a street, a firm, a suite, a post office box or group of boxes, or other specific geographic location.

There are some difficulties with the ZIP+4 coding when one attempts to use it to represent location(s) of traffic incidents. For example, the ZIP+4 coding of highway segments in non-intuitive and tedious. Furthermore, on surface streets, any block (between two intersections) of the road is likely to have two codes, one for each side of the street. Any two consecutive blocks of the street are likely to have different codes and the coding sequence is not necessarily regular. This makes it difficult for a motorist to search and input the location codes that are assigned to a certain route of choice.

Note that the first three digits of the 5-digit Zipcode provides a well-established, non-overlapping naming scheme based on geographic locations of areas. The first digit divides the country into ten regions from east to west. The second and third digits identify large cities or other geographical areas. The fourth and fifth digits of the Zipcode identify a post office or a local delivery area. Obviously, the last two digits suffer the same drawbacks as the "+4" code when used to code locations for ATIS.

Our proposed scheme is to augment the Metropolitan Area Dependent Coding with Zipcode-specified districts. It uses the first three digits of the 5-digit Zipcode to uniformly identify (relatively) large areas, or districts, across the country. Within each district, we devise a new scheme to name local locations for the purpose of reporting traffic incidents. With this scheme, a "district" as depicted in the Loral-IBM IVHS Architecture report [IVHS America, 1994] can contain one or more Zipcode-specified areas.

The naming of locations within a given district can be divided into four categories for ITS reports and inquiries: residential streets/areas, freeways/expressways, major arteries, and points of interest. The coding of each of these four categories is described below.

ZIP+LOCAL LOCATION CODING

Residential Street in General

Residential streets are small roads that connect residential areas with major arteries and freeways. The distinct features of residential streets are: (1) each carries a small amount of traffic, and (2) an incident on them causes less delay (than that on other categories) and a driver can normally re-route through residential streets in the immediate vicinity of the incident. It is these features that have led us to encode residential streets in small areas, instead of encoding each individually.

The area within each district is divided into small areas as shown in Figure 1. For each district a rectangle is defined to contain the entire area. A geographic center is selected for the rectangle which can be its geometric center or a logical (based on importance) center of the area, such as the center of the city within the district. The center is used as the origin. Two lines are drawn horizontally and vertically to divide the area into four quadrants. The northeast quadrant is coded as "00," the northwest as "01," . . . , etc. The same process is then applied recursively for each quadrant obtained in the previous step (see Figure 1) for a total of eight steps such that the smallest identifiable areas in the district are coded in sixteen binary digits or bits (for a total of 256 x 256 such areas). Once the process is completed, the following format is used to code a specific area:

(District Code, Location Type, Resolution, Location Code)

where *District Code* is ten bits long; *Location Type* is specified with two bits ("00" for *area* in this case); *Resolution* is encoded with four bits. Its value multiplied by two gives the number of bits in *Location Code* should be used for this location; *Location Code* consists of sixteen bits. This coding provides us with means to uniquely and easily identify locations of different sizes. For example, the entire district, Area A and B in Figure 1 are uniquely identified as:

Whole Area:	011110001:00:0000:xxxxxxxxxxxxxxxx
Area A:	011110001:00:0001:10xxxxxxxxxxxxxx
Area B:	011110001:00:0001:000011xxxxxxxxxx

Note that districts differ in size geographically. A district at the center of a metropolitan area is smaller than the one in a less populated or less business area. Thus, the same resolution in location coding gives more details in a large city. For example, in district 481—Ann Arbor, Michigan and surrounding areas—the smallest identifiable areas are approximately 1000 x 1000 ft^2).

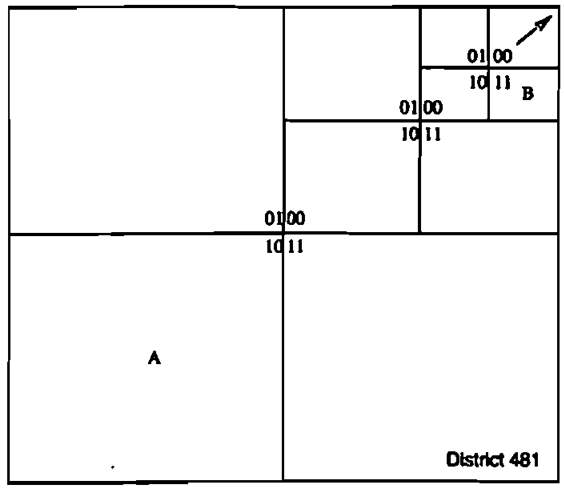

FIGURE 1 Encoding of Small Areas in a District

Freeways and Expressways

Freeways and expressways have characteristics quite different from residential (or local) streets:

- Distance between two adjacent exits on them (excluding interchanges) is usually larger than a street block.
- They are unidirectional. Traffic or incidents in one direction of an expressway/highway normally do not interfere with the traffic in the other direction.
- Areas alone cannot correctly identify an expressway that lies on top of another at an interchange. This is important because an incident occurred on one does not necessarily affect the traffic on the other.

Expressways are represented with ten bits, retaining their normal road number, if any, such as I-75. A (road) number that has not already been used will be assigned to an unnumbered expressway. Each expressway is broken down into "segments," each of which is the stretch of the roadway between two exits. The segments of one expressway are numbered with eight bits to provide the ability of numbering 256 segments within one district ("00000000" can be reserved to code the entire highway). For each segment, an additional 2-bit directional information is given to denote the direction(s) in which the traffic is affected:

(District Code, Location Type, Highway Number, Direction, Segment)

where *District Code* is ten bits long; *Location Type* is two bits long with "01" representing expressways; *Direction* means: 00 = entire expressway, 01 = East/South, 10 = West/North, and 11 = both directions.

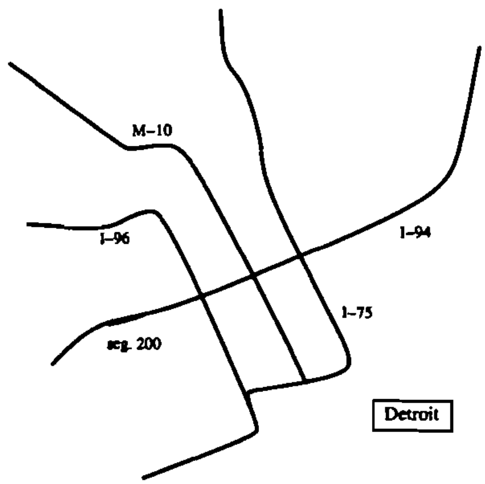

FIGURE 2 Example Highways in Detroit Area

Figure 2 shows several highways in the Detroit area, with the following naming examples:

I-94, Seg. 200, East:	0111100011:01:0001011110:01:11001000
I-94, Seg. 200, Both:	0111100011:01:0001011110:11:11001000
I-94, Entire:	0111100011:01:0001011110:00:00000000

Major Arteries

Major arteries are surface roads that carry large amounts of traffic, like freeways, but, unlike freeways, they intersect numerous smaller roads. The traffic in the two directions of arteries may or may not be physically divided. These roads often have on- and off-ramps to freeways and constitute a significant part of commute routes. This makes it necessary for motorists to be able to uniquely identify an artery when s/he makes location-related inquiries.

Like expressways, arteries can be numbered. Each artery is then broken down into "segments." Each segment is the stretch of the roadway between two major intersections with other arteries or freeways (i.e., with ramp accesses to freeways).

A location on one artery is coded in the following form:

(District Code, Location Type, Artery Number, Direction, Segment)

where *District Code* is ten bits long; *Location Type* is two bits long with "10" representing arteries; *Direction* has the following meanings: 00 = the entire artery, 01 = East/South, 10 = West/North, 11 = both directions.

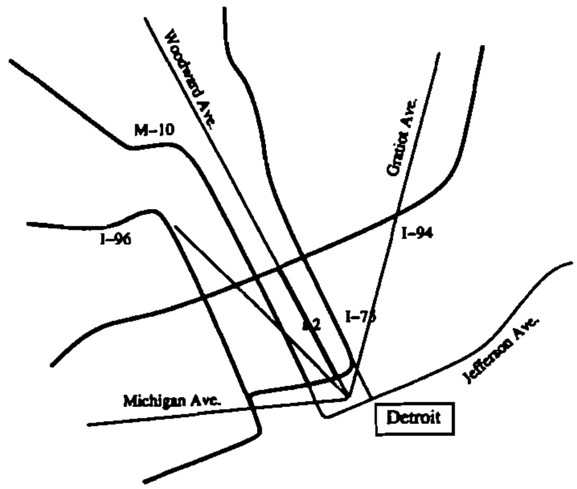

FIGURE 3 Example Arteries in Detroit Area

Figure 3 shows several arteries in the Detroit area, where Woodward Avenue is arbitrarily chosen with a number "1." Example namings are:

Woodward, S.2, North: 0111100011:10:0000000001:10:00000010
Woodward, South: 0111100011:10:0000000001:01:00000000

Points of Interest

Another type of location is points of interest. This is specific to traffic and/or weather condition reports and inquiries. For example, in case of a special event, one would be interested in knowing the location of that event, and the extent to which the traffic would be affected around that location.

Point's of interest are coded as:

(District Code, Location Type, Location Number, Extent)

where *District Code* is ten bits long; *Location Type* is two bits long with a value of "11" representing points of interest; *Location Number* is ten bits long containing the identification of a location of interest in the district; *Extent* is ten bits long and represents the radius of a circle around the point that the report/inquiry deals with (in units of street blocks, for example).

Figure 4 shows several points of interest in the Detroit area. These example namings are coded as follows:

Henry Ford Hospital: 0111100011:11:0000000001:0000001010
Belle Isle Park: 0111100011:10:0000000010:0000000000

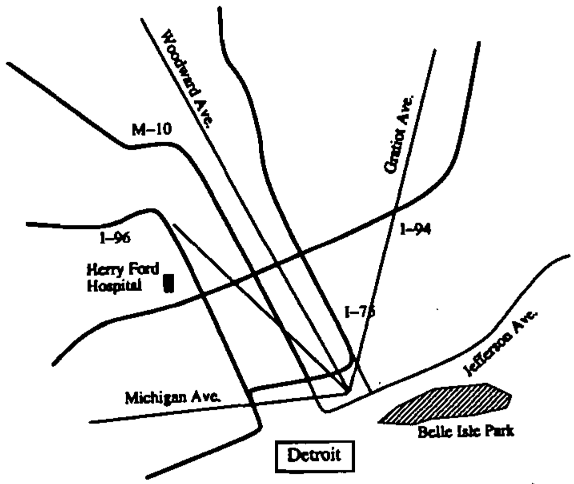

FIGURE 4 Example Points of Interest in Detroit Area

Mapping

A summary of location coding formats for the proposed scheme is shown in Table I. Note that the number in the pair of parentheses "()" in the Location Type column denotes the binary value for each type while the number in the pair of parentheses in the Coding column denotes the width of the field in bits.

The locations as specified above are internal notations for an ITS. To a user, the more natural naming of locations is by street names, exit numbers, intersections, or points of interest. A mapping between conventional names and their IDs is provided after the naming is completed. It provides an easy, conventional interface to the users. Where possible, graphical user interfaces can be used to click and point the locations of interest.

A user-selected route such as a commute route can be identified as the concatenation of stretches along different highways/arteries within one or more districts. Each stretch consists of two locations: the starting segment and the ending segment. Not all the segments in any one route need be specified in the scheme.

With the help of a map database, raw GPS location information can be translated into the coding for that location through the following two steps: (1) the GPS location is first converted into a road name and segment number via the map database; (2) they are

Table I. Coding for Different Types of Locations

Location Type	Coding
Small Areas (00)	District Code (10): Location Type (2): Resolution (4): Location (16)
Highways (01)	District Code (10): Location Type (2): Hwy. No. (10): Direction (2): Segment (8)
Arteries (10)	District Code (10): Location Type (2): Art. No. (10): Direction (2): Segment (8)
Points of Int. (11)	District Code (10): Location Type (2): Location No. (10): Extent (10)

converted into location ID through a location-definition database. The district information can be obtained through one separate layer of the map database superimposed over the road network. The entire conversion process should be repeated continuously by the on-board computer or as frequently as desired.

Comparison with Road Name ID

The proposed scheme is compared with the Road Name ID scheme [Ramakrishnan et al., 1994] in Table II, where RID is the road ID; DIR specifies the direction of the road identified by RID; TYPE specifies the type of road that RID is; CID identifies the administrative region in (city, state) pair.

While Road Name ID scheme is capable of identifying smaller, residential roads, they normally constitute only a very small part of the congestion problem. They are less important, because there are typically alternative routes the driver can take to avoid or go around the problematic location on smaller roads. It is our design choice not to include them since a route guidance or navigation system with a map database can provide sufficient coverage for this purpose. On the other hand, the proposed Zip+Local scheme is able to express more types of locations and do so more efficiently. Specifically, the ability to identify small geographical areas is desirable in many applications. For example, weather report/forecast, or traffic updates such as "water main broken makes roads in area X extremely hazardous," or "downtown surface roads are all clogged up by rush hour traffic. However, freeways and toll ways are running smoothly." In addition, this capability can easily support "fuzzy" inquiries such as "Where can I find a fast-food restaurant in area Y?" The support for points of interest in the proposed scheme allows convenient specifications such as "how is the traffic around Tiger stadium right now?"

The disadvantage of using Zip+Local scheme alone is the fact that it is relatively difficult to express specific points on the roadway such as ADA pickup/drop-off points, location of public transit stations on small roads, etc. However, these are the problems that are best addressed with the help of a map database and GPS-capable units, which co-

Table II. Zip+Local vs. Road Name ID

	Zip+Local	Road Name ID
Efficiency: segments	32 bits	3 x (*RID* + *DIR* + *TYPE* + *CID*) (RID alone needs 17 bits)
Efficiency: areas	32 bits	Difficult to express
Efficiency: points of int.	32 bits	Difficult to express or 2 x (*RID* + *DIR* + *TYPE* + *CID*)
Roads covered	Highways and Arteries	All
Expandable	Limited by bit assignment	Limited by bit assignment
Language dependent	no	no
Architecture dependent	no	no

exist with applications that use Zip+Local scheme to specify locations. A possible conversion process between the different location formats can be found in Mapping.

CONCLUSIONS

The Zipcode provides a well-established means of uniquely naming a (relatively) small geographic area in the United States. Alternative naming schemes such as the ones presented in this paper have their weaknesses. We proposed a method that uses part of the Zipcode scheme to identify ITS/IVHS "districts" and a local area/segment coding scheme to identify small areas, segments of highway/major arteries, and/or points of interests.

The proposed scheme provides a naming scheme that: (1) identifies a location uniquely nationwide, (2) does not depend on the service delivery agent (such as a particular FM station or a particular base station in a cellular network), (3) provides unique naming of small areas in a metropolitan area, (4) addresses the special needs of ATIS applications such as regional, even nationwide inquiries, and finally, (5) is more efficient and/or expressive than the other schemes discussed. It is also suitable for inclusion in the framework of the Location Reference Message Protocol (LRMP).

ACKNOWLEDGMENTS

The work reported in this paper was supported in part by Loral Incorporated and the US Department of Transportation under Grant No. DTFH61-93-X-00017. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the funding agency.

REFERENCES

Getting, I. A., "The Global Positioning System," IEEE Spectrum, 1993, 30 (12) pp 36–47.

Goodwin, C. W. H. and D. Xiong, White paper: "The Location Reference Message Protocol for ITS Applications," technical report, Oak Ridge National Laboratory, December 1994.

Goodwin, C. W. H., Xiong, D., and S. R. Gordon, "Functional Requirements for National Map Databases for IVHS," technical report, Oak Ridge National Laboratory, January 1995. Review Draft (Final).

IVHS America, "IVHS Architecture Development Program," (interim status report), April 1994.

Kady, M. A. and M. P. Ristenbatt, "An Evolutionary IVHS Communication Architecture," Vehicle Navigation & Information Systems Conference, 1993.

NRSC. United States RBDS Standard (draft), April 1992.

Ramakrishnan, R, C. Collier, and D. Behr, "Standard Location Referencing (SLR) System: Road Name ID Scheme," Moving Toward Deployment, Vol. 2, pp. 620–24. IVHS America, 1994. Proceedings of the IVHS America 1994 Annual Meeting.

RDS ALERT Consortium, "Alert C: Traffic Message Coding Protocol Proposed Pre-Standard, November 1990.