An Integrated Approach to the Conceptual Data Modeling of an Entire Highway Agency Geographic Information System (GIS)

Vorgelegt von
M. Sc. Eng. Hande Demirel

von der Fakultät VI –
Bauingenieurwesen und Angewandte Geowissenschaften
der Technische Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
- Dr. Ing. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr.-Ing. Dieter Lelgemann
Berichter:	Prof. Dr.-Ing. Lothar Gründig
Berichter:	Prof. Dr.-Ing. M. Orhan Altan
Berichter:	Dr.-Ing. Habil. Martin Scheu

Tag der wissenschaftliche Aussprache: 15. Februar 2002

Berlin 2002
D 83

Acknowledgements:

During the preparation of this thesis a large group of people have supported the personal effort, whom I like to acknowledge their contributions. I would like to express sincere appreciation to my supervisors Prof. Dr. Lothar Gründig and Prof. Dr. Orhan Altan for giving me several excellent ideas as well as carefully examining the many manuscripts of this thesis and overall being distinguishing “doctor father”s. I wish to thank Dr.-Ing. habil. Martin Scheu for carrying out the official examination of this thesis.

I wish to acknowledge the intellectual contributions made by faculty colleagues, especially Dr. Frank Gielsdorf. I appreciate his enthusiasm and interest about spatial information theory as well as for his brilliant suggestions and discussions. A special thanks to Erik Moncrieff for his great assistance during checking and correcting the written English of this thesis.

I would like to thank for their cooperative attitude among the various components of the highway community-researchers, government executives, engineers and others with commitments and enthusiasm about spatial information systems.

Last, but surely not the least, thanks and appreciation for my parents, Ayşe and İsmail Demirel, for their support and encouragement.

Berlin, 2002

Hande Demirel

Titel der Dissertation:

Integrierter Ansatz zur konzeptionellen Datenmodellierung eines geographischen Informationssystems (GIS) für Daten der Straßenverwaltung

ZUSAMMENFASSUNG:

Um ihre Aufgaben zu verwirklichen, sind Straßenverwaltungen in aller Welt dazu aufgefordert, neue Technologien einzuführen. Grund dafür ist die große Menge an anfallenden Informationen des Straßennetzes und die Notwendigkeit Datenquellen effizient zu nutzen. Geoinformationssysteme für das Transportwesen (GIS-T), welche speziell für Straßenverwaltungen entwickelt wurden, bewirken eine erhebliche Effizienzsteigerung, da sie am besten in der Lage sind, dem räumlichen Charakter der Daten Rechnung zu tragen. Häufig wurde dieser räumliche Charakter der Informationen bei der Systementwicklung ungenügend beachtet, was dazu führte, dass die Möglichkeiten solcher Systeme nicht voll ausgeschöpft wurden. Die Implementierung eines Systems kann nur dann zu vollem Erfolg führen, wenn eine detaillierte Informationsstrukturanalyse durchgeführt wird und wenn die Datenmodellierung formalisierten Entwurfsmethoden folgt. Im Verlauf der Untersuchungen wurde festgestellt, dass gebräuchliche Systeme verschiedene Anforderungen von Straßenverwaltungen nicht erfüllen. Die Probleme können wie folgt zusammengefasst werden: Die Beziehungen zwischen geometrischen, topologischen und Sachinformationen wurden nicht strukturiert. Die Abbildung von geometrischen Informationen in unterschiedlichen Referenzsystemen war nicht redundanzfrei möglich. Die Verwaltung topologischer Informationen in unterschiedlichen Abstraktionsebenen wurde nicht realisiert. Spezifische Funktionen der Straßenverwaltung wurden nicht in ihrer Gesamtheit abgebildet. Nicht alle existierenden Informationen und Methoden konnten in die Systeme integriert werden.

Es ist erforderlich, Metadaten wie Konsistenzbedingungen, Qualitätsangaben und Historisierung im System zu berücksichtigen. Speziell für die Definition von systemübergreifend eindeutigen Objektidentifikatoren sind neue Ansätze erforderlich.

Um die Effizienz von GIS-T zu verbessern und die beschriebenen Anforderungen zu erfüllen, wird in der vorliegenden Arbeit schrittweise ein Ansatz für eine konzeptionelle Datenmodellierung vorgestellt, welche den Bedürfnissen einer Straßenverwaltung Rechnung trägt. Der Grundgedanke des vorgeschlagenen Modells besteht in der Abstraktion und der strengen Unterscheidung von geometrischen, topologischen und Sachdaten. Um die Integration aller Daten, die Kontrolle von Redundanz und eine Optimierung der Datenpflege zu erreichen, wurden Trassierungselemente durch datumsinvariante Parameter abgebildet. Das vorgeschlagene konzeptionelle Datenmodell wurde erfolgreich implementiert. Dabei kam ein objektrelationales Datenbanksystem zum Einsatz.

Title of Dissertation:

An Integrated Approach to the Conceptual Data Modeling of an Entire Highway Agency Geographic Information System (GIS)

ABSTRACT:

World-wide highway administrations are stressed to implement new technologies, due to the large amount of information associated with highway networks and the necessity of using sources efficiently in order to realize their tasks. Geographic Information Systems-Transportation (GIS-T), which are specifically tailored for highway administrations, are identified having the highest information technology payoff potential by the highway administrations due to road information spatial character. Contrarily, road information spatial character is not adequately considered during system design, as a result, many of the benefits of GIS-T are not fully realized and efficiency of this technology is mainly under estimated. The relative success of implemented system is not clear without a detailed information analysis and a data model, which rely on formal data model design methodologies. It was determined during this study that several demands of highway administrations were not responded by means of current systems. These topics can be summarized as follows; firstly, relationships among geometry, topology and thematic information were not structured. The geometry information can not be mapped in various reference systems without redundancy. Thirdly, the non-planar multi- abstraction topological information was not exist. The entire highway administrations business rules can not be performed in the current systems. The existing information and methods were not integrated into the system.

The metadata including consistency rules, quality specifications and history information needed to be incorporated into the system. Especially in order to determine permanent, non-spatial and a unique object identifier, regulations and new approaches are required.

In order to increase the efficiency of GIS-T and fulfill these requirements, this study considered a progressive approach appropriate to the conceptual data modeling requirements of an entire highway agency. The main approach of the proposed data model was abstraction and decomposition of geometry, topology and non-spatial data. In order to achieve data integration, control of redundancy and optimization of data maintenance, linear elements were mapped by means of datum invariant parameters. The proposed conceptual data model was successfully implemented using the integrated approach in one object-relational system and results were discussed.

CONTENTS

FIGURE INDEX:

TABLE INDEX:

Chapter 1 : Introduction

1.1 Introduction

World wide, a very large amount of information associated with the national and international highway networks exists. Coupled with the necessity of using these sources efficiently, highway organizations are stressed to research and implement new technologies for realizing their tasks. The rapid development in technology has also given rise to the improvement of decision support systems. Consequently, many highway administrations moved to Geographic Information Systems (GIS) technology to assist their decision-making.

Geographic Information Systems tailored for highway administrations are called Geographic Information Systems-Transportation (GIS-T). According to a study done in the USA (1992), almost 80 % of the transportation departments responded that GIS-T technology had the highest information technology payoff potential of any technology identified *[ISTEA,1995]*. Although many other solutions for project management and engineering applications exist, there are two very important reasons why GIS-T is of interest.

1. All highway administration data, with some exceptions such as legal consultance or accounting division, has a spatial nature.
2. GIS integrates highway administration data and methods, which other systems do not offer. GIS-T can be used as a logical and physical data integrator of all types of data necessary to the highway sector.

However, due to its spatial nature, many of the benefits of GIS for highway administrations such as integrating data and methods, enforcing rules and standards, cost reduction and quality improvements, are not fully realized. The spatial information in highway administrations is either not recognized or not fully considered during system design. Because of such reasons, efficiency of GIS-T technology is mainly under estimated.

1.2 Overview of Problems

Highway administration specific spatial information problems, including multi-dimensional spatial road data and the relationship between these, integration of methodologies and topology abstraction levels are the main focus of this study. In addition to these, common problems of GIS, such as data integration, data inconsistencies, and non-adequate data update are examined in this context.

Highway administrations are very large governmental organizations established in order to support transportation requirements at a national and international level. Accordingly they are charged with many diverse tasks. In order to fulfill these tasks, data from various sources is acquired by road administrations in different forms and formats. Each department has different usage requirements. Although the collected data for organization tasks is in very large amount, full efficiency is lost due to its analogue format. In addition there are many methodologies, enterprise-rules and terms, which are not generally common, even between departments. Research conducted in the United States reported that thirty-eight referencing methods were being used in a single federal state highway administration *[NCHRP,1997]*. Additionally due to workflow needs information needs to be transferred between many divisions, departments and organizations.

In order to facilitate these tasks, an information technology infrastructure, especially databases and GIS, is needed. But in some situations existing databases are not updated with new information. One of the reasons for this is that organizations can be easily frustrated by the high cost of GIS implementation and data maintenance. GIS generally have long term profit expectations and high implementation costs. The typical variation in benefit and cost expectations for GIS with respect to time is illustrated in Figure 1-1. It has been noted that data costs, consisting of data acquisition, modeling and maintenance, comprise 80% of the total cost of GIS.

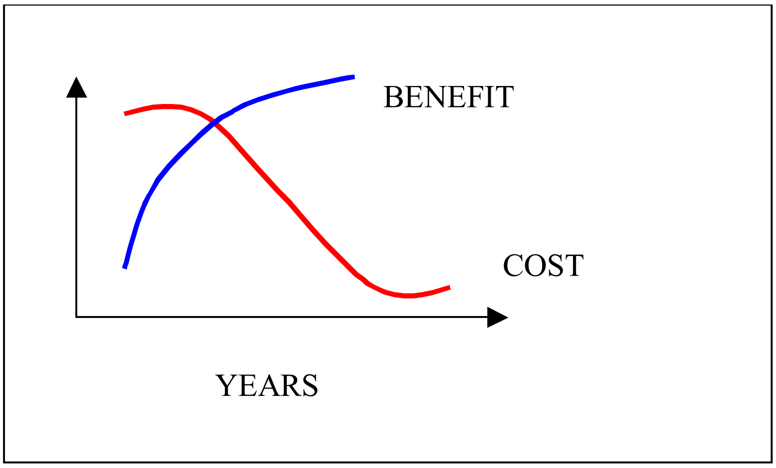


Figure 1-1: Benefit – Cost Expectations

Secondly in its early stages, GIS-T technology was applied to a variety of project-oriented transportation applications. Consequently problems known as "data islands" are very common. Available digital data is isolated within divisions and there is no data exchange between divisions and departments, even when it is desired. With many highway administrations there are difficulties in gathering and analyzing information. Inaccurate information and inconsistent data persist in the information systems.

Thirdly, road information has multi-dimensional spatial character with various levels of abstraction due to diversity in the highway agencies' requirements. Road objects can be static or dynamic, referenced to one, two, three or four-dimensional coordinates. The most common information referencing method used by highway agencies is the linear referencing system. This is based on a one-dimensional specification of the unknown point in terms of direction and distance from a known point. Many other spatial frames are in usage such as numbering systems, addresses, topology, administrative reference systems and road names. In GIS-T roads are defined using two-dimensional reference systems. In order to integrate these various dimensions, typically, geographical location by two-dimensional coordinates is used, and linearly referenced road data is considered as attributive data. However, highway information spatial character is continually changing through new alignments and construction, therefore its reference system is also continuously changing. Therefore, linearly referenced data is badly affected by such geometrical changes, requiring a new referencing for the sections after the modification. Thus, maintenance of this data is clearly necessary. Unfortunately, because of its attributive storage in GIS-T the practical realization is often insufficient.

Additionally, since transportation facilities and phenomena exist in three and four dimensions, the restriction of current data models to two and 2.5 dimensional space limits the ability of GIS to effectively model the real world *[VONDEROHE ,1993]*. Due to historical reasons, GIS concepts and implementations were initially closely associated with the requirements of land information systems. Therefore, highway administration with its complex analytical network-based models, enterprise-wide business model and topological requirements could not be well modeled using standard GIS.

Even when all these issues are considered, without a detailed information analysis and data model, the relative success of implemented system is not clear. Most organizations, and highway agencies are not exceptions, do not rely on formal data model design methodologies. This is considered one of the major causes of information systems failure. Lack of a structured approach to database design often leads to inadequacy or inefficiency in meeting the demands of organizations.

Some open issues are:

1. In which context do current systems respond to highway administration demands?
2. Is it possible to map all road related information in an integrated manner using GIS-T for a complete enterprise?
3. How can the above-mentioned problems be solved, particularly with respect to the spatial nature of road information?
4. Is it possible to solve the defined problems with current technology?

1.3. Aim of the Study

This study aims to provide a progressive integrated approach to the conceptual data modeling requirements of an entire highway agency, with the intention of addressing the above-mentioned issues.

In order to examine the current situation, clarify problematic areas and identify the requirements of highway administrations in a wider perspective through this study four countries were selected for particular attention namely; Turkey, Denmark, Germany and the USA.

During this study the current status of GIS-T usage and available information sources are examined in order to describe the highway agency specialized view of reality. Available conceptual data models are evaluated to identify problematic areas, especially due to the spatial nature of information. Highway administration specific aspects such as linear referenced data integration, cross-sectional design information, spatial data integration and identification of uniqueness are particularly considered. In addition, implementation possibilities of a proposed conceptual data model are examined and the problems encountered are discussed.

The proposed conceptual data model concentrates on special aspects of highway administrations in detail such as;

- Relationships among topology, geometry and thematic road information
- Multiple topologic abstractions
- Analyzing multi-dimensional spatial road data and realizing transformations between dimensions.
- Modeling highway administrations business rules
- Integration of existing road information and methods
- Modeling the metadata

In methodological terms, the following procedure is adopted;

1. Defining requirements, analysis of the data flow in highway agencies, examining existing data sources and relevant standards, regulations and systems.
2. Development of the data model.
3. Drafting the data schema.

4. Development of proposals for implementation.
5. Implementation of proposed concepts.

1.4. Contents of the Study

In Chapter 2 the fundamental concepts for analysis of spatial data and a general overview of data modeling steps were provided. Various data modeling approaches were compared with specific reference to GIS-T aspects.

In order to clarify problematic areas and the requirements within highway administrations in a wider perspective, organization structure, needs assessment, data acquisition techniques, existing data sources, methodologies and system architectures were evaluated. The current status of GIS-T in highway administrations was discussed, including GIS technology state, and standards established for GIS-T. Four different conceptual data models were evaluated. An overview of problematic areas is also provided in Chapter 3.

After identifying problematic areas, requirements and experiencing the perspective of highway administrations, a conceptual data model was proposed. Special aspects of highway information systems, such as linear referencing, cross-sectional design information, geometrical data integration, feature identification and proposals developed within the conceptual data model, were introduced in Chapter 4.

In Chapter 5, the proposed concepts were implemented for a sample project in order to recognize gaps and unfulfilled requirements from the conceptual data modeling.

Finally in Chapter 6 the developed concepts were discussed concluding the study.

Chapter 2 : Concepts

2.1 General Approaches

The general approach of information system development can be separated into two main categories; data-driven and method-driven. With the data-driven approach the entire focus of the design process is on data and its properties. After identifying user data requirements, a conceptual data model needs to be designed and to be implemented in a database, then applications that use the database are developed. With the alternate method-driven approach, working activities within an enterprise are determined and then application programs are designed according to user requirements *[BATINI ,1992]*. Both approaches has advantages and disadvantages during the design of a complex system, such as GIS-T. With the data-driven approach, a complete view of the system exists, although there is a possibility not efficiently to response specific application requirements of single users. With method-driven approach, due to consideration of single repositories separately, application requirements are identified in detail. However, there is a risk of not recognizing the data exchange or common activities, which exist due to not taking a complete view of the system.

In general when designing complex information systems, for example an entire highway agency information system, a joint data-method driven method is preferable in order to benefit from both approaches and to close the gaps of each. After feasibility studies have been carried out and the decision has been taken to implement the system, according to a joint data-method approach data structure and business data analysis are carried out separately. In order to construct the optimal conceptual data model, overlapping time intervals should be taken. This provides a specialized view of reality, and actually helps in the understanding of the users point of view. During the study this goal was achieved by surveying user requirements and analyzing existing models as well as existing systems.

Following business and data structure analysis, the conceptual data model should be defined. It is a formal way of describing an abstraction of real world phenomena for a specialized view of reality, which reflects decisions about features and their relationships. Without describing these different specialized views of reality, system success could not be guaranteed. Because of this, information design is not a one-way processes. Generally, designing a successful system requires repetitive cycles. Although the conceptual data model is the core of GIS system design, lack of the data modeling issues result many problems that are faced today in GIS. Such problems mainly concludes insufficient user requirement responses, unpredicted data integration problems and finally not efficient usage of the system. Because of above mentioned reasons this study is mainly considered a progressive approach to the conceptual data modeling requirements of an entire highway agency in order to accomplish such problems and increase the efficiency.

After completing the conceptual data model, a logical design should be made which is the description of the database structure in a formal language. The conceptual data model should be converted into one of the database structures, which can be hierarchical, network, relational, object-oriented or object-relational. During this study object-relational logical design approach is preferred, due its benefits during GIS data modeling that are highlighted in further sections.

Physical database design expresses the implementation of the developed concepts for a selected database, mainly with respect to the description of the storage structures and data access methods. Validation and implementation of the designed system can be conducted in parallel.

For data maintenance, further extensions to the system or data integration the above mentioned steps should be documented in a formal language. An overview of the data modeling steps can be seen in Figure 2-1.

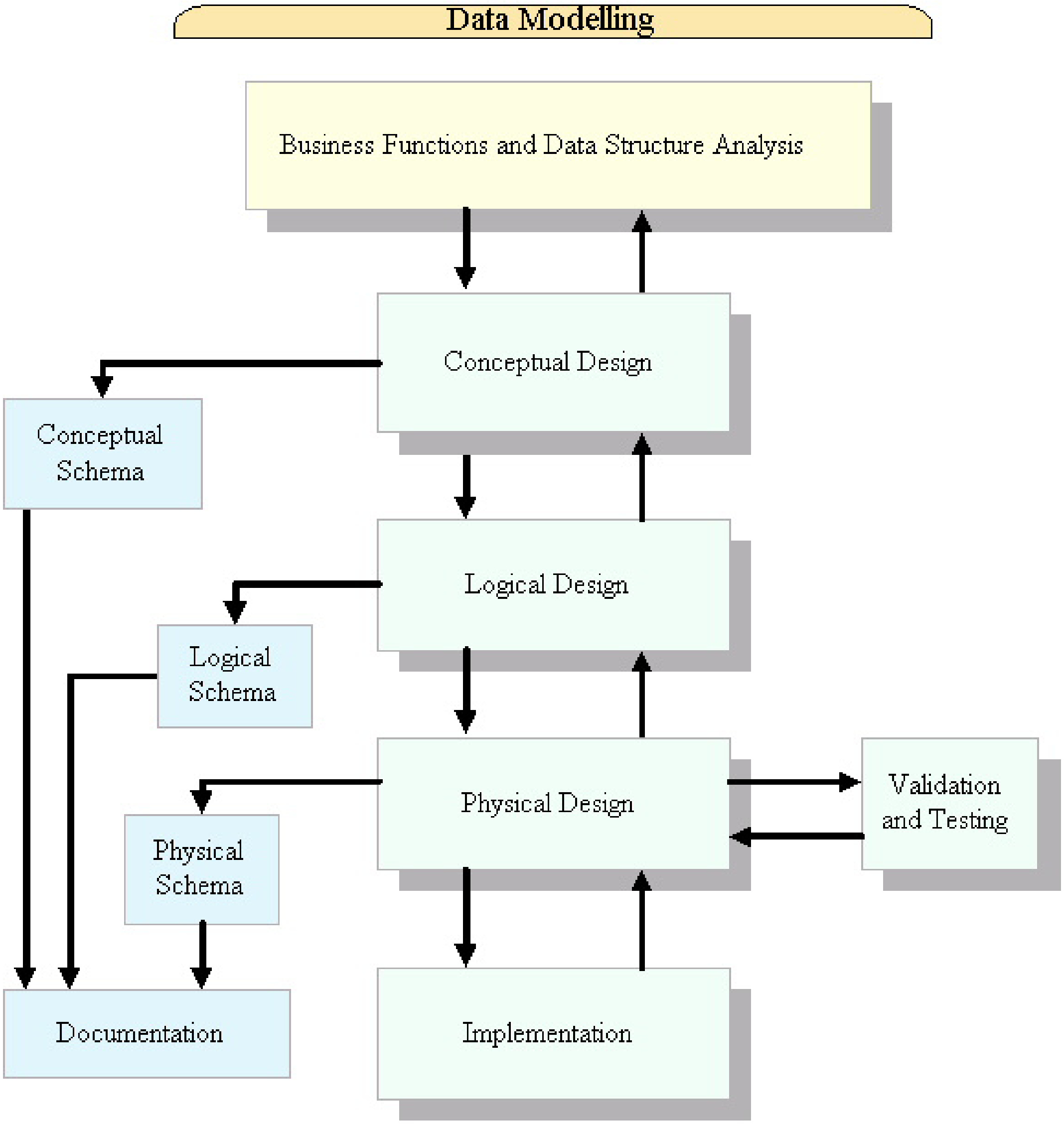


Figure 2-1: Data Modeling Steps

2.2 Information Structure Analysis

In order to analyze the data structure and to map specialized views of reality as mentioned in the first step of Figure 2-1, fundamental concepts for the analysis of spatial data should be clarified. Analyzing the data structure refers to all types of data used by an enterprise, where the emphasis is on which things are of interest and descriptions of the relationships between them. A real world phenomenon is specified in three main information categories as spatial properties, non-spatial properties and role, behavior or method.

The first two components, being, spatial and non-spatial information, will be examined in this chapter. The third component; role, behavior and method of road information, is studied during the conceptual data modeling.

2.2.1 Spatial Information – Geometry and Topology

Spatial information can also be sub-divided into two main components, geometry and topology, although these components are usually indistinctly handled in GIS. One of the fundamental concepts for analysis of spatial data is a formal understanding of geometry and geometrical relationships between objects.

Geometry is described as " the study of figures in a space of a given number of dimensions and of a given type "*[WEISSTEIN,1999]* or " the study of invariant properties of given elements under specified groups of transformations " *[JAMES,1976]*. Therefore, geometry can be also be separated into two categories; datum dependent and datum independent. The datum dependent indicates "given number of dimensions" in the definition, specifying that element is referenced on spatial dimensions. The datum independent indicates "the given type or elements"; that elements are defined using a parameterized approach. Examples of these categories are;

▪ Datum dependent:	Point (X, Y, Z) coordinates referenced to coordinate system, in this case geodetic datum dependent.
▪ Datum independent:	Straight line described by its parameter length.

Topology arises from the geometry by generalization, which means in this case by abstraction of the metric information. This describes an explicit knowledge of the mutual connectivity such as; order, connectivity and adjacency. Road networks are generally represented as graphs in which intersections are nodes and roads are links, although some disadvantages arise in the case of GIS-T, because of the complexity of real world objects which will be examined in Section 4.2.3.2.

Generally graph, which is represented by a diagram, is defined in mathematics as follows: *[BALAKRISHNAN ,2000]*

> *"A graph is an ordered triple $G = (V(G), E(G), I_G)$, where $V(G)$ is a non-empty set, $E(G)$ is a set disjoint from $V(G)$, and I_G is an "incidence" map that associates with each element of $E(G)$ an unordered pair of elements (identical or distinct) of $V(G)$. The nodes of G, and elements of $E(G)$ are called the links of G. If, for the edge e of G, $I_G(e) = \{u, v\}$, then $I_G(e) = uv$ "*

Example: Figure 2-2 shows a graph description.

If $V(G) = \{v_1, v_2, v_3\}$, $E(G) = \{e_1, e_2, e_3\}$ and I_G is given by

$I_G(e_1) = \{v_1, v_2\}$, $I_G(e_2) = \{v_2, v_3\}$, $I_G(e_3) = \{v_2, v_3\}$, representing the connectivity.

Therefore, $(V(G), E(G), I_G)$ is a graph.

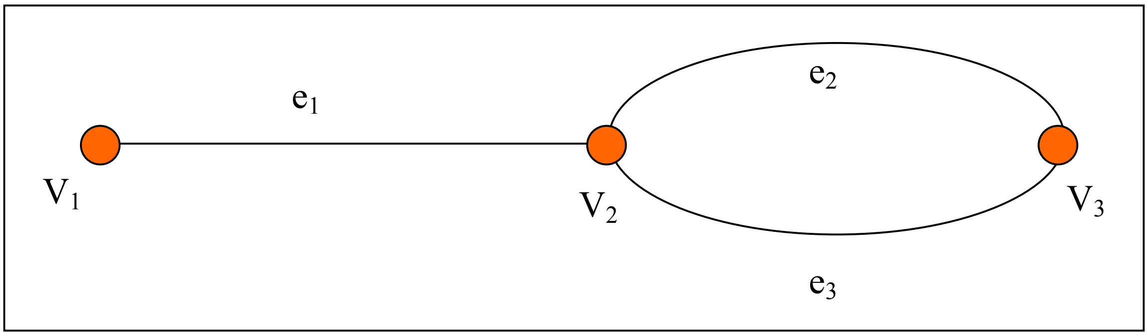


Figure 2-2: Graph Description

The topological relationships between the objects are invariant with respect to their position, orientation, transformation, shape and size. Figure 2-3 illustrates the spatial components of a road network which has different geometrical information, but identical topological information.

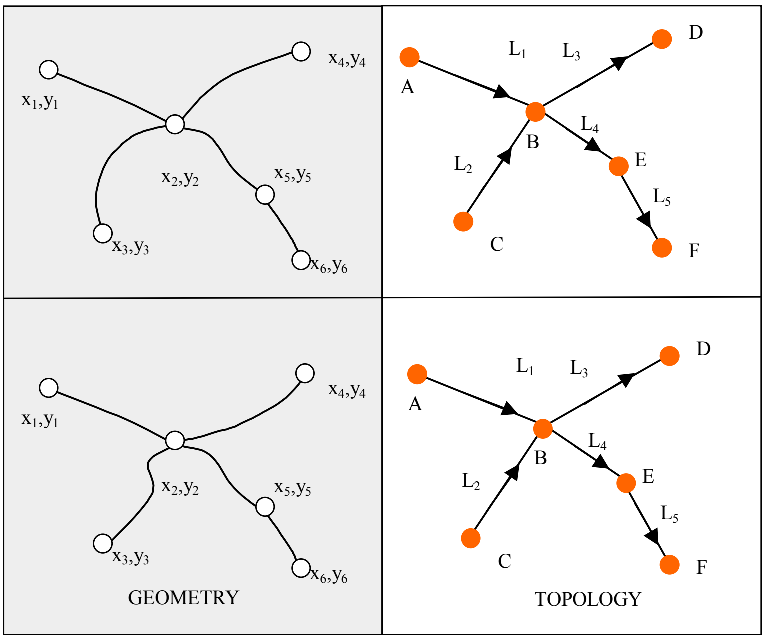


Figure 2-3: Spatial Components of a Road Network

2.2.2 Non-spatial Information

Non-spatial information, also called thematic information, is defined as descriptive information of a real world object, which does not contain any spatial character. In the case of highway administrations the road type, pavement type, accident statistics and capacity analysis information can be given as examples of thematic information. In highway agencies, this information is generally documented in analogue semantic sketches, referencing road information on a linear system. As this information is referenced on a linear system, it has also spatial components. However, in GIS-T, spatial components of this information is not considered. In Section 4.3.1, the problems encountered due to not considering this spatial component, and solution proposals are discussed. An example of linearly referenced thematic road information is provided in Figure 2-4.

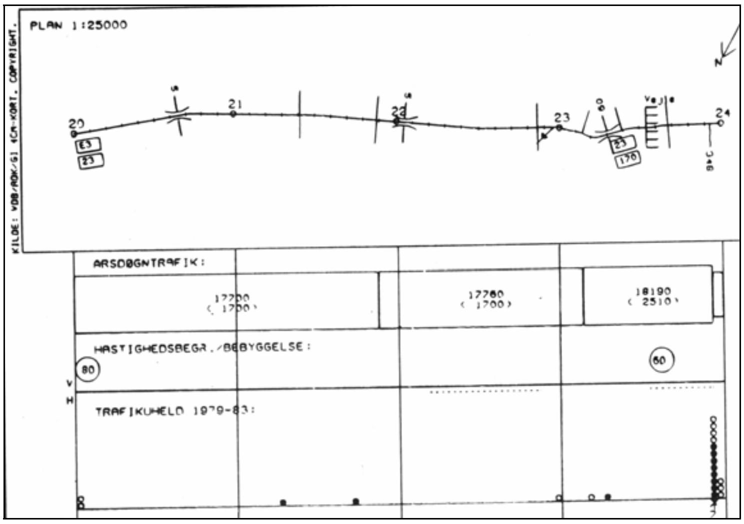


Figure 2-4: Thematic Road Information

2.3 Conceptual Data Modeling

A conceptual data model, is an unique central description of real world phenomena. In the conceptual data model real world phenomena is designed to combine various subsets of the reality, which describes relevant information and methods in the user perspective, and information that may be held in a database. The success of GIS is highly dependent on information structure analysis and conceptual data modeling. Geographic Information Systems are composed of hardware, software and data. Of these components, data, including the data model, has the longest life span and is the most costly. However, the data component has never had the first priority during GIS establishment, especially compared to the software component.

The basis of a data oriented approach to database design was presented in the ANSI/X3/SPARC (1975) report. The report described a three-schema architecture, illustrated in Figure 2-5. It identified the following three schemas: external, conceptual and internal *[ROLLAND,1992]*. Briefly;

- The external schema describes the varying views of real world phenomena according to users
- The conceptual schema formally describes the conceptual data model
- The internal schema describes the physical storage structure of conceptual schema that may be required at any given time.

The external schema is constituted after completing data structure analyses and identifying business methods. For highway administrations the most difficult aspect of database design is the definition of the external schema, due to the differing abstraction levels and scales of user's expectations and requirements. Some highway administration applications can require road network information for whole countries, while others, such as tunnel design are only related to a specific road section.

Three approaches exists in order to perform a conceptual data model; being top-down, bottom-up and integrated. Primitives are objectified in two groups, top-down and bottom-up. Top-down primitives correspond to pure refinements which apply to a single concept, the starting schema, and produce a more detailed description of that concept; the resulting schema *[LAURINI, 1994]*. The bottom-up approach begins with low-level programs and develops the system to a high level gradually, successively integrating the user's need's as much as possible. There is always a risk of the top-down approach missing important details, which should be included in the conceptual data model, and of the bottom-up approach losing the wider view. A good methodology for conceptual design should ideally be a compromise between the two contrasting approaches *[BATINI ,1992]*.

For highway agencies the data modeling process is mostly realized top-down. Since the conceptual data model is designed for an entire agency and highway administrations are hierarchical organizations, using a top-down approach permits a more comprehensive and generalized data model. However, during this study it is realized that, in order to fulfil some specific requirements such as cross-sectional design information, the bottom-up approach is also required due to perform better information data structure analysis.

The conceptual schema typically includes conceptual entity objects, relationships, attributes and methods which express the system behavior required by the user's community. The conceptual schema should be free from the physical structure of the database. This means it should be independent of the software and object storage techniques. This makes possible a change at the physical data level without involving any modification of the conceptual schema. Three important characteristics of a conceptual schema, which must be satisfied, are:

- Consistency with the business infrastructure and be valid across all application areas.
- Extensible, such that new data can be defined without altering previously defined data.
- Transformable to both the required user views and to a variety of databases and system architectures.

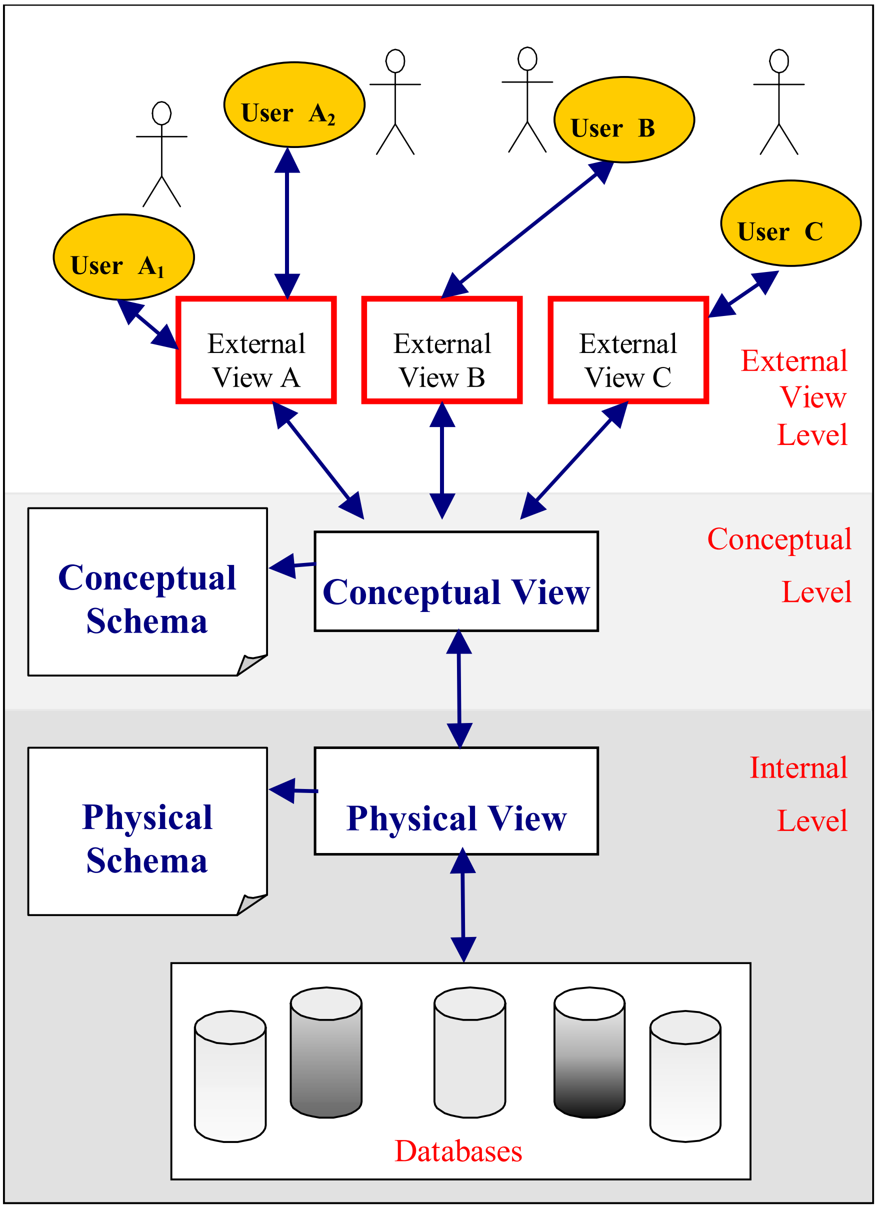


Figure 2-5: Three-Schema Architecture

2.3.1 Methods of Conceptual Data Modeling

Many methodologies exist for conceptual data modeling including semantic, functional, Entity-Relationship (ER) and object-oriented. During this study, the two main methods will be considered, namely ER and object-oriented. Currently both two approaches are highly used during conceptual data modeling. Conceptual modeling in relational database design often makes use of a formal approach known as ER modeling, first represented comprehensively in 1976 (Chen,1976) *[LAURINI, 1994]*. With the advent of object-oriented technology, an other modeling technique, object-oriented approach has emerged.

2.3.1.1 Entity- Relationship Approach

The Entity – Relationship (ER) approach adopts the view that the real world consists of entities and relationships between them which are characterized by properties.

In this formalism the basic components are:

- Entities, which are defined as clearly distinguishable real world phenomena and atomic.
- Relationships between entities are defined as associations between two or more entities.
- Attributes for both entities and relationships, where attributes are properties of these with specific meaning with respect to the conceptual data model.
- Cardinalities describing possible relationships for each participating entity.
- Integrity constraints, which are functional relationships between entities.

ER generally contains groups of entity types, which are high-level classifications of a major topic of interest. It is formally defined, using a very expressive language. ER data modeling approach is well known, easy to read and wide-spread. However, ER does not provide adequate concepts for representing methods designed in the conceptual data model. Additionally due to required atomic entities, expressing composite data structures it is difficult.

2.3.1.2 Object-Oriented Approach

The Object-Oriented approach is a method of design encompassing the process of object-oriented decomposition and a notation for depicting both logical and physical as well as static and dynamic models of the system under design *[OMG, 1999]*.The formal language of the object-oriented approach is the Unified Modeling Language (UML), a general-purpose notational language for specifying and visualizing complex, object-oriented software or projects.

The four major elements of this approach are abstraction, encapsulation, modularity and hierarchy. These concepts will be examined in Section 2.4.2. An object is an entity with a well-defined boundary and identity that encapsulates state and behavior. State is represented by attributes and relationships, behavior is represented by operations and methods. Relationships in UML can be classified as dependency, association, generalization and realization *[BOOCH,1994]*.

With object-oriented technology entities and relationships are still used, generally with an extended notation. The main advantage of object-oriented approach is ability to define and represent methods of objects, which will save many development and maintenance efforts and provide clear understanding of the conceptual data model. Additionally the object-oriented approach allows objects having complex structure and therefore model the reality in a better way, whereas with ER the entity is atomic. However, a disadvantage of object-oriented approach could appear because of complex structure concept, since the modeled complex objects representation must satisfy the required detail level for further purposes such as data maintenance.

As an example, geometrical element point and topological element node entities and their relationship with respect to each other are presented. These entities are purposely chosen as atomic, in order to highlight differences between the two approaches.

Nodes are generalizations of points. Therefore, they cannot be represented without points. In order to simplify the example, it is assumed that the point entity attributes are the point identifier and (X,Y,Z) coordinates, and the node attributes are the node identifier and node number provided by the agency. There is **0..1:1** relationship between node and point defining that every node should be assigned to a point, but a point may or may not be assigned to a node. In Figure 2-6 these entities, their attributes and relationships are illustrated according to both the ER and object-oriented approaches

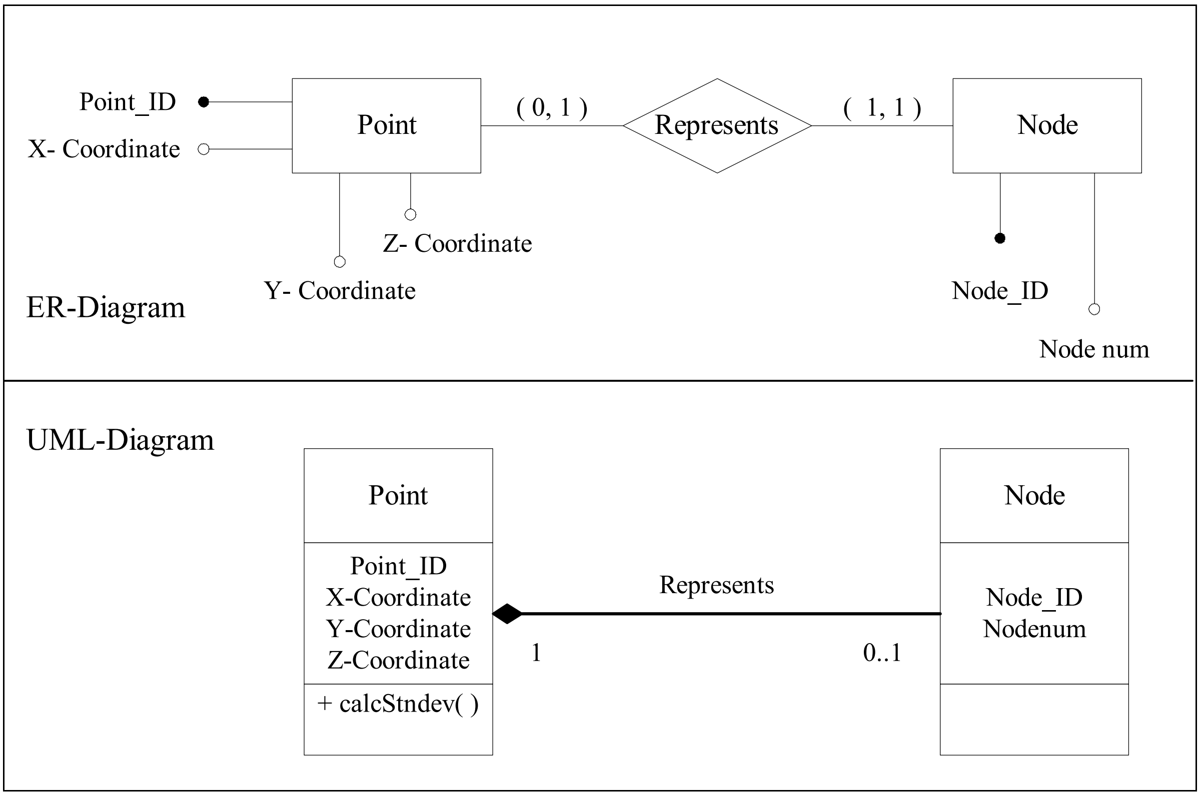


Figure 2-6: Conceptual Data Model Approaches

Between these two very common approaches the advantages of UML can be easily recognized. The conceptual data model documentation is clearer through definition of methods such as calcStndev( ) which stands for calculating standard deviation values of (X, Y, Z) coordinates. In addition the UML notation is more expressive. For example the relation between a point and a node can be presented as aggregation which also provides the information that **Node** cannot be represented without **Point**.

2.4 Logical Data Modeling

The goal of logical design is to translate the conceptual data schema into a logical schema tailored to the specified database's management system. A logical schema is a description of the structure of the database that can be processed by the database management software.

Until recent past two decades, the traditional approach to information system design mainly concentrated on conventional systems, including the flat file, the hierarchical and the network data model. In GIS-T point of view, these are important, since existing data sources are still partly stored in these systems in highway agencies.

According to information technology historical developments, firstly flat file approach was promoted. The flat file approach is based on application programs, where each program defines and manages its own data. All data could be entered into one large table or a flat file. They were often developed individually to meet the requirements set by a particular department of the organization. With its simple data structure and rapid data access, the flat file approach was highly used. A tabular model clearly allows association of entity instance attributes, but is not effective for different levels of aggregation or for complex situations with many entity types. Consequently some of the data in the file was either duplicated (data redundancy) or inconsistent, since no coordination existed between files belonging to different groups of people. Data sharing was limited and there was no enforcement of standards, requiring extremely careful data maintenance.

Due to above described deficits, database systems had been promoted. Database systems can be objectified according to the following types:

- Hierarchical
- Network
- Relational
- Object-oriented data model
- Object-relational data model

The hierarchical model has two basic data structuring types; records and parent-child relationships. Record types describe the structure of a group of records. They were stored in a general tree structure with one record type, which has zero or more dependent record types. Hierarchical data models were widely used between 1960 and 1980 due to the widespread use of IBM's Information Management System (IMS), so it is possible to come across hierarchical databases or applications, including GIS, also today in highway administrations. The advantages of hierarchical models were easy usage and high speed of data access. However, many-to-many relationships can not be realized with the hierarchical data model because the parent and child object structure enforce one-to-many relationships. A child object can not exist without a parent object. Additionally every new relation in records resulted redundancies in the system. Difficulties appeared especially during modeling many to many relations and non-hierarchical structures.

Problems of flat file approach, progress the development of the network data model, also known as the CODASYL model, in parallel with the hierarchical approach. The network data model has two basic structuring type, record and set. While record types are defined similarly to the hierarchical data model, set types define one-to-many relationship between record types. Network data models support many-to-many relations. Redundancies in the system was highly decreased. However, updates were limited. Additionally it was not feasible to solve spatial queries. The network data model has limited flexibility for changing data and access requirements, which is necessary for GIS.

These systems were highly desired compared with relational database initially, because of their high performance. However, the limited data exchange, unsatisfactory access requirements and expensive maintenance, promoted the development of the relational data model.

2.4.1 Relational Data Model

The relational data model is based on mathematical relations, where data is logically structured in tables. The relational model was first defined in 1970, when E.F. Codd introduced the idea of using the mathematical concept of relations (in the set theory) as the means to the data model. *[PAPAZOGLOU, 1989]*

Principally, set theory and predicate logic is used. Supposing two sets, S_1 and S_2, where $S_1 = \{x_1, y_1\}$ and $S_2 = \{x_2, y_2, z_2\}$, the cartesian product is the set of all ordered pairs.

$$S_1 \times S_2 = \{ (x_1, x_2), (x_1, y_2), (x_1, z_2), (y_1, x_2), (y_1, y_2), (y_1, z_2) \} \tag{2.1}$$

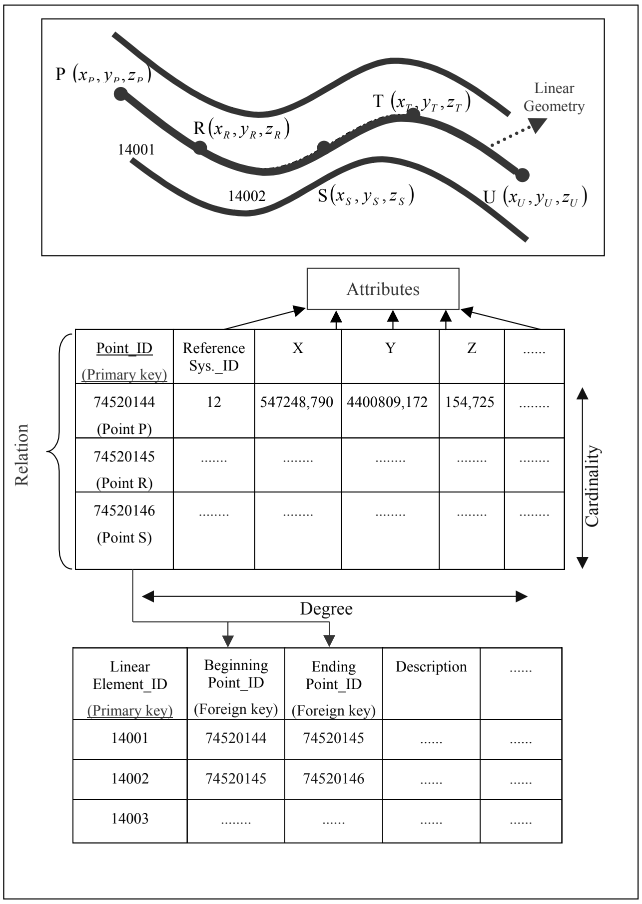
Any subset of this cartesian product is a relation. In order to define a general relation between n domains S_1, S_2, S_3,..... S_n, the cartesian product is defined as

$$S_1 \times S_2 \times S_3 \times \ldots \times S_n = \{(s_1, s_2, \ldots s_n) \mid s_1 \in S_1, s_2 \in S_2, \ldots, s_n \in S_n\} \qquad (2.2)$$

The Entity-Relationship (E-R) approach, described in Section 2.3.1.1 is the basic approach used in the relational databases. Entities and relationships among them are stored in tables. They are tabulated into rows and columns. With the relational data model, each element in the n-tuple consists of an attribute-value pair ,where tuple is a row of a relation. An attribute is a named column of a relation. Degree of a relation is the number of attributes it contains. Cardinality is the number of tuples it contains. Relations between entities are implemented through foreign keys. The foreign key is a code, which is stored in a table and refers to rows in another table. The primary key of the linked table is stored in the foreign code column of the other table. Relational databases can also be defined as normalized relations. Normal forms are guidelines for relational database design that increase the consistency of data. In the relational database systems methods and integrity constrains, which are defined in the conceptual data model are realized through external transactions.

An example, in order to clarify the relational database terms is provided in Table 2-1, showing the relationship between point and linear elements. Two **1: 0..N** relationships can be defined between **Point** and **Linear Element**, where every linear element must have at least one beginning point and one ending point, and every point may be assigned to none or many linear elements. In order to realize this, the primary key of the point, named point_ID, is used as the foreign key in the linear element table.

Table 2-1: Basic Elements of the Relational Data Model



Point_ID (Primary key)	Reference Sys._ID	X	Y	Z	......
74520144 (Point P)	12	547248,790	4400809,172	154,725	....
74520145 (Point R)	....	....	....	....	....
74520146 (Point S)	....	....	....	....	....

Linear Element_ID (Primary key)	Beginning Point_ID (Foreign key)	Ending Point_ID (Foreign key)	Description	......
14001	74520144	74520145	......	......
14002	74520145	74520146	......	......
14003	....	......	......	......

The relational database model presents operations on algebraic expression language in order to realize data query and manipulation. With data manipulation languages the types of operations allowed on the data is defined through relational algebra. It was originally proposed eight operations in relational algebra, but several others have been developed. These operations are selection, projection, cartesian product, union and set difference. In addition, there are also join, intersection and division operations, which can be expressed in terms of the five basic operations *[CONNOLLY,1999]*. These operations are also the basis for other data manipulation languages, where the Structured Query Language (SQL) is the international standard query language of relational databases.

2.4.2 Object Oriented Data Model

The object model captures the static structure of a system by showing the objects in the system, relationships between the objects, and the attributes and operations that characterize each object.*[RUMBAUGH, 1991]* Communication between objects is realized using a message passing system. A message is a request sent from one object to another in order to execute one of the objects' methods.

One of the advantages of an object oriented data management system is that real-world phenomena can be modeled closer to reality with non-atomic objects. Object models could be structured around conceptual objects rather than geometrical properties. With the object-oriented approach, geometry can be modeled like other information. Explicit stored spatial information, which is required in GIS, can be realized without redundancy. For example geographical objects can be defined using methods, which can automatically generate these with the use of their parameters. In addition it is possible to generate user-defined types without being limited to vendor specific data types.

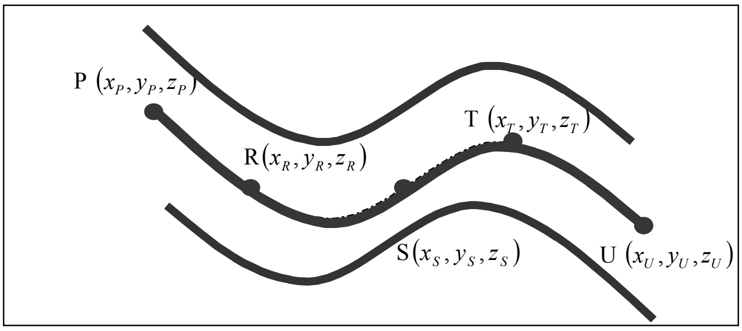
Another advantage is the encapsulation concept. This permits possibilities such as;

- Integration of different abstraction levels for one object.
- Designing spatial, non-spatial and methods in one object.
- Expressing business rules as methods belonging to the object.

In Table 2-2 a road segment is presented as an example to illustrate the object-oriented concepts. Spatial information, non-spatial information and business rules are encapsulated in one road segment object, preserving relationships between each other. Spatial data for the provided example includes; coordinates of the middle axis, reference coordinate system, parameters, spatial accuracy obtained via defined methods and topological elements. Thematic data belonging to the segment is represented by road type. Finally, business methods are introduced such as distance( ) and calalignment( ) which is defined for the calculation of distance of the road segment and the alignment parameters.

With reference to the relational database example in Section 2.4.1, as this example is not an atomic object, some additional information is required for this example. However, with additional tables and their assigned relationships, it is possible to map the same information using relational databases. Additionally methods and integrity rules needed to be provided in the relational database in order to acquire the same result and to maintain the data integrity.

Table 2-2: The Object Road Segment



Road Segment								
RoadSeg_ID (Object Identifier)	Link _ID	Measured Length	Road Type	BPoint _ID	EPoint _ID	LinEl _ID	LinEl Type	…
10015	5001	42,521	08	74520144	74520145	14001	line	…
				74520145	74520146	14002	arc	…
…	…	…		…	…	…	…	…

+distance( )
+calalignment( )
+accuracy ( )
+Move ( )
+Animate ( )

2.4.3 Comparison of Relational and Object-Oriented Approaches

In order to find the appropriate approach for designing GIS-T conceptual data models, the advantages and disadvantages of both approaches needed to be highlighted.

One of the benefits of relational database technology considering GIS-T is that; it offers solutions to the issues of security, versioning and referential integrity. In addition, it is mature and available across a wider variety of platforms. Established standards such as the Structured Query Language (SQL) for querying large databases are available. The main drawback lies in not providing adequate facilities for specifying constraints on the data. Additionally, relational data models are built upon elementary elements, with only atomic features being permitted, which leads some limitations during complex features. In the areas of data semantics, model extensions, object identity and programming interface, weak points of the relational model can be found.

Compared to the relational approach, object-oriented methods provide a model for integrating data with business rules. To achieve the same semantics, relational database management systems usually require complex control methods, generated through a combination of third- and fourth-generation languages. *[BOOCH,1994]* With object oriented modeling complex objects can be generated. Concepts of abstraction, user defined data types, encapsulation of object properties with business rules are the main benefits introduced by means of the object-oriented approach. In addition, it provides a more semantic substance by allowing the user to explicitly specify constraints on the data. However, the object-oriented modeling approach has several disadvantages. The object oriented languages run slower than procedural ones due to message expression, degrading query performance. Additionally, very few GIS-software vendors have been successful in using an object-oriented database for storing and retrieving spatial data. Object oriented databases do not provide a standard query language. Another bottleneck is that the amount of data in the GIS database component is very large compared with other systems, which makes the performance considerably low.

A comparison between the two approaches with respect to GIS is provided in Table 2-3.

Table 2-3: Comparison of Relational and Object-Oriented Approaches

Relational Approach	Object – Oriented Approach
Advantages: ▪ Standard query language, SQL ▪ Enhanced spatial query opportunities ▪ Versioning ▪ Wide-spread, mature ▪ Security	Advantages: ▪ Integration of business rules with spatial information ▪ User defined types and functions ▪ Enhanced abstraction concept ▪ Simplicity in interfacing to other sources ▪ Providing solutions for generalization problems
Disadvantages: ▪ Weak support in the integration of business rules ▪ Pre-defined data types ▪ Atomic entities	Disadvantages: ▪ Management weakness with large amounts of data due to messaging ▪ No standards for query languages ▪ Not widespread, many concepts are still at a test stage

In GIS, it is necessary to be able to use enhanced spatial query opportunities, versioning of the relational concept. Additionally, existing digital information is generally stored in relational databases. Contrarily, the object-oriented approach propose solutions to problematic areas of GIS such as multi-dimensionality, various abstraction levels and integration of business methods. When both approaches are evaluated in the GIS context with respect to their respective advantages and disadvantages, an hybrid approach is emerged, namely the object-relational approach.

2.4.4 Object-Relational Data Model

The object-relational data model combines the concepts of objects and methods from the object-oriented model with the concept of relations from the entity-relationship model. The object-relational database management system is an extended relational database supporting abstract data types, procedures, encapsulation and complex objects of the object-oriented concept. Limited operations of relational databases can be extended, defining new operations and methods. The advantages of relational databases such as standardized query languages, security, versioning and referential integrity facilities can still be used.

The object-relational data model shows its strength in performing queries of complex structured data. Since with the object-relational data model, it is possible to define complex structured data in the same manner as relational database. The object-oriented messaging approach is not required between objects, which causes lack of performance. Additionally, the enhanced spatial query possibilities of relational databases are easily performed, since the complex objects are stored in tables associated with their object identifiers.

The main requirements of GIS such as searching capabilities, multi-user support, handling complex data and extensibility of systems are best satisfied with object-relational database management systems. Stonebraker proposed a classification between database approaches with a four-quadrant view as illustrated in Figure 2-7. *[CONNOLLY,1999]*

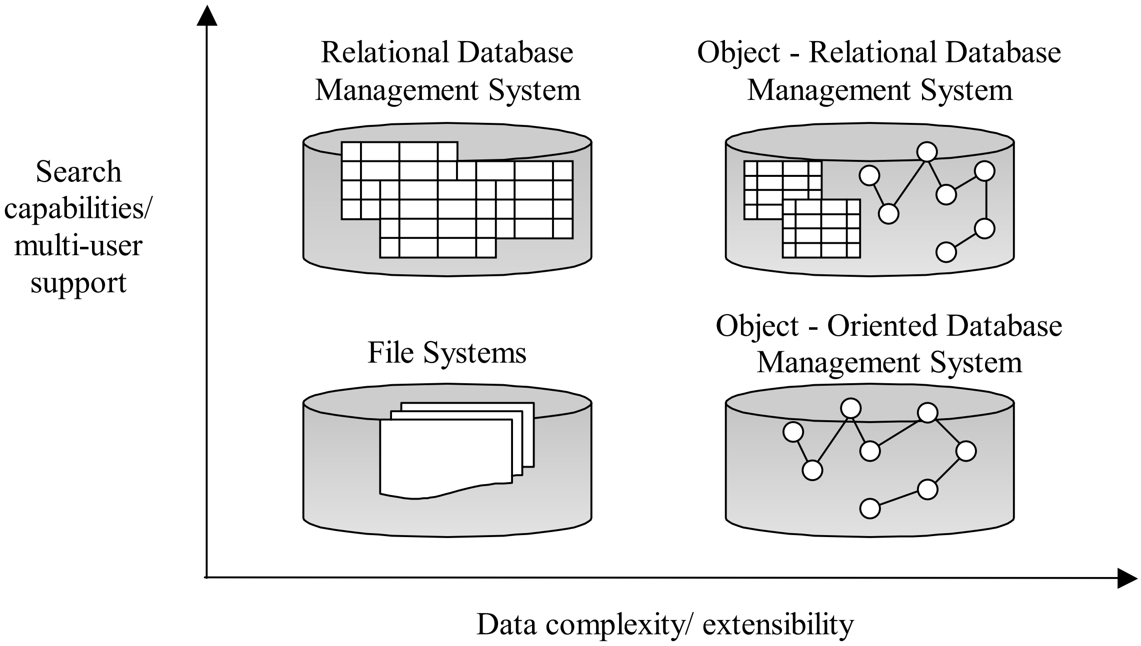


Figure 2-7: Classification of Database Management Systems(DBMS) [CONNOLLY,1999]

During this study, since in highway administrations large amount of information are mainly stored in relational data bases and the object-oriented concepts provide solutions to identified problematic areas, the object-relational data model was considered as the most suitable approach and applied.

Chapter 3 : Highway Administration and GIS

3.1 Overview

Highway administrations, being one of the major governmental organizations, provide services and coordination at a national and international level to ensure the mobility of people and goods. Examples of highway administration tasks are:

- Planning new constructions
- Maintenance of existing roads
- Traffic information
- Regulations about road usage and safety
- Installing facilities on highways
- Conducting projects
- Program evaluation

In order to fulfill these various tasks at a technical and administrative level, information is needed from diverse sources.

In addition to the above mentioned tasks, due to an increase of environmental awareness, other tasks have arisen such as:

1. Protection of environment, referencing existing road net information during the planning phase.
2. Rising qualitative requirements in terms of the traffic route network, e.g. increase of road safety and separation of the traffic method.
3. Coordination of road, rail and waterways, which can only be achieved through new technologies and uniform networks.
4. Evaluation of:
    a. the improvement or adjustment of design elements in planar and vertical positions.
    b. cross-section and node design is only possible through new standard software and the development of new conceptions.*[BMVBW,1998]*

Technological developments, especially GIS technology, are considered to be inevitable by highway agencies for the realization of their tasks, since GIS provides rapid and accurate decision support tools in a more integrated and economical manner. GIS capabilities, which have major significance to highway administrations, are *[ALTAN, 1996]*:

- Ability to share spatial information, which often leads to better cooperation and decreases inconsistency and redundancy.
- Ability to integrate loosely related data, with respect to usage of spatial references, which can lead to discovering new properties of data and cooperation at many levels.
- The ability to aggregate data into larger geographic units that are more appropriate for the large scale applications, which can lead to far better understanding, cooperation and management of the operational units within an organization.

Due to the necessity of close coordination between departments and divisions, the efficiency of GIS-T technology in highway administrations, can only be achieved using an integrated approach. Additionally, the integrated approach has many economic benefits such as the minimization of data collection.

3.2 Analyzing Highway Administrations

In this chapter, an evaluation of the current GIS-T situation within highway agencies is presented for four countries in order to achieve a wider perspective through this study. Concentration is mainly given to the following;

- Organization structure.
- User requirements.
- Data acquisition techniques.
- System architectures.
- Existing information systems and the conceptual data models.

The usage of GIS-T technology differs between the mentioned countries, which helps in concentrating on more generalized problems of highway agencies. Some statistical information on the countries evaluated is provided in Table 3-1:

Table 3-1: Statistical Information on the Countries Studied

Country	Area (km^2)	Road Nets(km)				Number of Motor Vehicles.
		Motor-ways	National	Regional	Total	
Denmark	43.094	880	3.690	7.090	71.600	2.040.000
Germany	357.022	11.300	41.600	75.800	633.000	43.350.974
Turkey	774.815	1.405	31.412	28.813	381.631	4.327.885
USA	9.363.520	88.400	727.000	694.000	6.420.000	203.659.000

3.2.1 Organization Structure

The organizational structure of the highway administrations studied is grouped into two main categories according to governmental type: federal and centralized.

The organizational structure of the German Highway Administration can be given as an example of a federal governmental type. State highway authorities are responsible for the administration of interstate roads and highways on behalf of the federal government. With respect to GIS, this will lessen the amount of data managed compared to a centralized structure.

However, problems arise with the transfer of information, standardization and data formats. This is due to there being different conceptual data models and different systems in use. This issue needs to be examined in detail, since this is one of the core problems of GIS-T and valid for both centralized and federal structures.

Highway administrations scope and business methods are same, therefore it is expected that conceptual data models are similar. However, there are huge differences between existing conceptual data models. The reason of using various conceptual data models includes;

- No formal conceptual data model has been designed, including spatial and non-spatial data. Software vendors' proprietary databases are used for storing spatial information. Other information sources, which are unstructured, are then linked as non-spatial information. Software vendors have various "black-box" data models and systems, which are not generally fully available for third-party developers. Therefore, the designed conceptual data model is unknown and dependents on selected GIS software.
- The conceptual data models are designed in order to respond various user requirements separately, such as pavement information system, tunnel information system or traffic safety information system.
- User requirements are rapidly changing due to increase of the GIS usage and technological developments.
- The identification of user requirements can be inadequate, due to complexity of data-flow and information structure.

If data integration is intended in federal structure, it is needed to integrate the various conceptual data models together. Additionally, other organizations or governmental bodies, with the task of increasing the information transfer between highway administrations, are required.

In a centralized governmental structure these problems will tend to decrease. However, some similar problems will be apparent concerning data management, integration of data and the coordination between regions and headquarters. The variety in conceptual data models are also observed. In order to give an overview of the size of highway organizations in a centralized governmental structure, the organization chart for Turkey's General Directorate of Highways (KGM) is provided in Annex A.

In this chart several departments have administrative or financial duties involving spatial information. However, these departments have a different level of interest in the spatial data. During user surveys, it has been noted that there is a necessity to structure, model and maintain this information, considering the information's level of detail and spatial characteristics. Facility management systems are recommended for the management of information such as building management, equipment stock and personnel, since for such requirements the efficiency need to be higher. The information concerning existing hardware, software and peripherals, such as plotters and scanners should also be included in facility management systems.

In contrast, GIS-T user requirements are generally concentrated more on technical aspects of road information and upper-management administrative aspects, such as alternative routes in emergency locations. The inclusion of other requirements, which are more adequate for facility management, would require additional data modeling efforts and enlarge the system with inefficient methods in order to integrate both different levels of information. These additional information and methods will definitely results to performance problems in queries and additional costs.

Since there is more interaction between the information users, the effectiveness of both GIS and facility management systems can be increased in this way. Mutual agreement on requirements will increase the information quality and reduce the costs of development and implementation.

3.2.2 Assessing the Requirements of Highway Administrations

In order to design the conceptual data model, it is necessary to identify the external views. The external view is composed of user duties, expectations from the system and information related aspects such as content, source, usage and methods. However, identifying the external views is not an very easy task, since the conceptual data model must be kept abstract in order to be able to extend it where required, and to fulfill all the diverse user requirements. The candidate objects should be identified which represent diverse user views at the required level of abstraction.

A variety of information sources, including documents describing the systems, were studied in order to determine user requirements. For example, the user requirements as defined by a special working group of the Turkish highway administration are given in Annex B. In addition a questionnaire was prepared which specifically focused on conceptual data models, existing data sources and user duties. During this study the questionnaire was used during interviews with professionals within highway administrations in Germany, Denmark and Turkey. This questionnaire can be found in Annex C.

Several requirements were mentioned during these interviews, which are of importance. These include crisis management, public travel security, emergency vehicle management**,** information exchange with other organizations, internet roadway condition map and “What if” reports. Additionally, the information requirements of international organizations, such as the European Union (EU), should be considered concerning future transport information systems in Europe. These are partially listed below *[EU-APAS,1996]*;

1. *Road traffic of passengers.*
2. *Long-distance traffic of passengers.*
3. *Traffic of goods for all modes.*
4. *Road, rail and inland waterways transport infrastructures and networks.*
5. *Improve the inter-regional flow data between countries of the EU.*
6. *Transport of dangerous goods, for all modes.*
7. *Environmental variables required by the integrated scenarios.*
8. *Qualitative variables required by European scenarios: border effect, development of information technologies, “just-in-time” techniques.*

Generally, these requirements highlight different views of reality listed such as design information, environmental impact analyses, accident data, cross-section design, road inventories and project monitoring, but they have a common denominator: road.

The road information can be categorized into several various groups including; new construction, existing data, traffic data and history, which is illustrated in Figure 3-1.Due to simplicity of data maintenance and topology analysis, in figure geometry and topology should be distinct. The main approach to separation of geometry and topology was in Section 2.2.1 defined and other requirements promoting this separation can be found in further sections of this study.

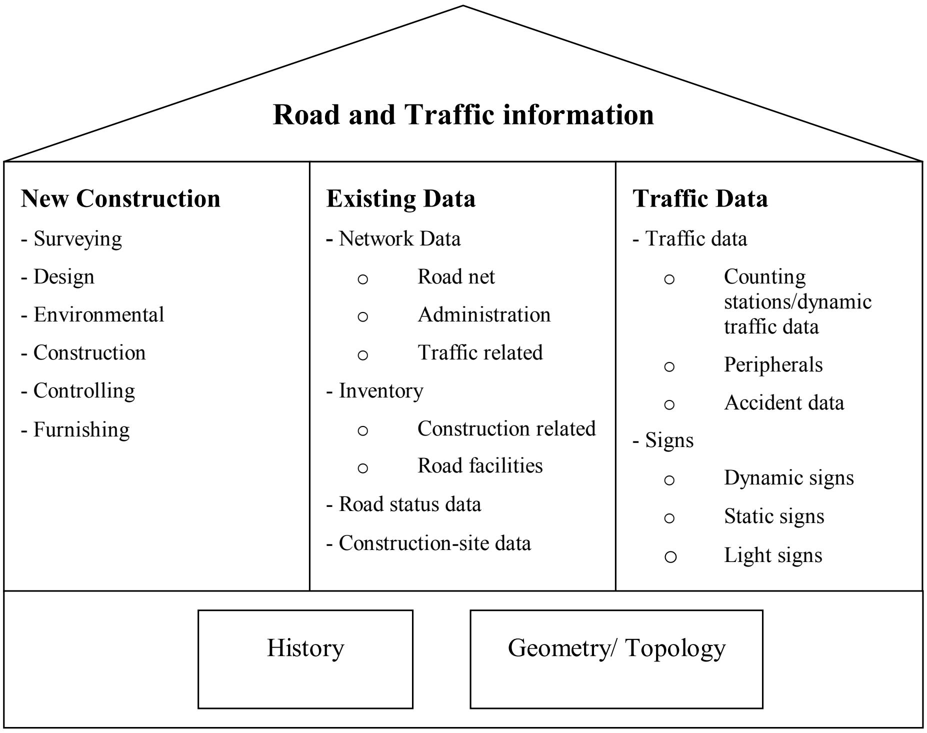


Figure 3-1: Road and Traffic Information[PORTELE,2001]

In order to incorporate all these views in the data model and to integrate them, the data structure needed to be examined. It was important to determine whether data has a spatial or thematic nature. In order to represent the variety of spatial information used in highway administrations, three application requirements of highway administrations were selected. These were; the generation of accident black spot maps, integration of inventory records into the system and the analysis of existing pavement layers. These requirements are associated with three different referencing systems illustrated in Figure 3-2.

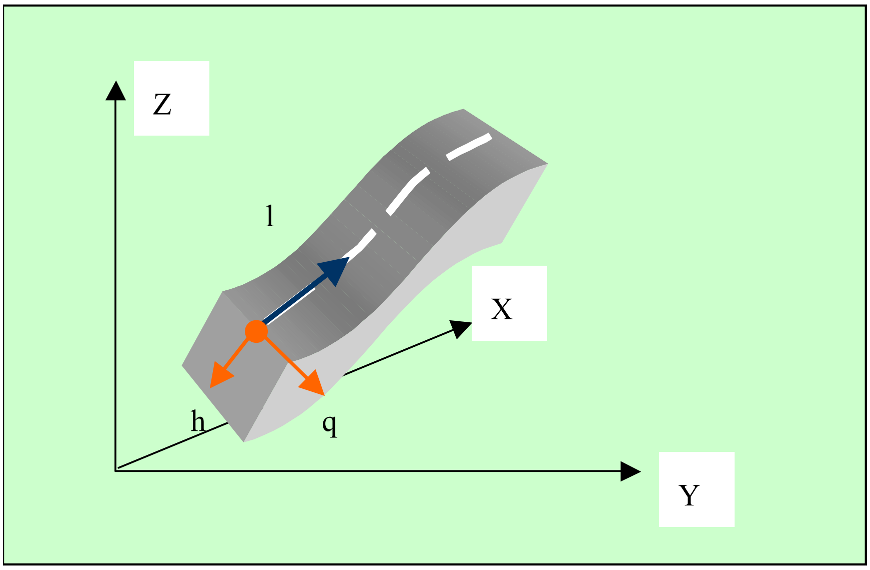


Figure 3-2: Reference Systems

Accident black spot maps are produced using three dimensional reference systems (X, Y, Z). These are generally represented as (x, y) planar orthogonal coordinate systems, although integration of third dimension will increase the efficiency of analysis.

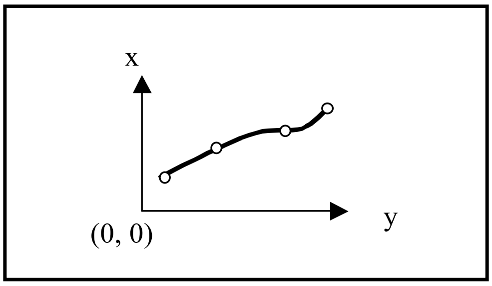


Figure 3-3: (x, y) Planar Orthogonal Coordinate System

The second example is the integration of inventory records into the system. The inventory data collected for maintenance purposes is one of the main information source. It shows the physical properties of roads and the current state. The road inventory contains information such as control section identification, road surface, intersection, cut-fill, road lighting, geometric elements, altitude, slope and critical section.

Due to the information variety, several other departments, such as planning and the traffic division also emphasized the importance of road inventory information being present in the system. Road inventory data has a one dimensional reference system, generally called linearly referenced. The linear referencing approach is used, due to its simplicity and low costs, in nearly all application areas that are based upon networks, such as infrastructure management, utilities management and hydrological analysis. Additionally, this method is the most natural method in order to collect linear objects. This methodology is shown in Figure 3-4.

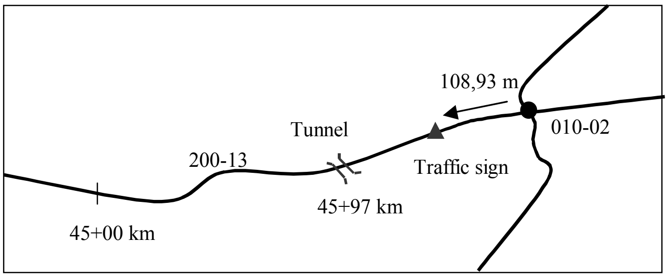


Figure 3-4: Linear Referenced Road Information

For integrating road inventory data into the system another reference system needs to be defined. We can derive from the simple linearly referenced system two planar systems namely; (l, q) and (l, h). These systems are illustrated in Figure 3-5. In both reference systems, measured length along the linear element defined by l. In the (l, q) horizontal system the q-coordinate represents distance normal to the linear element. In the basic linearly referenced system it would be regarded as being zero. With the (l, h) vertical coordinate system, h is the height of points along the route.

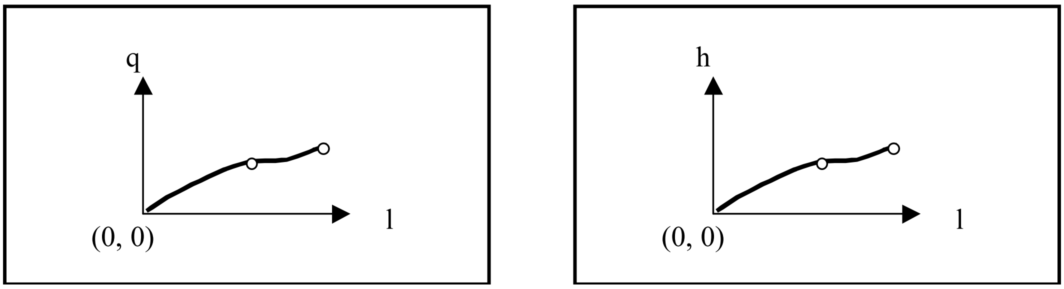


Figure 3.4.a: Horizontal System *Figure 3.4.b: Vertical System*

Figure 3-5: Linear Reference Systems

In order to analysis of existing pavement layers, due to historical convention, another reference system is used in highway administrations. In this work this system is termed as the cross-sectional reference system and represented as (h, q) reference system. The coordinate system origin is defined to lie in the middle of the road axis. This reference system is illustrated in Figure 3-6. Further details on referencing system issues will be discussed in Chapter 4.

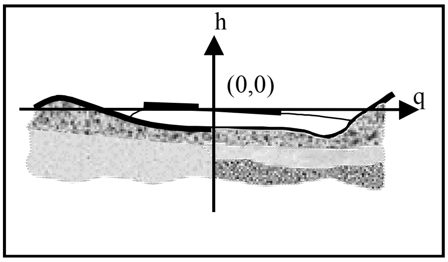


Figure 3-6: (h, q) Reference System

Several other requirements such as; project information, traffic monitoring and capacity analysis are selected, in order to present an overview of spatial data and the reference systems used by highway agencies. For these requirements spatial characteristics, the analysis cycle in the agency, data collection frequency and methods are provided below in Table 3-2.

Table 3-2: Data Requirements

User Requirements	Spatial Data		Analysis Cycle	Data Collection	
	Linear Referenced	Referenced By Coordinates		Time	Method
Project Information	☑	☐	One-time/ annual	As Built	Contract
Road and Motorway Information	☑	☑	Annual	Daily/ Monthly	Inventory
Traffic Monitoring	☑	☑	Annual	Hour/Daily	Counting Information
Maintenance Costs	☑	☐	Annual/ Multi-year	Current	Record
Accident Information	☑	☑	Annual	Take Place	Accident Report/ Inventory
Pavement Maintenance	☑	☐	Annual	Seasonal	Report
Bridge Information	☑	☐	Annual	As Built/ Monthly	Inventory
Planning Applications	☑	☑	Annual	Monthly	Report
Capacity Analysis	☑	☐	Annual	Daily/ Monthly	Record
Work Program	☑	☐	Annual/Year	Seasonal	Report

In addition to these information variety in highway organizations, there is an enormous exchange of data between departments and divisions. Since methods and definitions vary, there are many information gaps in the technical data-flow, making the analysis of information flow very difficult. The existing data in the highway administration is examined, in order to achieve better understanding of external view. In highway administrations commercial GIS and database software are in widespread use. These are generally supported by in-house and requirement customized programs.

3.2.3 Data Acquisition

In order to realize above described tasks, the first priority is given to data acquisition in highway agencies. The data acquisition is a continuos process, since various phases of road information is needed in order to fulfill the requirements. These various stages are planning, construction, current information such as traffic data and maintenance, which are shown in Figure 3-7.

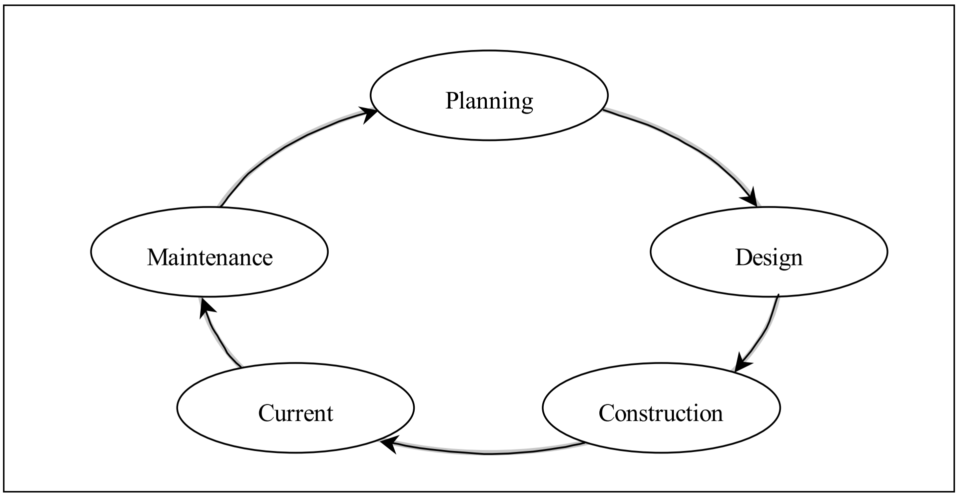


Figure 3-7: The Road Information Phases

These diverse requirements are stressing agencies to collect the required information by means of various methods, depending upon required generalization level and quality. In addition to these, collected data have generally spatial character. The spatial data acquisition requires the highest efforts and costs, due to the high level necessity of completeness, actuality, correctness and well-defined data structure *[BILL-I,1999]*. In the past few decades, conventional data acquisition approaches were mainly used in highway agencies and the information was stored in analogue format. Main reasons for this situation were, digital data acquisition techniques either were not available or economic. In addition to this, new approaches require generally a change in the traditional way of thinking. However, analogue data collection techniques are very expensive and not rapid. Additionally, the analogue information could not be easily exchanged or shared with other users, therefore efforts needed be paid only on one side.

In highway administrations, digital data requirements have increased in importance due to the needs for increased efficiency with possible low costs. Developments in information technology, especially in the areas of database and GIS-T, now provide the means to realize such optimization. In parallel to this, digital data acquisition techniques, such as Global Positioning Systems (GPS), photogrammetry and automatic counting machines, became commonly used in highway administrations, due to their data collection speed and cost reduction. Although the new techniques reduces the costs of conventional techniques, data acquisition from sketch is still not affordable for the entire highway information. As a result, although there is an increasing trend towards data acquisition with the new technologies, conventional techniques are also in usage. Consequently, in highway administrations both analogue and digital data is available.

The road inventory data can be given as an example. It is mainly acquired using vehicle odometers and documented in analogue semantic sketches. This existing information is entered manually, if it is needed to be used in a system such as; Computer Aided Design (CAD), road database or GIS-T. In addition to these conventional approach, other technological possibilities are being considered and applied successfully. In the case of Denmark the road information related to network profile, surface conditions, length and cross-section profile is collected by gyro and laser combined vehicles. The data collected with this vehicle is automatically sent via a network for further processing. It is planned to combine this system with the Global Positioning System (GPS) in the near future. This system, developed by the Danish Highway agency, is being used for roughness measurements in other countries such as Sweden and Greece. In other countries also similar integrated techniques are in usage. Especially in the USA and Germany, GPS integration was already realized. Additionally, evaluated countries are investigating systems for obtaining geometric information in an accurate manner, especially road widths and geometric elements. In particular the use of differential global positioning system (DGPS) in this field is being developed. In Germany GPS measurements have been made especially at network nodes in order to improve the node location accuracy. Also in Germany, a research group was founded with the task of studying the relationships between road alignment parameters and traffic safety *[NWR,1995]*.In particular the correlation between possible accident causes and short and long curve parameters is being considered.

Monitoring road information is another popular data source for highway administrations, especially for the determination of current road condition and in the clarification of legal issues. Video image includes time, date, speed, route number, location on the road and the direction of measurement at Danish highway administration.

Figure 3-8: Sample Video Image from Danish Highway

Another image database source is maintained by the Federal Highway Research Institute of Germany to document the direction signs of the German federal motorway network, including main lanes as well as connecting, exit and entrance ramps. Access to images is realized via a search method either through selection of the motorway number and the direction or the name of junction. After displaying an image, it is possible to maneuver by jumping from image to image. The database is marketed in CD-ROM format. It can also be considered to use this information for automatic extraction of data such as road numbers, distance measured or the location of traffic facilities.

Photogrammetric data acquisition is carried out by highway agencies, especially during the planning phase. Within the Turkish highway administration, digital terrain models (DTM) are available for areas where photogrammetric data acquisition has been realized. In Germany, experimental three-dimensional GIS applications have been made using Digital Terrain Models with 50 m resolution.

However, although these new techniques has many benefits and successfully applied, these applications are specific to requirements or applications, therefore are not fulfilling the entire highway agency requirements, especially in case of GIS-T. In order to fulfill entire highway agency requirements, the information in various sources, formats and generally in analogue form, is needed to be structured and integrated.

The main requirement of GIS-T is digital base-maps. Indirect methods, which are digitizing and scanning, are commonly used in order to acquire digital base-maps. In some cases these are provided from other organizations, such as the federal governmental land registry agencies in Germany. However, generally base-maps are digitized by the highway agencies, as is the case in Denmark. In base-map production, direct methods including the remote sensing and photogrammetric data acquisition techniques are still not widely in usage as it was expected to be. Since costs of these techniques depends highly on production scale and highway organizations requirements varies form very large scales to very small scales. However, these techniques provides more rapid and up-to-date solutions compared with the digitization, therefore it is expected that, they are more commonly for base-map production in the near future. Additionally, with the increased usage of GIS-T, other requirements such as digital divided highway information is expected to increase.

Several other requirements appear during GIS-T establishment. Other road related information, such as linearly referenced data or cross-sectional design data needed to be integrated into the system. However, the available digital information is not efficiently applicable or requires enormous efforts before it can be used in GIS-T. This is because, due to organizational separation of the tasks, similar or identical data was collected with different data descriptions, formats and various accuracy. Additionally, there is lack of common terminology and various methods are used. In spite of such complications, reusability of existing data and integration within the system is essential because of obtaining high costs of digital information. This problem promotes the usage of integrated conceptual data modeling and adjustment techniques, since the efforts required for data integration can be reduced and integration can be automated. Additionally, such methods increase the quality of available data in the system. In the further chapters of this study, various examples of such methods were presented.

3.2.4 System Architecture

The entire highway administration GIS-T design is accelerating due to the integration of methodologies, provision of common basics, decreasing data acquisition efforts and the integration of existing data. In this context, as the efficiency of the system can only be realized with up-to-date information, an appropriate system architecture should also be considered.

In highway administrations, due to distributed user locations and various interests, a client-server architecture, intranet and internet facilities needed be designed in order to realize contribution of all users. The importance of the client-server architecture and communication infrastructure have been already appreciated and realized in the examined countries. However, in practice there are various implementations of client-server architectures. One of them is storing information in distributed databases, as an example on basis of regional divisions of highway administrations, and integrating them in the main server on a regular defined time interval. Another approach is storing information on a main server. In this case, it is also possible to distribute the information on several databases but user data retrieval and queries are realized on the server, without a secondary storage on the user side. Additionally, in both approaches security rules and user rights for updating the system needed be defined and established.

The Danish highway administration provides a good example for the establishment of system architectures in highway administrations, since both approaches were applied and evaluated. In the Danish highway administration, on the basis of regional districts, initially the first approach was implemented. Every region had been stored a copy of the server information related with its own region. The maintenance of this information was under responsibility of regional divisions and within certain time intervals it was sent back to the main server in Copenhagen. The advantages of this approach is quick responses to user queries, as amount of data was reduced. Additionally, the amount of information, which needed to be maintained in the region database was extremely decreased. However, this architecture did not provide the expected efficiency. Firstly, the coordination ability of highway administration headquarter was reduced, since regions and headquarters have different versions of the same information. Secondly, the main server information was not completely updated. Thirdly, there were redundancies considering integration of various data sources, since road information is generally shared and used by different regions and headquarter, but different data sources were processed separately. Due to inherited difficulties the second approach was performed.

Within the second approach, on the basis of regional districts, clients update their particular portion of the server database. Each client has pre-defined rights and responsibilities. In order to guarantee correct system functionality, these rights and responsibilities are designed to be non-overlapping. Users are connected to a main server via network facilities. The communication between the central and local systems is ensured by Integrated Services Digital Network (ISDN). Their queries are submitted to the server where it is processed and the response returned to the client. Database update is performed by server. Fourteen federal countries are connected to the server database, and updates are realized on daily-bases. The central data processing division undertakes database integrity and general maintenance. With this approach, first approaches disadvantages are reduced to minimum. Only one version of information exists, which highly increase the coordination in the agency. Additionally, redundancy is avoided. However, performance is decreased, due to data retrieval from a very large database. The performance plays a secondary role in this case, when the advantages of this approach is considered. Additionally, the performance issue can be lessen using new technologies, such as; Broadband-ISDN(B-ISDN), intelligent infrastructure or optimizing data model. In general, the system architecture established satisfy the user requirements in topics of up-to-dateness, obtaining complete data overview.

Since only database vendors existing pre-defined data maintenance facilities are in usage, in topics of security and up-date-rights with respect to data maintenance, more attention is needed. Consistency and data integration controls are not performed within the system, especially with respect to spatial data. Considering these issues, several proposals will be made in further chapters.

3.3 Current Status of GIS-T at Highway Administrations

3.3.1 Current Status of GIS Technology

Before evaluating established systems in selected highway administrations, since some problems such as data integration issues are correlated with GIS technological development, developments and current status of GIS technology needed to be discussed.

With the development of spatial theory and computer technology, in order to analyze data collected by Canadian Land inventory data and to produce statistics for land management, in the mid 60's the Canada Geographic Information System had been established. With the speed of computer technology today, IV Generation GIS offering many benefits especially in the area of worldwide web technologies, data modeling techniques, with their client-server architectures. The main concern of this study, the data modeling techniques, had been enormously developed from geometrical data modeling into object oriented (OO), recently into object-relational(OR) data models. Development of GIS technology is illustrated in the Figure 3-9.

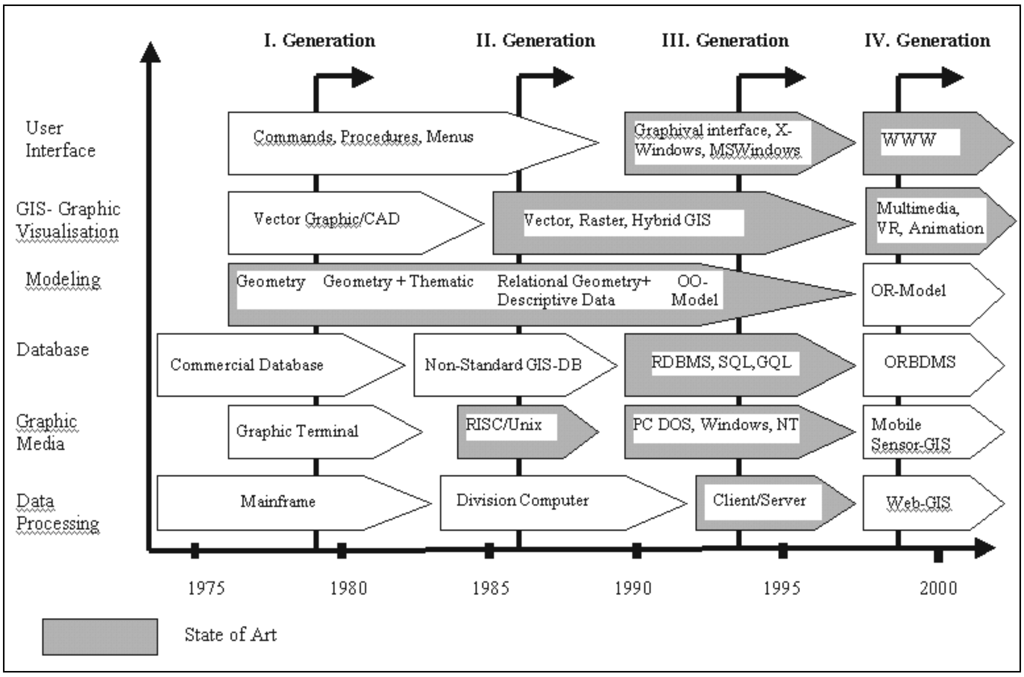


Figure 3-9: Status of GIS [SCHILCHER,2000]

Although it is not possible to classify approaches used data modeling definitely in GIS, the file processing, dual database and integrated database approaches are the main groups. With the file process approach, data files are stored in various file and many additional programs are provided in order to fulfil diverse applications. Although data is handled at proprietary databases, it is possible to access standard commercial databases. *[HELOKUNNAS,1995]* Environmental Systems Research Institute's (ESRI) ArcView and MapInfo can be given as an example of software using this approach.

The dual database approach is the most common approach used in GIS. Each geographic entity is decomposed into its respective spatial and thematic components and stored in separate "dual" databases. ArcInfo, Intergraph MGE and GeoMedia Professional GIS software can be given as an example for this approach. The thematic component is stored entirely within a commercial Relational Database Management System (RDBMS). The geographic component, conversely, is stored in a proprietary database with its own unique internal access, through feature identifier, and storage methods illustrated in Figure 3-10. The stability of system is achieved generally through feature identifiers, which are automatically assigned or integrity rules provided by software vendor. The decomposition of spatial an non-spatial data was inevitable in, since it was not efficient to store spatial information and non-spatial information in the same database. The reason for this non-efficiency was the lack of appropriate spatial indexing mechanisms for queries in databases. Additionally, visualization could not be realized. Due to these reasons, the proprietary storage mechanisms are used for the geographical data, in order to increase system performance and to be able to visualize the spatial information.

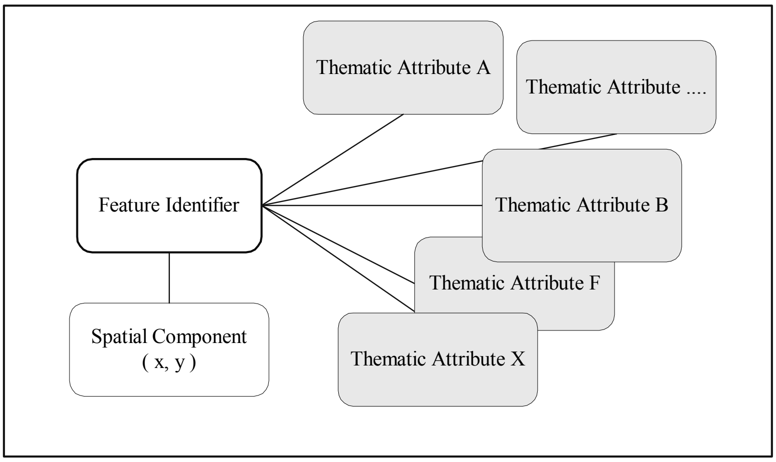


Figure 3-10: Components of Dual Architecture [MOLENAAR,1991]

Within this approach, the thematic aspects of data have the first priority in GIS. Structure and processing of spatial data was secondary. *[MOLENAAR,1998]* The spatial component of an entity, in this context, mainly refers to its geometric representation (i.e. point, line, polygon) and required only for visualization of thematic data. Generally at practice geometry is redundantly stored in the system. Although the concepts behind the design of these spatial database management systems are well documented, the actual systems are ultimately black boxes. *[CHURCH et al,1994]* Due to this non-open architecture, redundancy of geometry can not be easily detected. In addition relational databases can not handle required amount of integrity rules, including spatial and non-spatial information integrity, as they are not tailored for this purpose.

Because of data maintenance problems and data integrity issues several other approaches such as shell architecture and integrated architecture were developed. The shell approach, stores all data within one database with a spatial support "shell" on top of a RDBMS for geographic queries. In this architecture, spatial data is separated into its basic elements (i.e. points, lines, polygons) and stored separately in related tables. To retrieve the information, relational joins are performed to reconstruct the required geographic entities. Smallworld GIS software can be given as an example. The efficiency of this approach, therefore, highly correlated with the design of the spatial shell appropriation. In many ways the proprietary shell between the user and the underlying RDBMS has the same drawbacks as the dual architecture. The data model and the shell is designed and established by the GIS vendor, and the system is essentially non-extensible. *[CHURCH et al,1994]*

After developments in database technology, especially enhanced spatial indexing systems and object-oriented technology, the integrated approach has emerged. With the integrated approach, it is possible to store spatial and non-spatial information in one database management system with the required performance in spatial queries with the spatial indexing mechanism. With this approach, an extended DBMS with spatial data storage capabilities or object-oriented DBMS is required. With this approach users can extend the system through user defined objects and methods, contrary to other approaches, where only software vendor can decide which type of objects should be used. The Oracle8i with Spatial Option is an example where this approach is applied. There are software vendors also supporting this approach through their own spatial data storage concepts, such as ArcInfo with Spatial Data Engine (ArcSDE). With this approach unlimited amount of real world object can be modeled. Especially with its integrating approach and provided possibilities of storing spatial data in relational database systems, this category is taking very much attention. The redundancy is lessen compared with dual approach, since data is integrated in one database. However, due to user-defined objects, there is an risk in data model integration of diverse sources with respect to compatibility. This risk needed to be minimized with complete, formally described conceptual data models.

The evaluated GIS-T applications were realized using dual and shell architecture, that spatial and non-spatial data are stored in separate databases and then linked back to together, using unique identifiers. A special aspect of highway spatial information systems, which is linearly referenced data was non-considered in studied countries, with an exception of the USA. They are considered as non-spatial information, which abolishes advantages of dual architecture. In the USA, these problems are already recognized and there are some efforts of integrating linear referenced data with three dimensional reference system but these efforts are yet not satisfactory in many cases.

3.3.2 Standards Established for GIS-T

With the prevalence of GIS usage and the resulting raise in data sharing issues, standardization efforts are increasing worldwide.

At the international level the International Standardization Organization (ISO) has three technical committees (TC) for the definition of data standards relevant to the topics of this research. These are TC 211 which deals with general GIS, and committees TC 204 and TC 22 which relate to road information. The scope of TC 211 is to establish a structured set of standards for geographic information. The standard proposal aims to standardize general aspects of GIS, such as methods, tools and services for data management. The TC 204 is responsible for the standardization of information, communication and control systems in the field of urban and rural surface transportation, including inter-modal and multi-modal aspects, traveler information, traffic management, public transport, commercial transport, emergency services as well as commercial services in the transport information and control systems (TICS) field.*[ISO,2001]* The TC 22 is concerned with compatibility, interchangeability and safety, with particular reference to terminology and test procedures (including the characteristics of instrumentation) for evaluating the performance of road vehicles and their equipment.

Another international standard is Digital Information Geographic Exchange Standards (DIGEST). This was defined by the Digital Geographic Information Working Group (DGIWG), an initiative established by eight NATO countries. DIGEST was planned for civilian in addition to military purposes *[BILL-II,1999]*. The aim of DIGEST is to assure compatibility between multi-national agreements for digital data standards with respect to supported data structures, feature and attribute coding scheme, exchange media, format and administrative procedures *[DIGEST,2001]*.

The International Cartographic Association (ICA) is an established working group on digital cartographic data exchange standards. It aims to identify research needs arising from the standards process and information exchange at the international level.

There are also activities by the Open GIS Consortium, which is an non-governmental consensus-based association of public and private sector organizations. The standard proposed by the consortium, Open Geodata Interoperability Specification (OGIS), has the scope of providing a single 'universal' spatio-temporal data and process model. The standard will cover all existing and potential spatio-temporal applications, a specification for each of the major database languages to implement the OGIS data model and a specification for each of the major distributed computing environments to implement the OGIS process model.

The Geographic Data File (GDF) has been developed as part of a European Community (EC)-sponsored project to develop a European Digital Road Map (EDRM). Since 2000 the ISO version of the GDF (ISO TC 204) format has become an international standard. The primary usage of GDF is for car navigation systems. It can also be used for many other transport and traffic applications including, fleet management, dispatch management, traffic analysis, traffic management, automatic vehicle location, road maintenance and public transportation. TC 278 and the GDF standards correspond to the needs of vehicle manufacturers in the area of routing systems.

In the USA another standard has been developed, namely the Spatial Data Transfer Standard (SDTS). This allows U.S. Federal Agencies to share spatial data between applications which use differing hardware, software and operating systems. SDTS was designed for use with data such as topology, raster data, hydrographic and topographic data. Also SDTS forms part of the DIGEST specification set used by European members of NATO.

At the European level, in the field of spatial data, European Committee on Standards (CEN) is represented by TC 287, which has the scope of identifying and defining a structured set of concepts and components in general terms. The standard includes defining, structuring, querying, updating geographic data and metadata with respect to geographic information. Metadata can be defined as data about data. This definition leads to the evaluation of current data according to various criteria such as logical consistency and completeness. Consequently an evaluation framework is determined, which extend beyond the required quality criteria within the field of geodesy and both suitable for spatial and non-spatial data.*[SCHEU,1995]* However, several evaluation criteria including positional accuracy, spatial attribute accuracy and logical consistency can be simply ensured with the help of approaches in the field of geodesy, especially with the use of adjustment techniques.

Additionally CEN TC 278, which was established specifically for road, transport and thematic data standards, is concerned with standardization in the field of telematics. The focus of the standard is the application of telematics to road traffic and transport, including those elements which need technical harmonization for inter-modal operation. Additionally, the standard seeks to support communication between vehicles and road infrastructure, in-vehicle human machine interfacing, traffic and parking management, public transport management and user information *[CEN,2001]*.

In Germany the Object Catalogue for Road and Traffic Information (OKSTRA) has been initiated *[OKSTRA,2000]*. This is an on-going effort in the standardization of road information exchange between state highway administrations and the divisions of these agencies. OKSTRA is a conceptual data model, established in 1998, which is especially focused on the modeling of the external data schemas in the highway administration divisions. OKSTRA coverage ranges from the road design to documentation of traffic data storage.

The quality, conformance and metadata aspects are also considered by the above-mentioned organizations. Although there is diversities between established standards in several aspects, considering the data and the data model quality presence; a general acceptance has been reached. Positional and semantic accuracy of data should be provided and the conceptual data models should be complete and logically consistent. Consequently, conceptual data models designed for GIS-T should ensure data modeling aspects and certificate data accuracy. A comparison of GIS standards is shown in the Table 3-3.

There are many national and international organizations and communities seeking to establish standards for GIS, some of which are referred to above, these organizations have different scopes. Consequently, each concentrates on a particular aspect of GIS. These diverse scopes are one of the reasons why none of these standards can be applied alone to GIS-T without adaptation. Specifically there is no standard yet available for GIS-T, which includes all of topology, linear referencing systems, graphical representation of road segments and abstraction levels of information issues.

Table 3-3: Comparison of GIS Related Standards [CASPARY,1998]

		FIP 173 SDTS 1988	DIGEST 1997 Ed.2	ICA 1995	CEN 278 1996	CEN 287 1996	ISO TC 211 1997
Meta Data → Sources	lineage	☑	☑	☑	☐	☑	☑
	purpose	☐	☐	☐	☐	☐	☑
	usage	☐	☐	☐	☐	☑	☑
	Temporal information/ accuracy	☐	☑	☑	☑	☑	☑
Meta Data → Model	resolution	☐	☐	☐	☑	☐	☐
	precision	☐	☐	☐	☑	☐	☐
	Clipping indicator	☐	☑	☐	☐	☐	☐
	Abstraction modifier	☐	☐	☐	☐	☑	☐
Conformance → Data Specification	Format Specification	☐	☑	☐	☑	☐	☐
Quality Elements → Accuracy	Position Accuracy	☑	☑	☑	☑	☑	☑
	Attribute accuracy	☑	☑	☑	☑	☑	☐
	Semantic accuracy	☐	☐	☑	☐	☑	☑
	Correctness	☐	☐	☐	☑	☐	☐
Quality Elements → Model	Completeness	☑	☑	☑	☑	☑	☑
	Logical consistency	☑	☑	☑	☑	☑	☑

According to established standards, some aspects can be used in GIS-T as a framework. Among the established standards, the closest to GIS-T are the intelligent transportation system standards (telematic). This is because the main interest is digital road information and data exchange between databases. Standards established for the telematic data are presented in Table 3-4.

Table 3-4: Quality Parameters for Telematic Data [WIDMANN,2000]

	Contents criteria	Spatial criteria	Criteria related with time
Completeness	Example: Thematic attributes	Example: Coverage of relevant area	Example: Capturing of dynamic information or time interval of data capturing
Precision	Scale of information structure	Accuracy of spatial data	Actuality and accuracy of date and time
Consistence	Independence of spatial and time related information	Independence of contents and time related information	Independence of spatial and contents related information

According to the GDF concept a "single hypothetical database" should have the following characteristics and components *[GRÜNDIG,1989]*:

- Positional accuracy
- Be transformable in order to be usable in applications requiring routes and linear referencing systems.
- Be available in a common, useful data format.
- Be subject to accuracy certification. The status of certification should be an attribute for each feature.
- Have a permanent feature identifier.
- Have a minimum set of attributes, including permanent feature identifiers, route numbers, names, road type and metadata
- Maintain topological connectivity
- Have maintenance standards.

3.3.3 Data Models Evaluated

The objective of examining current conceptual data models was to understand the capabilities and advantages of the established systems. Additionally, the problematic areas facing highway administrations were determined. As any conceptual data model requires some degree of interpretation, comparison of these conceptual models is out the scope of this study. Examples from the USA, Germany, Denmark and the international conceptual data model GDF will be discussed.

3.3.3.1 A Sample Conceptual Data Model Applied in the United States of America

GIS-T technology is widely used in the USA, especially by federal state highway administrations. Between 1996 and 1999 the number of GIS applications used by departments increased enormously (268%). These applications were mainly concerned with base map production, data management, linear referencing, GPS, information flow between departments, data distribution and internet technology *[GIS-T/ISTEA,2001]*.

National Cooperative Research Program (NCRP) was created in 1962, which is administrated by the Transportation Research Board (TRB) and sponsored by the member departments of the American Association of State Highway and Transportation Officials (AASHTO), in cooperation with the Federal Highway Administration (FHWA). NCRP project 20-27 was initiated in response to the need to define the basic structure of GIS-T based on current and anticipated needs and the characteristics of transportation agencies. During the studies carried out for that project, it was noted that the two major problems in defining transportation objects were *[NCHRP,1998]*:

- Various definitions of road
- Varieties of criteria used to break roads into logical segments.

In order to cover these issues two proposals were made. In their proposal, Dueker and Butler designed a GIS-T data model which defines relations between transportation data elements. It is based on a feature (object) database approach with legacy data of varying spatial accuracy. An alternative approach is a location (geometry) approach as suggested by Sutton (1999). This alternative was designed to work in an environment where the location of all transportation features would be re-collected using high precision GPS. This approach focuses on enabling the linking of spatially accurate tracking or events to a spatially accurate map base *[TRB, 1992]*. Because the Dueker-Butler conceptual data model is designed, not only to solve problems of linear referencing specifically, but also for the entire agencies, it will be examined in this study. In addition, various case studies in several agencies were pointed out in literature with respect to implementation of Dueker-Butler conceptual data model.*[SUTTON, 2000] [VONDEROHE,1997] [BENDER, 1999]*

The Dueker-Butler data model is designed for all modes of transport such as highways, railways, maritime lines, and airlines at universal enterprise level. The proposed data model is based on the independence of geographic datum, transportation system events, geometrical system representation and the topology of links and nodes comprising the transportation system.*[DUEKER,1997]* The model is constructed through a series of steps. These begin with the basic elements to which elements accommodating more complex needs are successively added. The entire database is designed to eliminate the need for complex GIS software for all methods except to display events and their derivatives using vendors' dynamic segmentation concepts. The basic data model is shown in Figure 3-11.

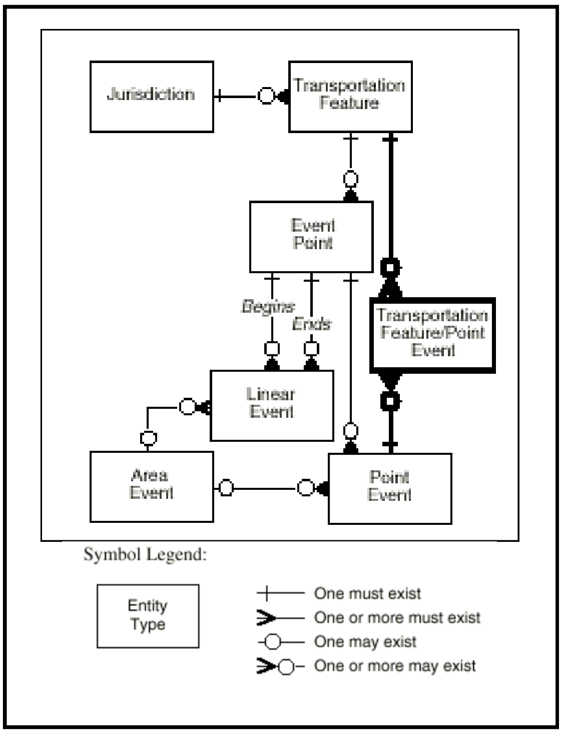


Figure 3-11: The Basic Elements of the Dueker and Butler Model [DUEKER,1997]

In the basic part of the conceptual data model, **Jurisdiction** is mainly used to represent political districts. **Transportation Feature** is an identifiable element of the transportation system. Within a jurisdiction, the transportation feature is identified uniquely *[DUEKER,1997]*. In order to define a transportation feature, the jurisdiction identifier and the transportation identifier are required. The location where an event occurs is called an **Event Point**. The **Event Point** is defined initially as an offset distance from the beginning of the transportation feature event, which is a similar concept with a linear referencing system. A road event is described as an attribute, occurrence, or physical component of a transportation feature. There are three event subtypes; namely **Point Event**, **Linear Event** and **Area Event**. **Point Event** occurs at a single event point location. A **Linear Event** is a component or attribute defined by two event points, representing the beginning and ending points. An **Area Event** is a transportation feature component or non-transportation event related to a **Transportation Feature**.

Although **Transportation Feature** possesses a subordinate point, and a linear and area event, the transportation feature to event relationships go through the event point entity. The use of only the event point entity is described in Dueker and Butler model thus; "An event is `owned´ by a transportation feature by virtue of its location on that feature; i.e.., its defining event point(s)" *[DUEKER,1997]* .This is because road events are defined using two event points, being beginning and ending event point similar to linear referencing concept.

In Figure 3-12[1] the completed basic GIS-T conceptual data model is presented. Further detailed requirements are realized by introducing entities to represent topology, cartography and linear datum.

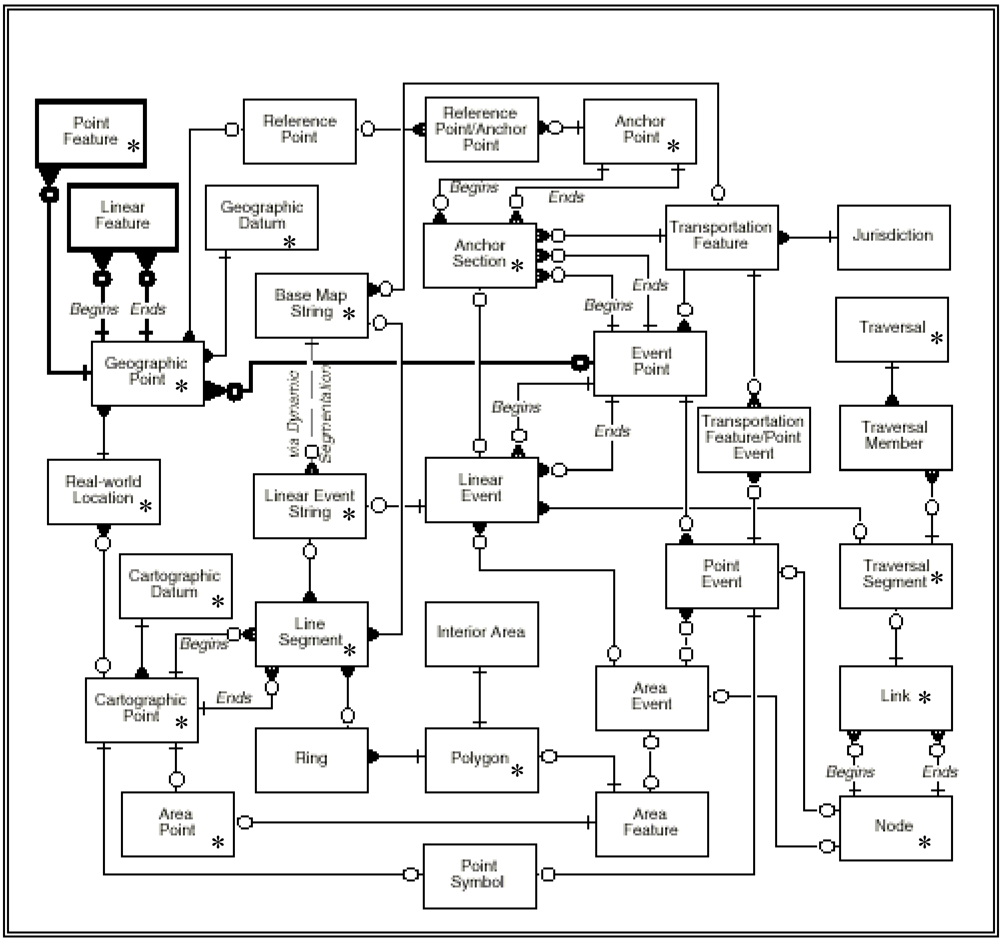


Figure 3-12: The Complete Conceptual Data Model [DUEKER,1997]

Topology's main element is **Traversal** which is a part or route consisting of one or more links. The atomic component of a traversal is the **Traversal Segment** defined by a link and its attributes. Additionally, **Link** and **Node** topological elements were introduced into the system.

The concept of the linear datum used by this conceptual data model had been developed by Vonderohe and Hepworth, 1996. *[VONDEROHE,1997]* The linear datum was used to reduce the impact of error sources including multiple network models, defining one linear referencing system and thereby improve the accuracy of location data collected according to a linear referencing system (LRS). The proposed linear datum is based on a set of well defined and precisely located anchor points and anchor sections, to which the linear reference

---

[1] *In Figure 3-12 entities with (*)sign are discussed in this study*

measurement methods may be calibrated. Anchor points are one dimensional control points, which their position can be determined and recovered in the field. Anchor sections are non-branching linear features, connecting two anchor points, whose real-world length (in distance metrics) can be determined in the field. Anchor sections are directed by specifying a “from” anchor point and a “to” anchor point. Anchor sections have a distance attribute, which is the length measured on the ground. *[VONDEROHE,1997]* In order to realize this, entities **Anchor Point**, **Anchor Section** were introduced into the Dueker and Butler model.

Cartographic and geometric elements were introduced into the model. Several new entities were added to support the typical cartographic elements found on transportation maps, including those created using the dynamic segmentation methods in GIS software. Some of the introduced entities are: **Geographic Datum, Cartographic Datum, Point Feature, Geographic Point, Cartographic Point, Real-world Location, Area Point,** and **Polygon**. In order to realize linear referencing requirements, new entities were introduced into the system, including **Line Segment, Base Map String,** and **Linear Event String**.

The reason for using various criteria in order to break roads into logical segments is discussed as; “In some ways, **Traversal Segment** is a resolution entity between **Traversal**, **Link** and **Linear Event**. A **Traversal Segment** may be used to define super-links by relating non-contiguous nodes located on the same transportation feature (Butler, 1995)” *[DUEKER,1997]*.

Within the proposed conceptual data model, the implementation recommendations are also provided, which will clarify the concepts introduced. Some of them are illustrated here in Figure 3-13 and Figure 3-14.

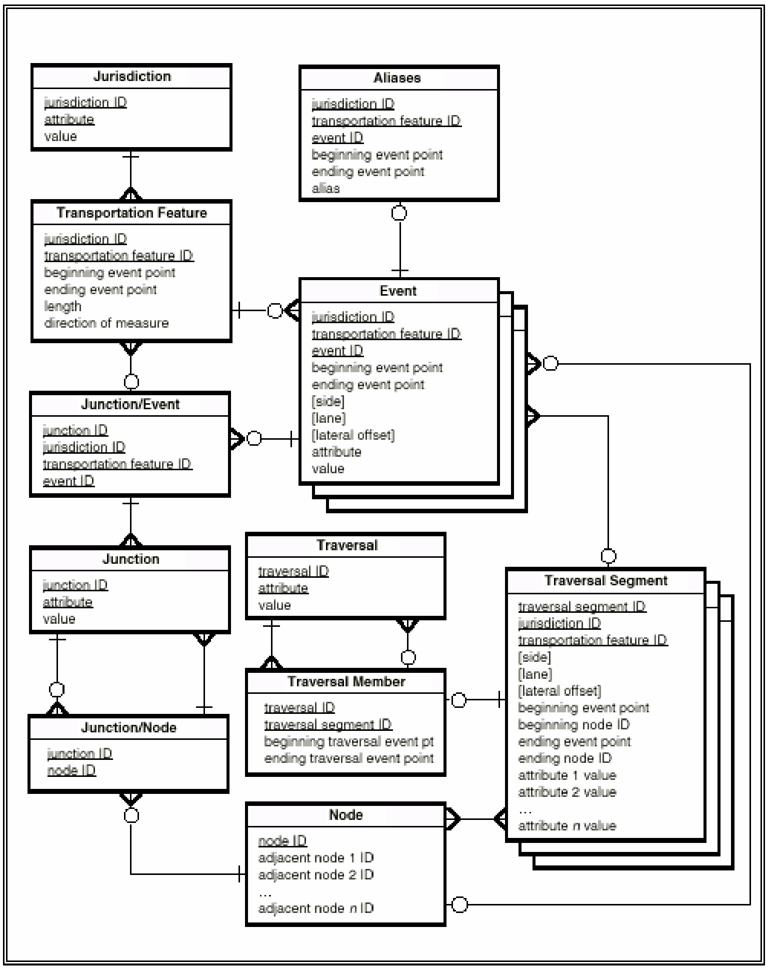


Figure 3-13: Implementation Recommendation of Road Events [DUEKER,1997]

In Figure 3-13 an **event_ID** is used to identify records in the **Event** tables, which are unique for only a given transportation feature. In the first sight, it can be seen that there are some differences between the conceptual data model and the implemented data model. For example, the feature **Junction** was not introduced in the conceptual data model and the **Event** feature was initially introduced as **Point, Linear** and **Area Event**. This can lead to confusions during implementation. Additionally, for example, in the implementation proposal adjacent node identities are stored as attributes within the **Node** entity. These spatial relationships could instead have been defined using **Link** entity and spatial relationships.

Although in the conceptual data model the **Link** feature was present, in the implementation tables **Node** is directly related to the **Traversal Segment** feature. The usage and implementation of **Link** feature is not clear. In the implementation recommendations below Figure 3-14, the length and direction of measurements are stored in two different tables, namely **Trans Feature/Anchor Sect** and **Transportation Feature.** This is redundant and will definitely lead to data maintenance complications.

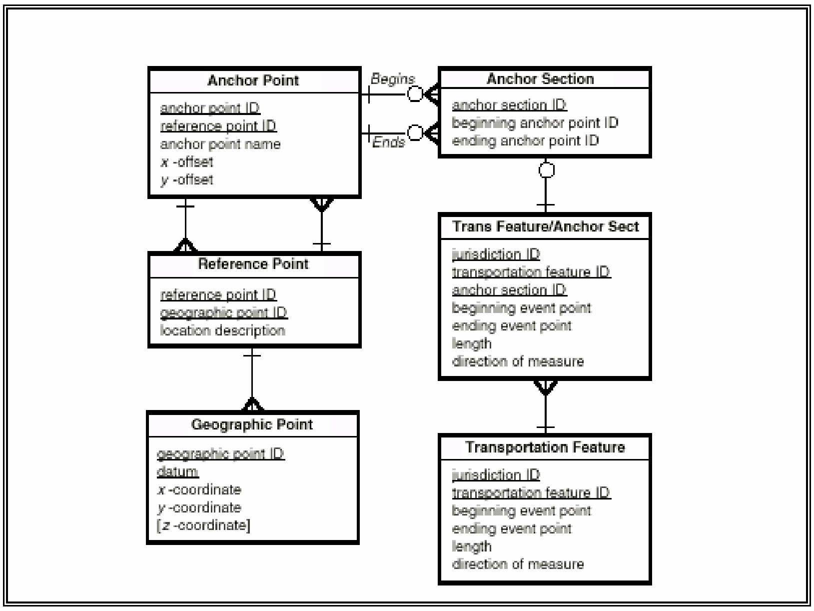


Figure 3-14: Implementation Recommendation of ***Anchor Section*** *[DUEKER,1997]*

In the **Transportation Feature** table, the beginning and ending event points are those of the entire transportation feature. In the **Transportation Feature/Anchor Section** table, the beginning and ending event points are for the limits of the subject **Anchor Section**. The **Anchor Section** table includes the beginning and ending anchor point_ID' s. This redundancy is required in order to perform maintenance of linear LRS locations of anchor points separately from the datum entities of anchor points and anchor sections. However, this needs to be carefully implemented and controlled by integrity constraints.

The main advantages of the proposed conceptual data model are; firstly, it is designed for an entire agency where all external transportation agency schemas are integrated. Secondly, the independence of road events, being thematic data, from topology is proposed. Thirdly, the topology is considered independent of geometry, which in many other data models is ignored. This approach will definitely bring many advantages for database maintenance and the simplification of analyses. Additionally, linear referenced transportation information is conceptually modeled. In the other data models examined, with some exceptions, this issue is generally ignored.

Open issues concerning the conceptual data model are:

- Efforts of integrating linear referenced data with other information in the conceptual data model is provided through another linearly referenced method, although this proposal results redundancy in both conceptual data model and the database.
- The linear referencing system is based on anchor point-section system. There are several disadvantages to use only one linear reference system in order to relate road events. This would require a change in the use of all other custom methodologies in highway administrations and reduce them into one method. During this conversion process, it is possible that some data will be lost. This issue will be discussed in Section 4.3.1.
- Anchor points must be complete and their field locations need be determined. This will incur extra costs for the maintenance of these points.
- The negative impacts of dynamic segmentation, such as the creation of additional entities, are also carried to the model. Additional entities introduced such as; **Line Segment**, **Basemap String** causes redundancy. Due to the requirements of dynamic segmentation software, Dueker and Butler conceptual data model illustrate the relationship of base map strings and the transportation features as many-to-one.*[SUTTON, 2000]* The many-to-many relationships can not be implemented.
- The proposal to establish unique feature identifiers, such as identifying them according to jurisdictions, can cause problems. This is because jurisdiction borders are non-persistent spatial frames.
- Definitions such as “**Event Point, Point Event, Point Feature, Reference Point, Anchor Point, Node, Geographic Point, Cartographic Point, Point Symbol**”, and the differences between these are confusing. This will cause many problems and redundancy at the implementation stage.
- The introduction of the **Traversal Segment** and **Anchor Section** entities define topological element **Link** with different properties. However during implementation a clearer description is required in order to perform integrity constraints.
- The requirement for transferring topology, imposed by the use of links and nodes to connect traversals (a.k.a., linear transportation features) to the datum and thence to the cartography, places a still greater burden on data sharing *[DUEKER,2000]*.

3.3.3.2 The German Highway Administration and A Sample Conceptual Data Model

Through the introduction of the "Instruction of Road Data Base (ASB)" in September 1968, the federally uniform structure of road data in Germany was ensured *[BMVBW,1998]*. ASB is a very useful and detailed document for the describing of data collection regulations in highway administrations. Within the ASB, roads are described in terms of construction information, geometrical elements, type, facilities, length, vertical sections, cross-sections, materials, width, thickness, and current road condition information. According to the ASB instructions the “Road Information Database (SIB)” was constituted. Research projects, including “Revision of structure and contents of ASB(2000) (09.120)”, “Federal-Information System-Road (98.192)” have continued with the target of extending or restructuring the ASB according to changes in the basic technical conditions and the growth of new fields like telematics or information systems *[BMVBW,1998]*.

The general concept of the ASB was developed using road networks as frames, where the node-link oriented linear referencing system is used. The link-node method uses node numbers, identified with physical features in the field as an intersection. Links are defined in terms of beginning and ending nodes, and are the logical connections between nodes. In the link-node method, attribute locations along each link of the roadway system are specified and an event that are recorded are measured as an offset from the nearest node number along a link. A node map is required in order to be able to identify the location of the nearest node and in order to track node number records. A link-node method has the advantage of not needing to maintain reference posts in the field. The disadvantage appears in usage, due to requirement of a node-map, since there is a lack of communication between third-party users. Within ASB, firstly, nodes are determined and uniquely identified using the numbering system. Each link, defined by two nodes, is measured and referenced with a sequential stationing method, which begins and terminates with a node. There exists four types of nodes being; planar node (without height difference), non-planar node (with height difference information), node partly independent of plan and fictitious node.

In highway administrations road data is stored within a relational databases, according to the ASB regulations. Approximately 600 features are present in databases. In addition to common information such as pavement type, road length, road type, other detailed information such as bikeway, tree age and tree type is available. The databases generally used by highway agency are Oracle and MSAccess. The desktop GIS software MapInfo is commonly used in order to visualize these databases.

Integration of the road database (ASB) and GIS is mainly realized using links, which are uniquely defined using assigned nodes. As a digital base-map, 1:10.000 scaled governmental topographical land registration maps (ATKIS) are used.

Since the ASB is designed as a regulation for data collection, not as a conceptual data model, the spatial information characteristics were not considered. However, the ASB is often used by highway agencies instead of a conceptual data model for GIS. Due to this, it is very common that, information also includes spatial character assigned to links. The spatial information is not modeled, instead it is directly stored in the database. It is typical that dual architecture GIS software is used. Because of this, spatial information in the relational database is generally also redundantly stored in software vendors' proprietary repositories. The geometrical information, such as alignment elements, relative height information are stored in database tables such as the sample shown in Table 3-5. Additionally, although data is stored in a relational database, relational data modeling techniques are unconsidered. An example can be given again in Table 3-5, which contradicts to the second normal form.

Table 3-5: A Sample Table From Information System

Road	Class	Nr. (LFD)	From Node (VNK)	To Node (NNK)	From Station	To Station	Length	Alignment	Type	Radius	Clothoid Radius
B 2	B	46	3347012	3347010	0	84	3028		G	0	0
B 2	B	46	3347012	3347010	84	208	3028	R	R	600	0
B 2	B	46	3347012	3347010	208	455	3028		G	0	0
B 2	B	46	3347012	3347010	455	575	3028	R	R	350	0
B 2	B	46	3347012	3347010	575	597	3028	R	A	80	350
B 2	B	46	3347012	3347010	597	617	3028		G	0	0
….	….	….	….	….	….	….	….	….	….	….	….

Due to the spatial character of information, GIS-T can not be used in order to integrate diverse information sources. Although there are efforts to integrate these different sources in GIS using existing network link information, which is available both in databases and in digital base-map of GIS, the required efficiency cannot be reached because of redundancy. As a result, generally both GIS and road databases are used in parallel by highway administration.

In order to overcome such problems, several projects at the federal state level have been initiated. One of these projects is the "Road Information Database Nordrhine-Westfalen (NWSIB)". It links all data with a highly detailed digital road map and visualize them. The NWSIB builds on available databases and telecommunications networks. *[NWSIB,1998].*

The basic component is **road,** which is ensured by consistent geographical referencing and monitoring of information. The object model of the NWSIB in its current stage of development is divided into four basic components: Road network, inventory, geometry and access rights. The network component does not depend on the inventory component and forms a basic part of the information system. During the design and implementation phases, Smallworld GIS system, which has a shell-architecture, is used. The advantages and disadvantages of this architecture was discussed in Section 3.3.1. The object model's geometry relates explicitly to the abilities of the Smallworld software. Relation of objects to the road system is made by point, line and area characteristics. The history concept, which is one of the most required information by highway agencies, is modeled in the conceptual data model. Since in the highway agency GIS applications are implemented using MapInfo, direct connection of MapInfo to NWSIB is planned. In order to supply data for other applications with road system references, Smallworld database is connected to an external Oracle database.

The conceptual data model, which is partly provided by the software vendor, is shown in

Figure 3-16[1]. The symbols used are described in Figure 3-15.

All linear referenced data in the NWSIB has implicit geometry in a planar coordinate system, indicating position and linear geometry to the road network. Object **Link or Shoulder** is used to calculate relationships between stationing and planar coordinate systems *[NWSIB,1998].* **Current Stationing**, **Operation Kilometer** objects are introduced to the proposed conceptual data model in order to integrate the linearly referenced data of the ASB with the two dimensional coordinate system. This is related to the **Road Point** and **Road** entities. **Road Point** represents point geometry.

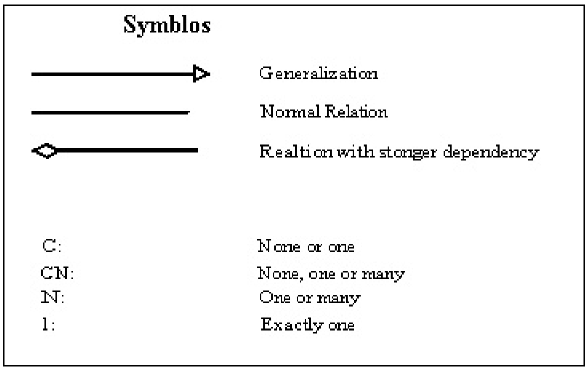


Figure 3-15: The NWSIB Conceptual Data Model Symbols

---

[1] In Figure 3-16 entities with (*)sign are discussed in this study

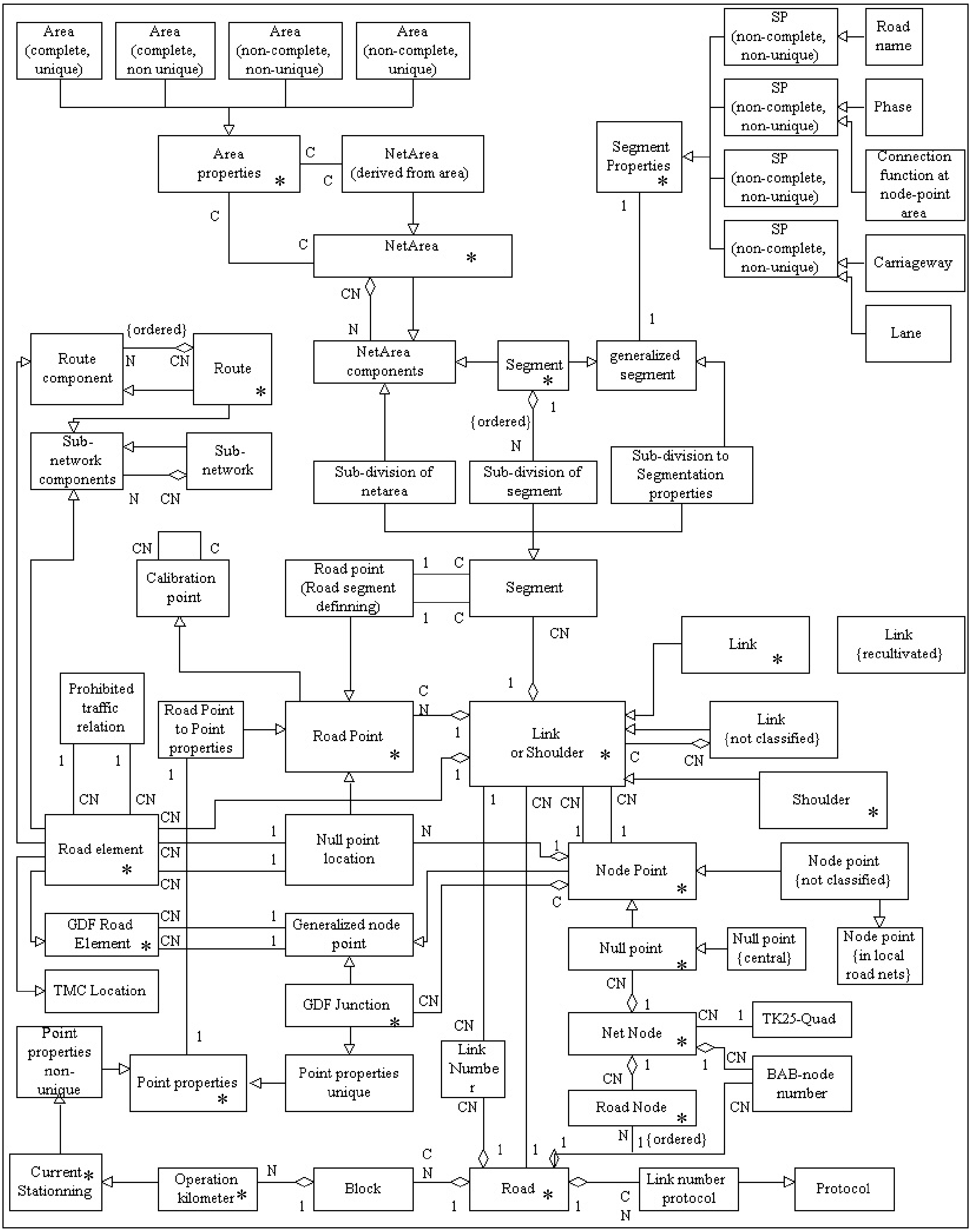


Figure 3-16: The NWSIB Conceptual Data Model Description [NWSIB,1998]

For the integration of linear referenced data, the data segmentation concept is used. The advantages and disadvantages of this are discussed in Section 4.3.1. Since a direct transformation is not available in the system, these information needed to be modeled separately and integration of these district information must be performed with the use of additional methods. Additionally, data maintenance must be performed separately. The **Route** entity is a compound road element. In the proposed system, there exists integration possibilities the GDF international standard and the Object Catalogue for Road and Traffic Information (OSKTRA). Two additional object **GDF Junction** and **GDF Road Element** are also incorporated into the data model in order to provide an interface to the GDF topological structure, since the GDF topological structure is not automatically derivable from available data. The **Road Element** entity is related with the **GDF Road Element**.

Non-spatial and spatial objects are modeled composite. Examples from the provided conceptual data model, which are designed following this concept, are **Segment** and **Segment Properties**, **NetArea** and **Area Properties** and **Road Point** and **Point Properties**. Topological elements are represented in the conceptual data model with various objects such as **Link, Shoulder**, **Node Point**, **Null Point**, **Net Node** and **Road Node**.

The system architecture of the NWSIB is shown in Figure 3-17. It can be described as follows: Firstly, the ASB conformed road database is imported into the system. In the figure the ASB conforming database is called database A. It is stored in the software vendors proprietary database. This holds both geometrical and thematic data, which are labeled here G(Geometry) and T(Thematic). In this database, A is stored without any decomposition of its spatial character. It is possible to import other information sources defined as C in the diagram, which can be both spatial and thematic. Planar geometrical features (G) and Thematic data (T), are extracted in order to be used for visualization and presentation. Finally, ASB database (A) is combined with other databases T and G, if additional information is required. During this process, if there is an update by the user, the updated information is stored in a GIS database, and decomposed into G and T.

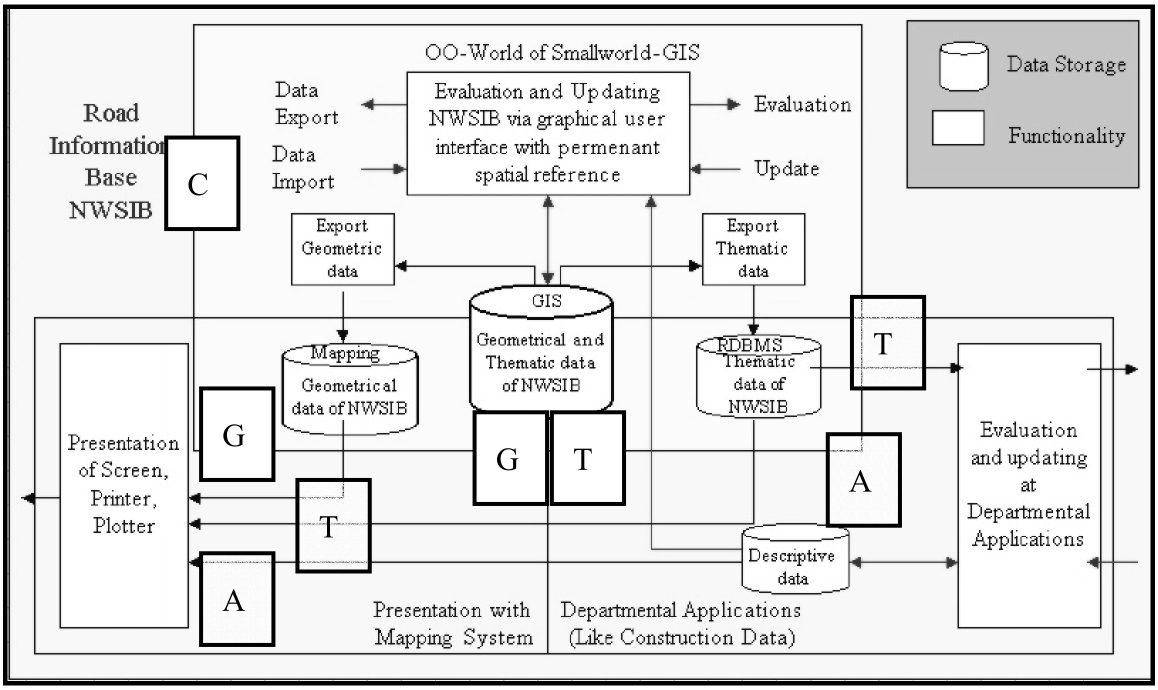


Figure 3-17: The System Architecture of NWSIB [NWSIB,1998]

The main advantage of the NWSIB conceptual data model lie in its integration of data sources and consideration of history information. Existing information and additional required sources are integrated into one GIS-T, in order to fulfill entire highway agency requirements. Using the history concept it is possible to model the current, past and future situations, which is highly required by all highway administrations. Additionally, compatibility possibilities with standards used in Germany such as GDF, ASB and OKSTRA, which was in other evaluated conceptual data models generally neglected, are existing.

The main disadvantage of this proposed data model and system architecture is the lack of decomposition between spatial and non-spatial information. Although geometrical, topological and thematic information is modeled in the conceptual data model, without conceptual separation of these components, redundancy cannot be controlled. Since the ASB is the basic information source, the linearly referenced data is stored as thematic information. With respect to the linear referencing issue, it differs little from the Dueker-Butler conceptual data model. Only one methodology, the node-link linear referencing method, is selected and spatial aspects are not considered.

It is expected that several problems will be encountered due to structural differences between the NWSIB and GDF. In the GDF, two topology abstraction levels were defined. However, in the described NWSIB conceptual data model, there is no topology abstraction level. Therefore, the information from one of GDF 's abstraction levels will be lost. Similar to the GDF, OKSTRA proposes a conceptual data model which is more detailed than the NWSIB. For example OKSTRA introduces alignment parameters, which are not included by NWSIB.

In the conceptual data model, network components are described as not dependent on the inventory component. However, in ASB, inventory components are dependent on network components due to usage of link-node linear referencing system. Consequently, a very careful implementation and data maintenance is required with appropriate integrity rules in order to control the redundancy.

Another problem arises in the conceptual data model during the description of geometry objects. Although a conceptual data model should be independent of any software vendor, realization of geometry objects is fully dependent on the software vendor definitions in NWSIB.

Referring to the system architecture proposed, and illustrated in Figure 3-17, the data flow can reveal problems. This is because, all information collected by highway agencies is considered as thematic data. Additionally, in practice the relational database management system (RDBMS) described in Figure 3-17 as T and A, are generally different databases. Therefore many various databases and parallel applications such as MapInfo and Smallworld need to be maintained at the same time.

3.3.3.3 Danish Highway Administration

A nationwide road information system, consisting of interconnected databases, is located at the Danish Road Directorate. The system holds the information about road and traffic conditions on national and regional roads. The Road Sector Information System (VIS) contains tools for extracting and presenting the information. *[VIS,1998]*

Various databases, each containing specialized information within a particular area, supply VIS information. The VIS provides map indices from which raster maps can be displayed according to submitted queries. Generally, 1:200,000 scaled maps are used in GIS applications. Due to this very large scale, VIS differentiates form the other evaluated conceptual data models. It is generally used for planning purposes. The road information in VIS is linearly referenced, based on the stationing linear referencing method. This method uses the white marker poles along roads. The linear reference system has been standardized throughout the agency. Measurements are taken by specially developed measuring vehicles, and data is subsequently transferred from the Danish Institute to VIS. GeoVIS is based on Mapinfo, which is widely used in the visualization of road databases in Denmark.

In the database, in order to authorize user update-rights, the users' Internet Protocol (IP) addresses are available, which was discussed in Section 3.2.4. Some other available information in the database includes; length, pavement, accident, industry areas, road alignment, route number, emergency line, date and statistic counting stations.

Road data is one-dimensionally considered and stored in the database using its road identifier and offset value. In a one dimensional reference system, a road is defined by:

- Road identifier (ID)
- Links
- Connections to other roads

In order to identify a road in the database, a road identifier naming regulation is used by the agency. For example:

000/ 0412/ 0 defines **Owner / Road number/ Side of the road**

Due to the its superior update and system architecture concepts, the Danish system is an effective system. The Danish GIS-T shares many similarities with above introduced German system, both positive and negative. It was noted that several databases, GIS systems and other engineering software are maintained in parallel. These include VIS, GeoVIS, Intergraph MGE and MS Geographic. Linearly referenced information is stored as non-spatial information. The naming regulations used in Denmark, will lead to non-persistency, due to identifier's spatial frames such as; owner federal country and side of a road, although the road identifiers used in GIS needs to be unique and persistent. Naming regulations for identifiers are discussed and a proposal in order to prevent this problem is made in Section 4.4.4.

3.3.3.4 The Geographic Data Files (GDF)

The European Community (EC) has sponsored several research programs to develop computer-assisted systems for route guidance, travel information, fleet management, traffic monitoring and many other applications. One of these research programs, the Drive Program, is concerned with the road infrastructure in Europe. The main object of Drive is to find a uniform way of collecting and exchanging traffic information between vehicles and the roadside. Drive has resulted in the establishment of the Geographic Data Files (GDF) standard which is now extensively used by both manufacturers and users of geographic information *[ERTICO,1996]*. The GDF, which was especially designed for vehicle navigation purposes, not only road objects but also various themes were incorporated in the basic data modeling concept.

The GDF conceptual data model defines three levels: Level 0 contains all the fundamental geometric and topological information. The topological information supports planar graph theory. The form and position of the Level 0 objects are defined in the three atomic units; Node, Link and Face. Level-I and Level-II hold additional information such as features and attributes. The objects in these two levels differ in their level of generalization. The simple features: point, line and area form the basic elements of the Level-I. The Level-II contains more complex features, which are constructed in terms of the simple features.

In the GDF, link is defined as a multi-abstraction level complex element. For example, in the case of divided highways it is possible to have eight nodes on one side and seven node on the other. Although the GDF is still restricted to two-dimensional applications, the possibilities of combining height information have been investigated. Metainformation is present in the data model, including areas such as; data acquisition date, accuracy or geodetic information. In this way quality aspects are covered.

The attribute concept of GDF is powerful, permitting segmentation as well as composite and time dependent attributes to be built up.*[GDF,1996]* Four category of attributes are introduced in the concepts of GDF as follows:

- Simple attribute
- Composite attribute
- Simple attribute with restrictive sub-attribute

- Composite attribute with restrictive sub-attribute

With respect to this concept, in GDF a hierarchical structure is proposed between attributes. An example of composite attribute concept GDF feature is **Divider,** which is defined with sub-attributes **Divided Road Element**, **Divider Type** and **Divider Width**. This attribute indicates the existence of a physical or legal divider (solid painted (double) line) along the center line of a single bi-directional **Road Element**. All sub-attributes together form the composite attribute **Divider.** *[GDF,1996]*

Attributes, which are called Segmented Attributes, are related to a feature in such a way that they reference a certain part of it. In the case of line features the particular segment is defined by a position **From** and a position **To** value. These positions represent the curvimetric distance, expressed in meters. Conversely, the position from and position to values may be left blank to indicate that the entire feature is subject to the associated sub-attribute. The GDF attribute concept is presented in Figure 3-18.

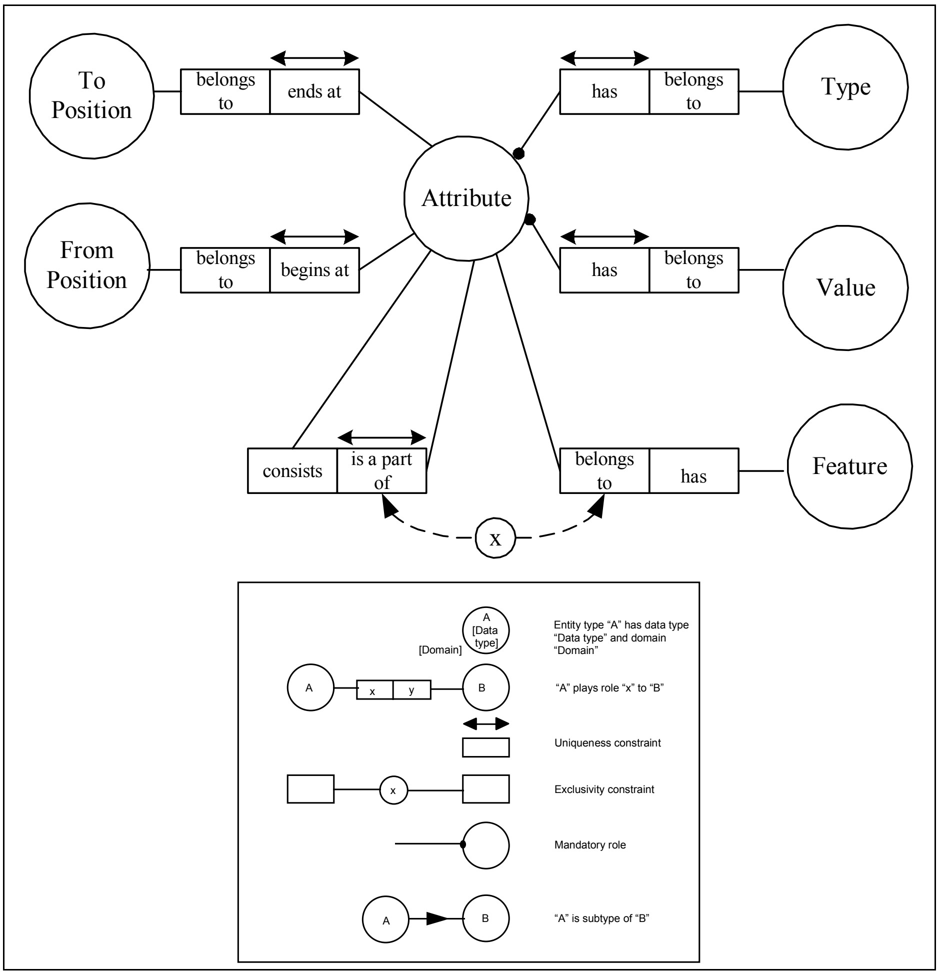


Figure 3-18: The GDF Attribute Concept [WALTER,1997]

The international standard GDF's data modeling concept is mainly adapted from other GIS-T data models. With its abstraction level concept and support for the representation of divided roads, many of the requirements of highway administrations can be fulfilled. Multiple-dimensions of road geometry information was designed, with the introduced complex elements. Metainformation was only considered in GDF, which has a vital importance in GIS.

However, there are several problematic areas. Some of these problems have already been pointed out in the literature. *[WALTER,1997] [DUEKER,2000].*With respect to GIS-T, some disadvantages are apparent, mainly;

- The GDF is based on planar topological concepts. This implies that every intersecting link of the planar graph defines a node. The planar topological basis has been widely used because of its simplicity. However, for applications where road network navigation is essential, such as the GDF, planar topology has many disadvantages. One of these disadvantages is that planar topology does not always represent real world phenomena appropriately. In particular constructions such as bridges, tunnels, overpasses are problematic. In order to alleviate this problem virtual nodes have been introduced into the system. These are termed "**Brunnel**"s from the combination of the words "Bridge" and "Tunnel". Such virtual nodes do not provide a complete solution. This topic will be further discussed in Section 4.2.3.2.
- In the GDF every simple feature is linked with a topology and corresponding geometry. Topology and geometry are not independent concepts in the GDF. Composition of these concepts requires additional efforts in data maintenance, in cases such as geometrical alignments in highway administration, although topology is invariant of geometry.
- In the terminology used by the GDF several descriptions can cause confusions, such as the terms **Node** and **Link** as used for geometrical features. For example in the GDF segmentation, a node must be inserted at each position where the road width changes where node term in this case is used instead of event points.
- In the GDF segmentation concept sub-attribute types are introduced, indicating an hierarchical order. In the case of changes in the main attribute, sub-attribute information will be disconnected. Additional efforts are required in order to connect this information again.
- Attributes are assigned using linear referencing methods in one dimension using **To Position** and **From Position** entities along a road segment. Several problems were encountered, especially during data maintenance, since multi-spatial dimensions and their integration is not considered. This will be discussed further in Chapter 4.3.1.
- The standard external feature identifiers do not exist for most of the features defined in the Feature Catalogue, apart from Administrative Areas.

3.4 Overview of Problems

The evaluated countries and conceptual data models present different aspects of the GIS-T current situation. Although levels of implementation and GIS-T usage vary; some common problems are:

- Due to diversity of highway administrations tasks, various abstraction levels of road information needed be coordinated, which was not fully realized.
- Road information data structure analyze is not generally performed. There are various definitions of road and road logical segments. However, highway administrations requirements were not fully satisfied.
- A huge amount of data needs to be collected. The existing data is mostly in analogue format. Data produced by one department can often not be shared with other users, due to lack of integrated approaches and adequate tools.
- There is insufficient, in some cases no, connection between GIS-T and databases. This results in;
    - Data acquisition and updates were realized separately.
    - Many different systems often need be used.

    - Systems were maintained in parallel.
    - No complete overview of the entire agency exists.
- Road databases are not tailored for GIS-T. Road databases were designed and implemented historically before GIS-T, without consideration of the spatial properties of road data.
- The necessary attention were not given to GIS data modeling. There is lack of formal description and documentation of designed data models.
- The difficulties in combining existing data in the database with GIS systems are mentioned by all highway administrations.
- The definitions and terms used in conceptual data models do not generally match.
- In evaluated countries, linear referenced road information is stored as attributive data. The integration of the various linear referencing systems used throughout an agency is generally not considered. Multi-spatial dimensions and their integration is neglected.
- The integration of linearly referenced data with other spatial data is provided using another linearly referenced method. Additional objects are introduced into conceptual data models, leading to redundancy. These solutions bring extra costs to implementation and maintenance.
- Representing road network topology by planar graphs causes many problems due to the inadequacy of the topological model for representing reality.
- Continuous update involving data entry is required. However, required update mechanisms and consistency controls have not been established.
- No standard is established with respect to GIS-T.
- For GIS-T purposes several regulations and definitions need to be established. Others must be redefined such as road identifiers.

The above listed technical issues will be discussed in further chapters of this study with proposed solutions. In addition to the technical problems, there are organizational problems. In particular there is widespread ignorance throughout organizations with respect to the fundamental principles and capabilities of the GIS-T. This issue has shown its effect, especially during user requirement assessments. However, this problem is correlated with technical issues, such as:

- Solutions were not provided for the entire agency. Therefore, this technology was not fully available to all users.
- In system architecture design, other technological possibilities, which can distribute the information efficiently, such as internet and intranet was not considered.
- In highway administration, user has various backgrounds. Therefore there exists various level of knowledge with respect to the GIS. This issue could be easily minimized with successive system designs, including user interfaces, which will lead minimum requirement of fundamental knowledge. Afterwards, the required minimum basic knowledge could be supported by means of workshops or other educational activities.

Chapter 4 : Proposal for an Integrated Conceptual Data Model

4.1 Criteria for the GIS-T

Based on the examined standards and conceptual data models evaluated in Chapter 3 an essential criteria list for GIS-T is constituted. This criteria list covers the problematic areas detected during the study, and aims to clarify the necessary properties of an successive highway administration conceptual data model. GIS-T data models should ensure;

- Topological, geometric and thematic information should be conceptually independent.
- Support for multiple topological representations and for various abstraction levels needed be realized.
- Non-planar topological model should be used.
- Spatial characteristics of road data, such as data collected using the linear referencing system, should be modeled with independent geometric and thematic character. An integration method is required.
- Thematic road data should not have a spatial nature.
- Highway administration business-rules should be modeled.
- Existing road information databases should be integrated into the system.
- Metadata, such as consistency rules, quality specifications and history information, must be incorporated.
- Interfaces to existing standards should be implemented.
- Permanent non-spatial unique feature identifier is required.
- Quick response to queries is needed.
- The conceptual data model should be designed independently of consideration of the specific software with which it is going to be implemented.

4.2 The Resulting Conceptual Data Model

The advantages of separating geometric and thematic data have been appreciated by the GIS community from the very earliest developments, which also remains the basic philosophy of most GIS approaches. However, such decomposition has not been implemented satisfactorily with the result that the problems associated with data maintenance have not been dealt with in an appropriate way. This is due to a lack of clarity in the separation of geometric and thematic data, as well as the widespread practice of integrating geometry with topology. Therefore, the main proposal of this data model is the conceptual independence of topology, geometry and road thematic data.

Due to the above observations, the data model is designed with three distinct components; topology, geometry and thematic road data. The basic component of the proposed data model is geometry, which is not usual in GIS. In most GIS, geometry plays a secondary role compared to thematic data. It is conventionally used for graphical visualization and is redundantly implemented. By considering geometry to be the basic component, many of the problems noted in Chapter 3 are avoided. Topology component is separately modeled in the proposed data model. Separating geometry and topology also leads to a consistent data model and increased topological analyses efficiency. In the proposed data model the thematic road data has no spatial character in the third-dimension, where today GIS are only capable of handling two and at most two and half dimension. The geometrical properties of the thematic data is provided by referencing the fully three dimensional geometry component of the model.

4.2.1 Overview of Highway Information Structure

In order to develop an appropriate highway information structure, a mind mapping diagram was used to identify clearly the required main components of the conceptual model. The main information used in the diagram refers to the analysis presented in Chapter 2 and Chapter 3.

Highway information can be structured into four categories; road events, topology, geometry and metadata. The road events describe the thematic information collected or required by highway administrations, which is of a non-spatial character. The geometry component is subdivided into three categories; point geometry, linear geometry and area geometry. The point geometry is defined in terms of a three-dimensional coordinate system, including height information. The linear geometry, which is again subdivided as; datum dependent and datum independent. It is defined in terms its parameterized geometrical elements which are either horizontal or vertical planes. Area geometry may be either planar or non-planar. According to the requirements of highway agencies, topology should be defined with two abstraction levels composed of the topological elements; link and node.

The metadata needed to be included with consideration of the database and data model standards defined in Chapter 3. Metadata is composed of: integrity constraints, history components, quality aspects and catalogue. The history component is used by highway agencies to fulfill the queries for past, current and future information. The catalogue is used to document the definitions and descriptions of the conceptual data model objects such as road event. An overview of the highway information structure identified is shown in Figure 4-1.

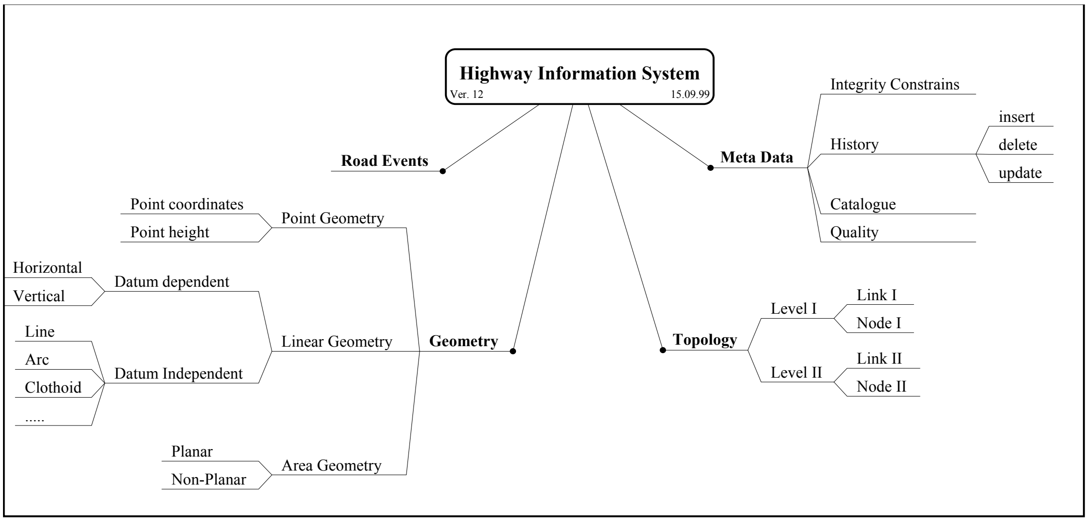


Figure 4-1: The Highway Information Structure

It is shown that all information required by highway administrations can be structured into four separate categories. The basic idea of the proposed conceptual data model depends on decomposition of these categorizes, in order to achieve the improved conceptual data model with well defined data structure and preventing redundancies.

4.2.2 The Data Schema

The proposed conceptual data model, with its relevant objects, is described using the Unified Modeling Language (UML). Every object is composed of attributes, instances and methods. According to the UML notation, relations between objects are defined through associations. A special relationship introduced by UML is **aggregation**. Aggregation differs from association in that it introduces a stronger relationship between objects, namely that one object is part of another object or has a relationship to an entire object. A multiplicity is a constraint on the number of objects that can be associated with another object. In Figure 4-2, the notation used is presented.

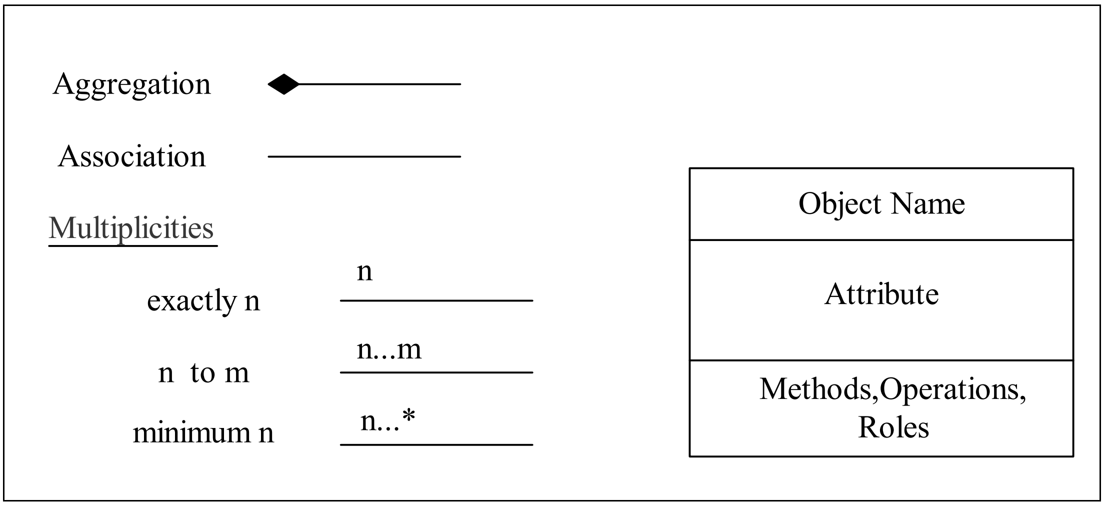


Figure 4-2: The UML Diagram Definitions.

4.2.3 The Conceptual Data Model

The proposed conceptual data model is introduced here step by step. Since the geometry is regarded as the basic element of the conceptual data model, the geometry component will be considered first.

4.2.3.1 Geometry

The geometry in the data model is composed of three main objects;

- Point geometry object.
- Linear geometry object.
- Area geometry object.

The conceptual data model designates points and linear elements in order to describe the road geometry. In GIS-T data models, the linear geometry is composed of either polygons or line strings in the horizontal plane. However, the requirements reported by many highway agency divisions indicated that it is necessary to model linear geometry by both planar and vertical elements. Due to this requirement, in the proposed conceptual data model, linear geometry is defined as being either planar or vertical parameterized linear elements. The vertical linear elements are defined in an (l, h) system as defined in Chapter 2. The reference system and transformation object, with its transformation parameters, are introduced. Attributes assigned to objects will be briefly described in this chapter. More information is also provided in the implementation explanations in Chapter 5.

Point Geometry is defined using a unique identifier shown as Point_ID in Figure 4-3. Other attributes shown in the figure are related to topology. These will be explained in the following section. The other geometrical elements **Linear Geometry** and **Area Geometry** are also defined using their identifiers as LinGeo_ID and AreaGeo_ID.

In GIS data models, information in different coordinate systems are transferred into one selected coordinate system. As a result the relationship between the reference system and point geometry is pre-defined as being **1:N**. This **1:N** relationship indicates that one reference system can have many points, but every point must be defined in only one reference system.

However, coordinates are properties of geometrical point elements capable of being defined in any reference system or any projection. It is therefore unrealistic to model points as being dependent on only one reference system. The conceptual data model should be independent of pre-definitions by software vendors. More importantly, in order to provide logical consistency in the conceptual data model, the relationship between **Point Geometry** and **Reference System** is modeled as being **N:M**.

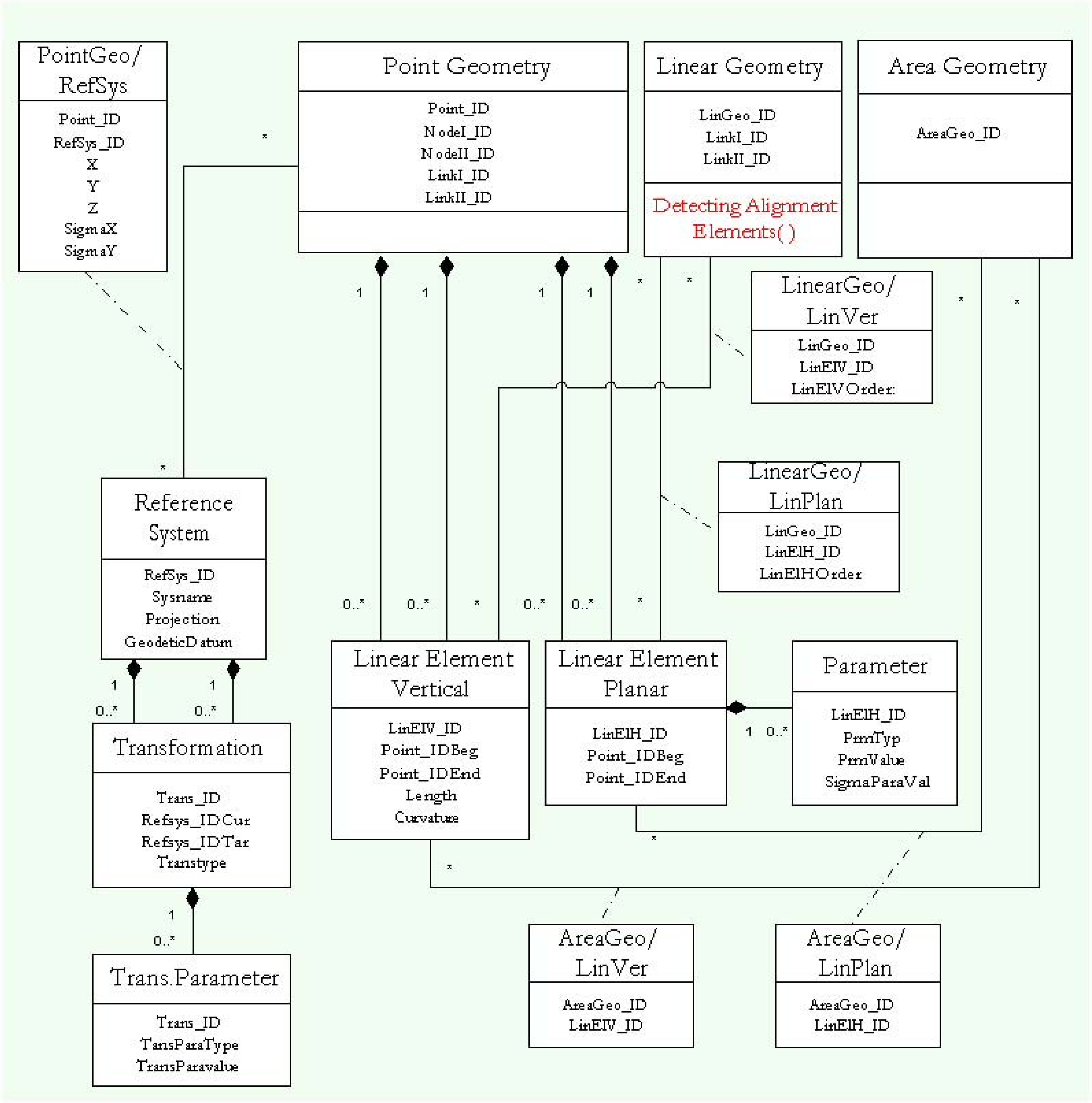


Figure 4-3: The Geometry Component of the Conceptual Data Model

Reference System is identified by a unique identifier RefSys_ID. Points can be assigned to many different reference systems and every reference system can have many points. The **N:M** relationships can not be directly implemented using current database system technology. This bottleneck was overcome in practice by the introduction of additional association objects between related objects. In order to establish a relationship between **Point Geometry** and **Reference System**, an additional table **PointGeo/RefSys** was introduced. The **PointGeo/RefSys** contains; Point identifier (Point_ID), Reference system identifier (RefSys_ID), point coordinates (X, Y, Z), as assigned by the reference system, and the standard deviation values of the point´s planar coordinates ($\sigma_x, \sigma_y, \sigma_z$).

In order to transform from one referencing system to another, two separate **1: 0..*** relationships are defined between **Reference System** and **Transformation**. These relationships describe the initial reference system relation and the target reference system relation. A reference system may have many transformation type and every transformation must be assigned to one reference system. These relationships are shown in Figure 4-3**.**

Transformations between reference systems requires the specification of transformation parameters in the model. This is achieved by assigning a **1: 0..*** relation between **Transformation** and **Transformation Parameters**.

In the proposed conceptual data model, each linear element is defined using its parameters e.g. length, radius. Three types of planar linear elements are defined: line, arc and clothoid. Roads are represented in the vertical plane by linear elements line and arc. These are pre-defined by highway agencies as the geometrical road design elements. Each linear element is specified by references to a beginning and an ending point, as well as a fixed number of additional parameters. An **N:M** relationship is described between **Linear Geometry** and the both **Linear Element Vertical** and **Linear Element Planar**. This **N:M** relationship means that a linear geometry may consistent of many linear elements and a linear element may belong to many linear geometry objects. In order to realize the **N:M** relationship **LinearGeo/LinVer** and **LinearGeo/LinPlan**, association objects are introduced. The additional attribute **Linear Element Order** is defined in order to store sequences of linear elements in the linear geometry.

Between **Linear Element Vertical** and **Point Geometry** two aggregated **1: 0..*** relationships are defined. Every vertical linear element must have one beginning point and one ending point and every point may be assigned to none or many vertical linear elements. Relationships between **Linear Element Planar** and **Point Geometry** are the same as those for **Linear Element Vertical** and **Point Geometry**.

The planar linear elements have three parameter types; line, arc and clothoid. These parameter values are introduced by defining the object **Parameter**. There is association between the **Linear Element Planar** and **Parameter** objects. This **1: 0..*** aggregation means that a linear element planar may have many parameters and every parameter must be assigned to only one linear element. However, the **Linear Element Vertical** has only line and arc parameter types. Considering this, it is not required to model an additional parameter object for a vertical linear element. These are defined within the linear vertical element using the curvature value.

The geometrical element information is essential, since a parameterized solution using these elements is proposed in the conceptual data model. The geometrical elements of highway are defined uniquely with respect to their horizontal and vertical alignment elements. Although this information is widely used in highway administrations, in many cases it is not in digital form. Occasionally this information is available in digital form as Computer Aided Design (CAD) drawings. Although these can be exported into GIS using software vendor provided interfaces, user defined interfaces and Open DataBase Connectivity (ODBC), these methods are not yet fully satisfactory. In GIS-T, alignment elements are generalized, depending on map scales. Therefore, a reverse process can be performed in order to detect the alignment elements, with the help of adjustment techniques. Consequently, the usage of an automatic calculation technique, namely ***Detecting Alignment Elements,*** is proposed. In the conceptual data model, in order to detect alignment elements the ***Detecting Alignment Elements*** method of the **Linear Geometry** object is defined.

This method is based on the detection of alignment elements realized over significant parameters and elements, with the help of curvature diagram. If the beginning point and beginning tangent angle is known, with an approximation, alignment element parameters and their sequence can be uniquely defined. *[GRÜNDIG-I,1988]*

The curvature diagram is a graphical representation of the curvature (k), where (k) is defined with respect to stationing length (l) as;

$$k = \frac{d\tau}{dl} \tag{4.1}$$

Therefore, alignment elements in this diagram can be identified with simpler functions being; straight lines parallel to axis, straight lines not parallel to axis and quadratic parabola. Relationships between geometrical design blue prints, curvature and angle diagram are illustrated in Figure 4-4.

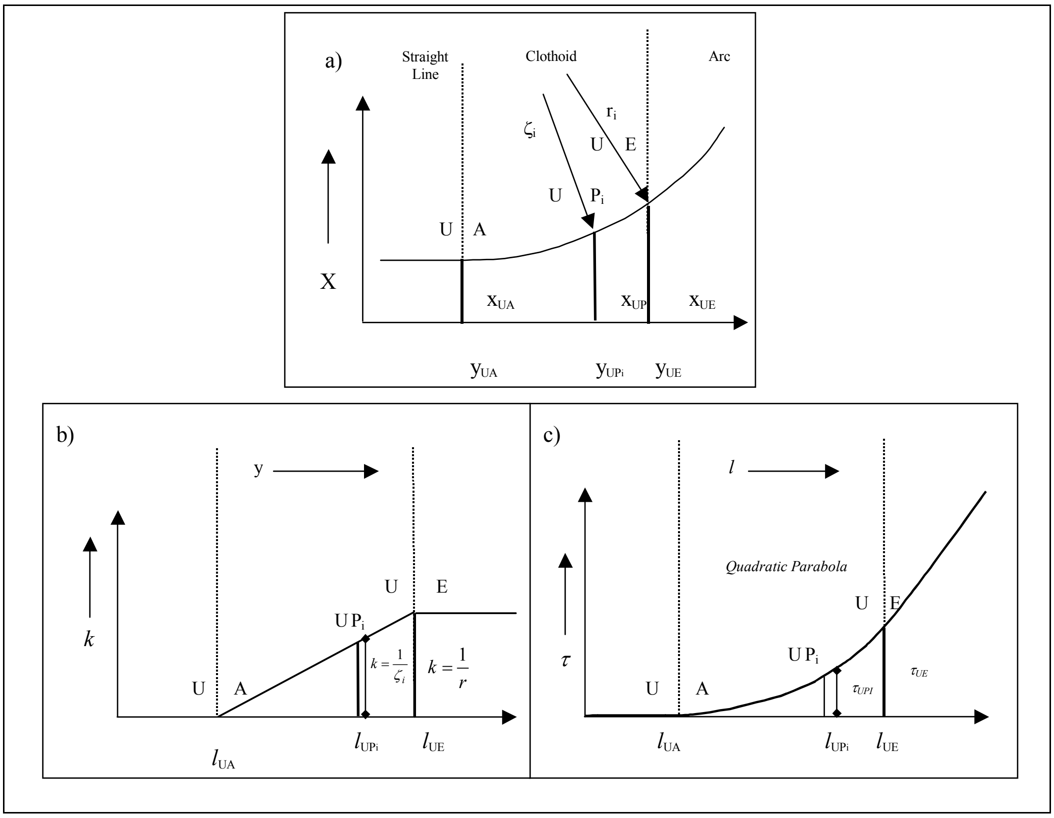


Figure 4-4: Relationships Between a)Blue Print, b)Curvature Diagram and c)Angle Diagram [MÜLLER,2000]

For each point of the alignment, where the bearing angle is a function of stationing value, the cartesian coordinates (x, y, z) of points of the horizontal (or vertical) alignment results detecting the alignment elements with the use of simple functions integration. Adjacent elements have to fulfil conditions of transition in order to enforce the smoothness of the alignment and of its first derivative. The approximation in the diagram of the first derivative of the alignment, corresponds to a spline analysis using parabolic curves of second order. *[GRÜNDIG,1992]*

The approximation of a sequence of points in the diagram leads to another task. It is necessary to find out the parabolic curve element to which the point has the closest distance. For every point, the bearing angle and the distance is required, where these information can be obtained automatically by means of adjacent points. Since the unavoidable very small distinguishing errors results undesired dispersions in the curvature diagram, during this process generalization effect is used. *[GRÜNDIG-II,1988]* In order to realize this task adjusted spline analyze with predetermined restrictions is used. Additional constraints for geometrical and driving dynamics are considered in the adjustment model as observations, in order to achieved the optimum result.

Obtained results needed to be optimized. In order to perform this optimization, parameters must be identified. The unique parameters for the straight line, arc and clothoid elements are;

Straight line:	l:	Length of a straight line
Arc:	R, l:	Radius and arc length
Clothoid:	R_P, R_f, l:	Radii of the arcs of the preceding and following element, the clothoid length

Any linear element between point a and point b can be uniquely identified with coordinates x_a, y_a, x_b, y_b and tangents t_a, t_b:

$$x_a + f(parameters) = x_b \tag{4.2}$$

$$y_a + f(parameters) = y_b \tag{4.3}$$

$$t_a + f(parameters) = t_b \tag{4.4}$$

With the use of available initial values, the functions of (f) can be linearized.*[BAHNDORF, 1994]*. A linear substitute system results which can be solved minimizing a weighted squared sum of residuals of the parameters in a least squares way. As result of analysis :

- Sequence of alignment elements, element type, radius, stationing values
- Coordinate list of alignment main points with approximation values, tangent bearings and stationing values for mentioned points

were obtained.

Points sequence, which are representing the road, is required in order to implement this method. This information is available in digitized maps and can be efficiently used. Beginning point and beginning tangent angle can be even obtained from large scaled digitized maps. However, the frequency of digitized points can be insufficient in some places, in order to detect main points of alignment elements. The appropriate frequency can be determined, with sampling theorem. A calculation according to Nyquist-frequency (f_N) is possible.

$$f_N = \frac{1}{2\Delta t} \tag{4.5}$$

(Δt)can only contain full information over the process, if no frequency is bigger than Nyquist-frequency (f_N) (TAUBENHEIM 1969) *[KULMANN,1996]*. In our case instead of time variable (Δt), the length variable (Δl) needed to be considered. The required frequency can be successfully achieved using integration of highway agencies rich data sources by means of adjustment techniques. This additional sources include; GPS data, geometrical road design regulations, digital terrain model and orthophotos. The Figure 4-5 illustrates an example where planar linear elements have been obtained by integrating different sources such as digitized highway network and geometrical design regulations, using the ***Detecting Alignment Elements*** method.

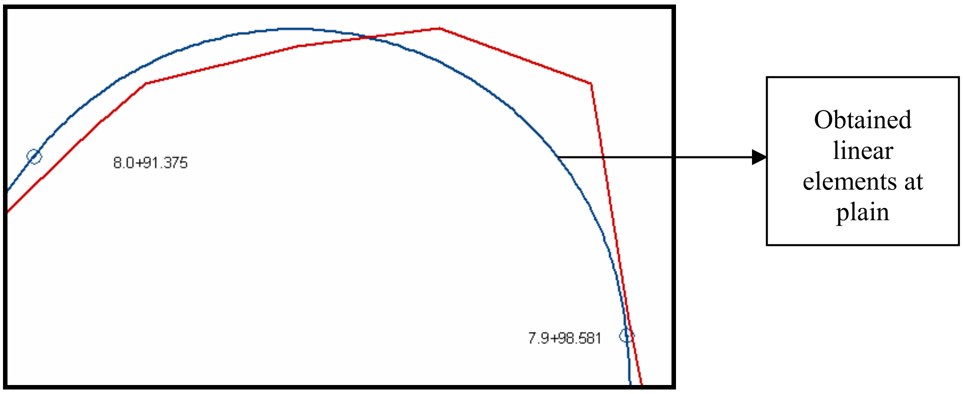


Figure 4-5: Integration of Different Data Sources

There is an **N:M** relationship between **Area Geometry** and **Linear Element Planar**, which is defined using **AreaGeo/LinPlan**. **AreaGeo/LinPlan** also contains the sequence of linear elements for additional requirements. A similar relationship is also defined between **Area Geometry** and **Linear Element Vertical** using the **AreaGeo/LinVer** association object.

Geometrical elements defined in the conceptual data model are shown in Figure 4-6.

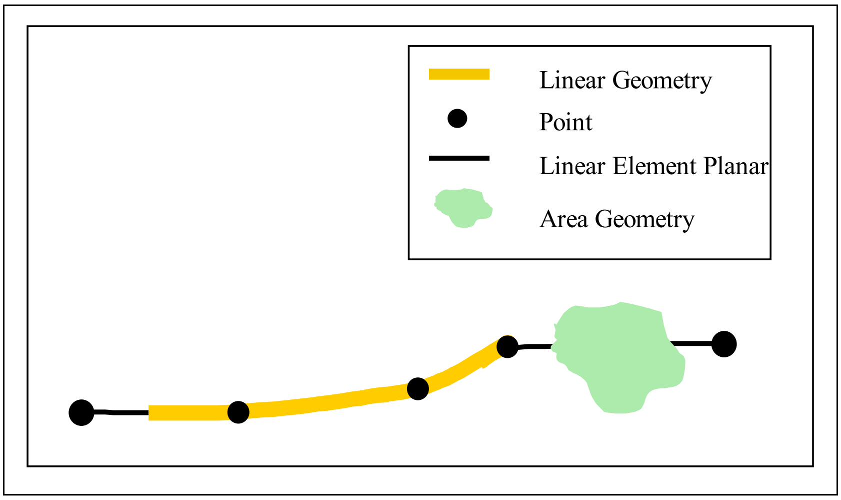


Figure 4-6: Geometric Elements in the Plane

4.2.3.2 Topology

Three main issues concerning the topology need to be considered during the conceptual data model design. These are:

- The relationship between topology and geometry.
- Topological abstraction levels.
- Non-planar topology.

With respect to the first issue, although topology is an abstraction of geometry, in many GIS data models topology and geometry are combined, such as in the example shown in Figure 4-7. In this figure only the red line is available in GIS-T systems, representing both geometry and topology. In spite of some advantages, such as the reduction of total data storage requirements and the increasing of query processing performance, many disadvantages arise with data maintenance. This is especially the case when there are geometrical changes, a situation which is very common in highway administrations. Since the geometry and topology are combined, displacements in geometry also effects topology. Although, topology is invariant under position, orientation, transformation, shape and size, data maintenance is required both for topology and geometry after every geometrical displacement. Additionally, the topology needs to be related to one-dimensional, two-dimensional and three-dimensional space. In order to avoid the above mentioned problems, topology and geometry are separated in the proposed model. Topology is considered as a logical abstraction of geometry.

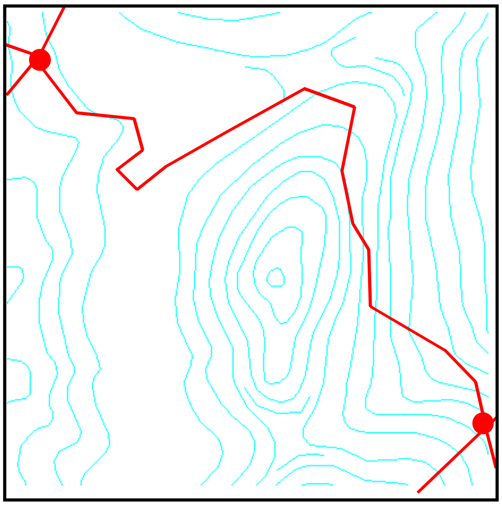


Figure 4-7: Road Segment in GIS, with Combined Topology and Geometry Information

Topology abstraction levels, which is the second issue, has not yet been considered in GIS-T data models. However, since there are many diverse applications in highway agencies, multiple representations of topology are required. Two topology abstraction levels are defined in the conceptual data model. These two topology abstraction levels are shown in Figure 4-7.

Figure a: An Air-photo, Representing Reality

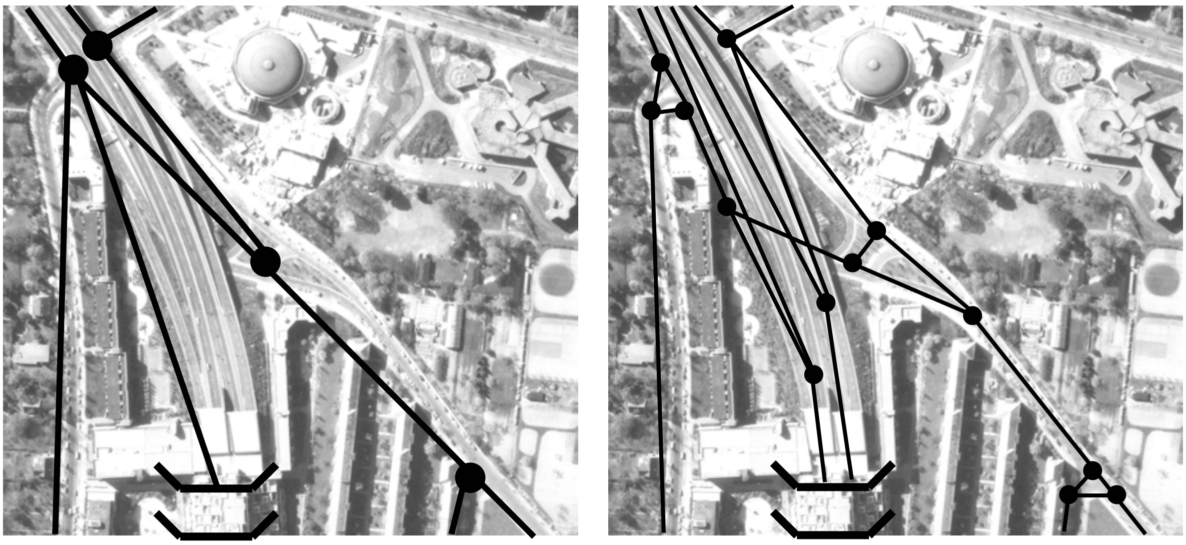


Figure b: I-Level Topology *Figure c: II-Level Topology*

Figure 4-7: Different Abstraction Levels of Topology

In order to highlight the necessity of multiple topology abstraction levels, divided highways provide a good example. Since topology is a logical abstraction of reality represented by geometrical elements, in the case of divided highways in many cases two different geometry exist for each.

Without a second abstraction level, the divided highway illustrated in Figure 4-8 would be modeled with one link, representing geometry a and b. Additionally, every piece of thematic information belonging to geometry a and b should be combined and assigned to this one link. As a result, the geometrical characteristics of thematic information, such as traffic accident location can not be identified. This results in a loss and/or mismatching of information in the data model. In GIS-T, divided highway information is not considered to be a basic requirement. However, it is certainly true that the modeling of divided highways is inevitably required by all highway administrations, therefore, the conceptual data model must fully support multiple topological abstraction levels. When this becomes the case, the second level topology can be implemented without any radical change in the data model. This is also required for analysis such as; route planning and intelligent transportation systems (ITS).

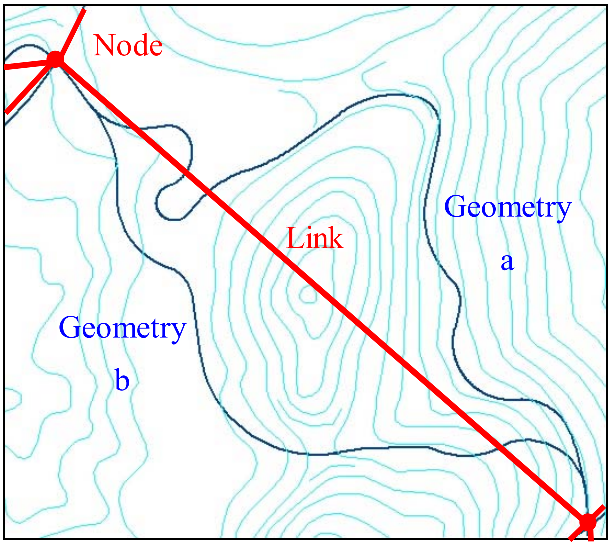


Figure 4-8: Sample of a Divided Highway

The third issue is non-planar topology. In the conceptual data models evaluated, all topologies were planar. With planar topologies links cannot cross each other without creating an intersection. The crossing links must therefore be split into several individual links. Planar networks have many advantages, principally that they are common and simple. However, this does not reflect the reality for road networks where links can cross without creating intersection. This is the case for example with bridges, tunnels, overpasses, underpasses and viaducts.

Highway administrations commonly use linear referencing methods including the link-node method. Therefore, there is a necessity to differentiate "real" nodes from "virtual" nodes which are generated due to the usage of planar topology. Due to this, as well as other reasons which were described earlier, virtual nodes were introduced by all the conceptual data models. However, this solution is inefficient in practice due to the high level of data collection requirement. In the conceptual data model developed, non-planar topology was implemented in order to avoid these problems.

In order to implement such non-planarity, use is made of the third dimension namely height information. With the designed objects **Linear Element Vertical** and **Point Geometry**, and **Linear Geometry** object methodology ***Detecting Alignment Elements,*** non-planar topology third-dimension is achieved in the conceptual data model. Since the geometrical vertical alignment elements were detected using the ***Detecting Alignment Elements***, vertical alignment geometry information is available as "build-in", where these are illustrated in Figure 4-9.

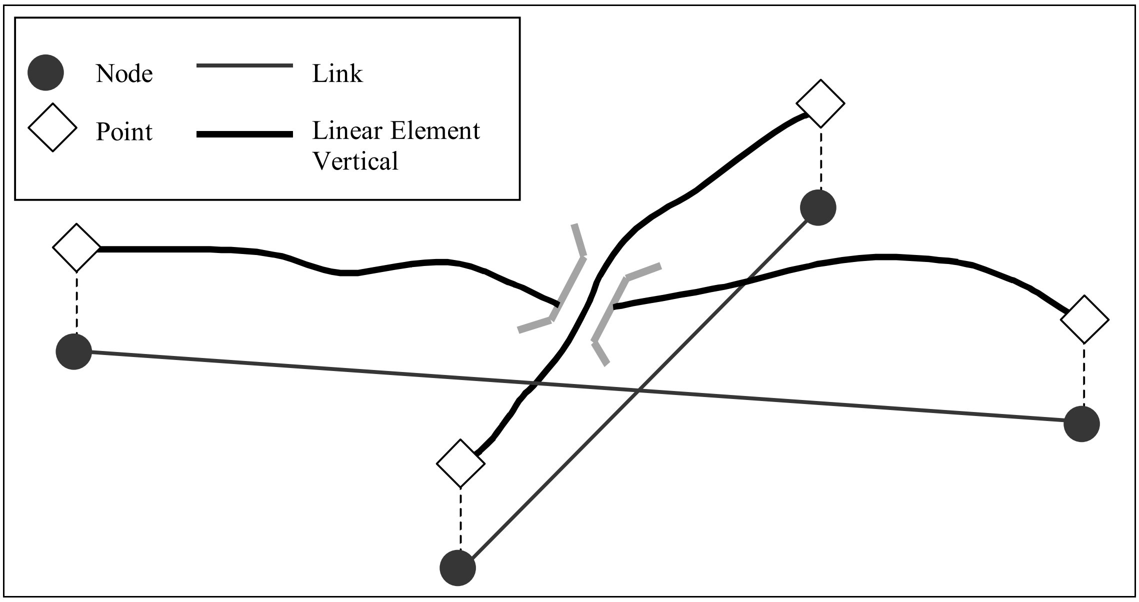


Figure 4-9: Non-planar Topology

In order to implement this information, there are diverse height information sources, due to departments' usage purposes. Since the relative height information is generally measured by means of surveying techniques, mainly leveling and real time GPS, obtained accuracy is adequate to fulfill user requirements such as; freight transportation and ITS technology. Several other solutions, which adjustment techniques can be applied, in order to achieve non-planar topology was introduced in Section 4.3.3.

The topology component of the proposed conceptual data model is shown in Figure 4-10.

The main elements of topology are **Node** and **Link** as described in Chapter 2. According to these definitions two different relationships between **Node** and **Link** are defined. Firstly, in order to define a link, beginning and ending nodes are required. Secondly, a node can be assigned to many links. These relations are shown in the data model with two **1: 0..*** aggregation associations. Associations between **Node** and **Link** are applied to both abstraction levels, using the same aggregated associations.

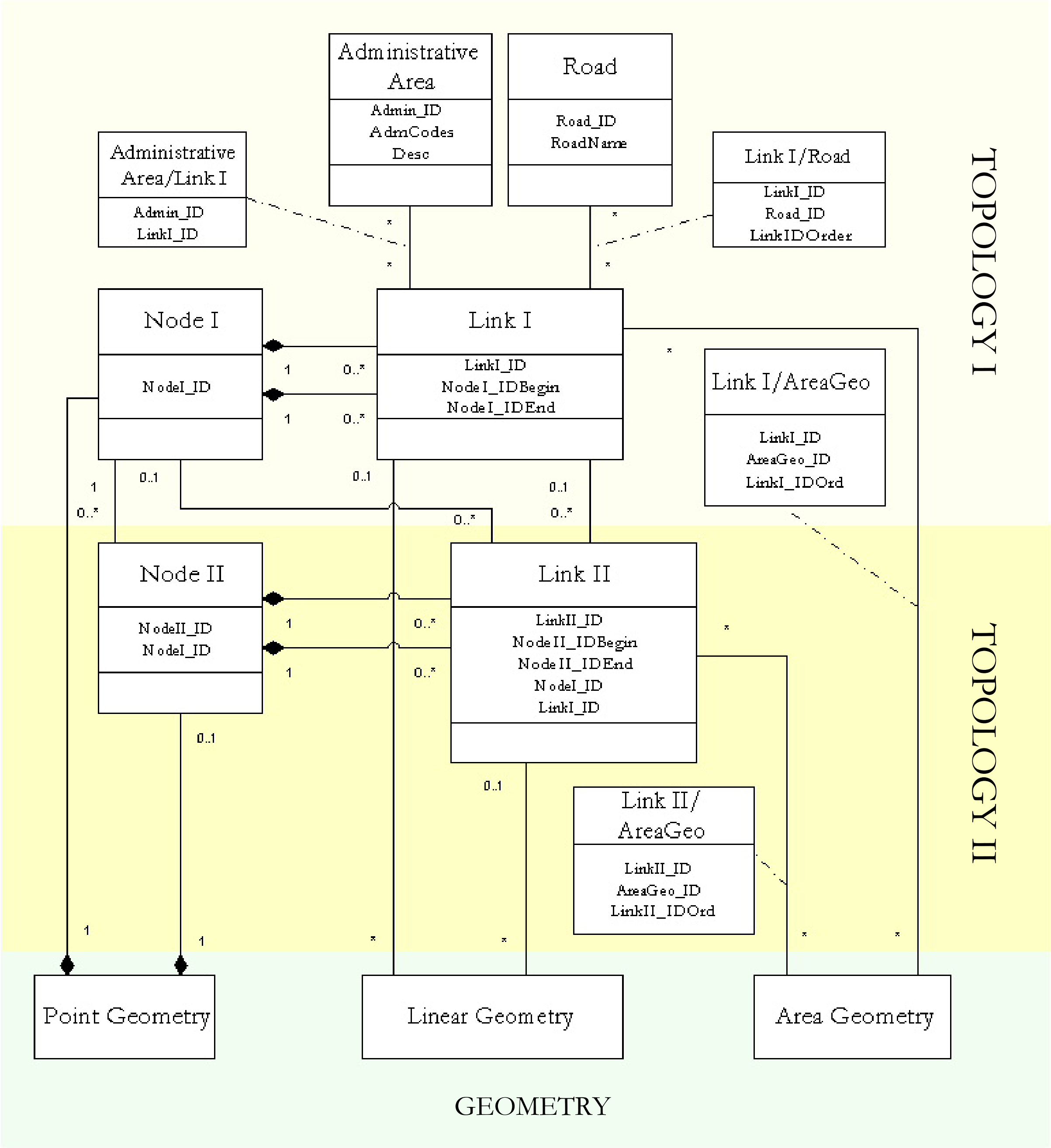


Figure 4-10: The Topology Component

Relationships between two abstraction levels are modeled as follows:

1. **Node I**, being a higher abstraction level may be composed of **Link II** and **Node II**. Between the first level topology element node (**Node I**) and second level topology element node (**Node II**) a **1: 0..*** relationship is assigned. This relationship maps reality adequately because **Node I** may be composed of many **Node II** and every **Node II** is assigned to the first level topological object **Node I**.
2. The relationship between **Node I** and **Link II** is modeled as **0..1: 0..***, where **Node I** may be composed of second level links (**Link II**) and a second level link (**Link II**) may be assigned to **Node I.** Road junctions are examples of such situations.

3. The relationship between **Link I** and **Link II** is modeled as **0..1: 0..***, where **Link I** may be composed of second level link's (**Link II**) and a second level link (**Link II**) may be assigned to **Link I**. The merging and subsequent separation of divided highways is an example of this.

Using the above defined relations, other required information can be extracted. Relationships between topological and geometrical components are described using the following associations.

The **Node** object is represented at the geometrical level by a point. **Node I** and **Node II** have a **0..1: 1** relationship with respect to **Point Geometry**. A node must be represented with a point, but a point need not be a node. This relationship is valid for both abstraction levels. **Linear Geometry** has a relationship between the two abstraction levels; **Link I** and **Link II**. Between **Link** (**Link I** and **Link II**) and **Linear Geometry a 0..1: 1..*** relationship is assigned. A link may be composed of many linear geometry and a linear geometry may be assigned to one link. The **N:M** association between topological element **Link** and **Area Geometry** is realized using association tables **LinkI/AreaGeo** for the first level, and **Link II/AreaGeo** for the second.

External objects, administrative areas and road objects, they are assigned to the topology component **Link I**. The object **Administrative Area** include governmental borders, highway division borders or national borders. The **Road** object only stores road names or codes provided by highway administrations, as attribute. Both objects are only associated with first level topology. The **Administrative Area** and **Link I** objects have an **N:M** association, realized using the **AdministrativeArea/LinkI** table. In the data model links are not dependent on administrative areas. A link may be assigned to many administrative areas and an administrative area may contain many links. The **Road** object has an **N:M** association with **Link I**. A link may have many names and/or codes. Many roads may be identified by one link. One road may reference many links. The association table **Link I/Road** is introduced in order to implement this **N:M** association. It also stores the sequence of links along the road in order to simplify possible queries.

The topology and geometry components introduced are illustrated in Figure 4-11.

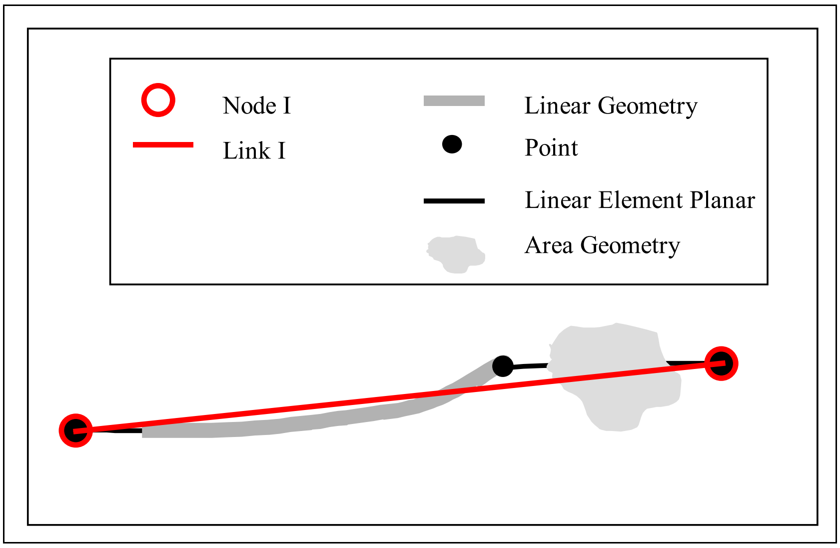


Figure 4-11: The First Level Topology and Geometric Elements

4.2.3.3 Road Events

The **Road Event** object stores all non-spatial information including the attributes, occurrence and physical components of the road. The **Road Event** is assigned to the geometrical objects;

Point, **Linear Geometry** or **Area Geometry**. Examples of typical road events are: accident, project, road facility, video image, pavement type and road type. Road events could be assigned to:

- A **Point Geometry** such as a traffic sign, a traffic accident or a maintenance workstation location.
- A **Linear Geometry** such as project information.

In addition to these, events can occur at the same time and in the same location. In this component of the conceptual data model **Road Event**, **Road Event Properties** and associations between the geometry components are introduced. These are illustrated in Figure 4-12.

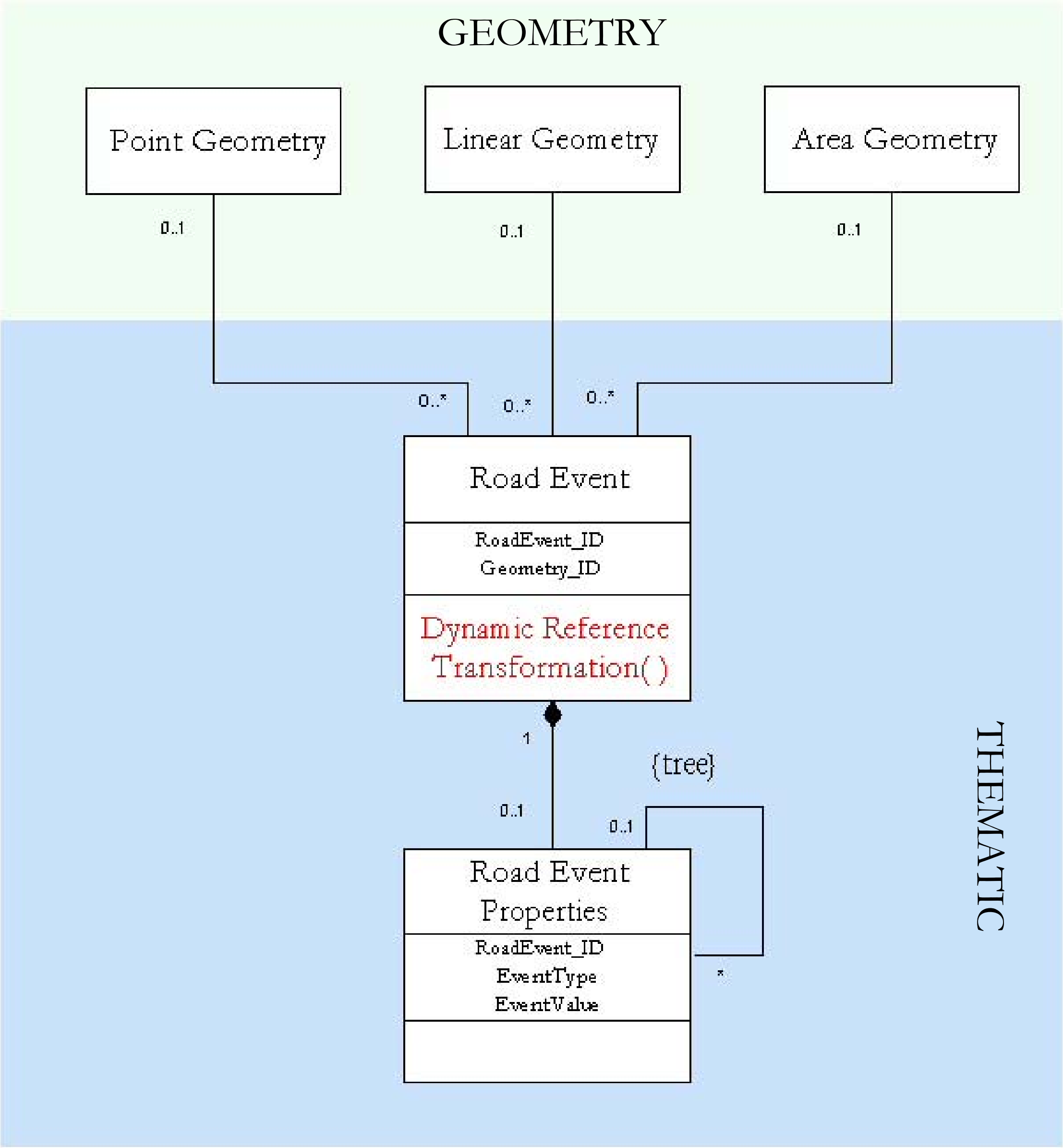


Figure 4-12: The Thematic Component of the Conceptual Data Model

In the Road Event object all non-spatial information needed to be stored according to the provided road event identifiers in the event catalogue. All the road events with their associated geometrical types, needed to be identified and documented in this catalogue in a standard way agreed throughout the entire highway agency. Having agreed on the standards for identification, associations between road events and geometrical elements can be made.

There is a **0..1: 0..*** relationship between the **Geometry** (**Point Geometry**, **Linear Geometry** and **Area Geometry**) and **Road Event** objects. These relationships mean that every road event should be assigned to one of the geometry types and many road events may be assigned to one geometrical object.

The **Road Event** object has a very important method, called ***Dynamic Reference Transformation***. This method is defined to transform linearly referenced road data into three-dimensional coordinates, as well as to re-transform it back into one of the user linear referencing methods, if required. Since this information is stored in a three-dimensional system, it can be transformed to any one dimension system. Usually, in highway agencies this will be a linear referencing system. This methodology is explained in Section 4.3.1.

All non-spatial properties of **Road Event** are stored in the object **Road Event Properties**. This object has a tree structure with which it is possible to expand properties of road events to any desired level. An association between **Road Event** and **Road Event Properties** is introduced here as **1:0..1**. This means that every road event may have many road event properties, or the necessary information is provided in the catalogue and every road event property must be assigned to only one road event.

With the objects introduced above, the requirements of highway agencies are satisfied in multi-dimensional space, including planar and vertical sections. Additionally, non-spatiality of road data is achieved.

In Section 3.2.2 four spatial reference systems were described. The remaining system which has not yet been introduced to the conceptual model is cross-sectional system (h, q). This can not be realized using just the introduced objects. The cross-section reference system is discussed in Section 4.3.2 in more detail. In order to realize the cross-sectional spatial information, five other objects are introduced into the data model. These are shown in Figure 4-13.

In order to realize this task, the **Road Event** object is divided using a generalization relationship into two objects. Generalization relationships declare that objects are fully consistent with the super-class. These are **Dimensional** and **Non-dimensional**.

In the object **Dimensional**, the cross sectional reference system and the reference system parameters are identified. After the determination of the reference system, the existing three dimensions are introduced. These are named; **Zero Dimensional**, **One Dimensional** and **Two Dimensional**. Between the object **Dimensional** and the objects **Zero Dimensional**, **One Dimensional** and **Two Dimensional** there exist **1: 0..*** associations. These associations mean that every **Dimensional** object may have many **Zero Dimensional**, **One Dimensional** or **Two Dimensional** objects, and every object must be associated with a dimensional object.

Other associations are defined:

a. Two **1: 0..*** aggregated associations between **Zero Dimensional** and **One Dimensional** objects exist. This means that every **One Dimensional** object is defined by a beginning and ending **Zero Dimensional** object, and a **Zero Dimensional** object may be assigned to many **One Dimensional** objects.

b. The association between the **One Dimensional** and **Two Dimensional** objects is **N:M**, realized using the association table **Onedim/Twodim**.

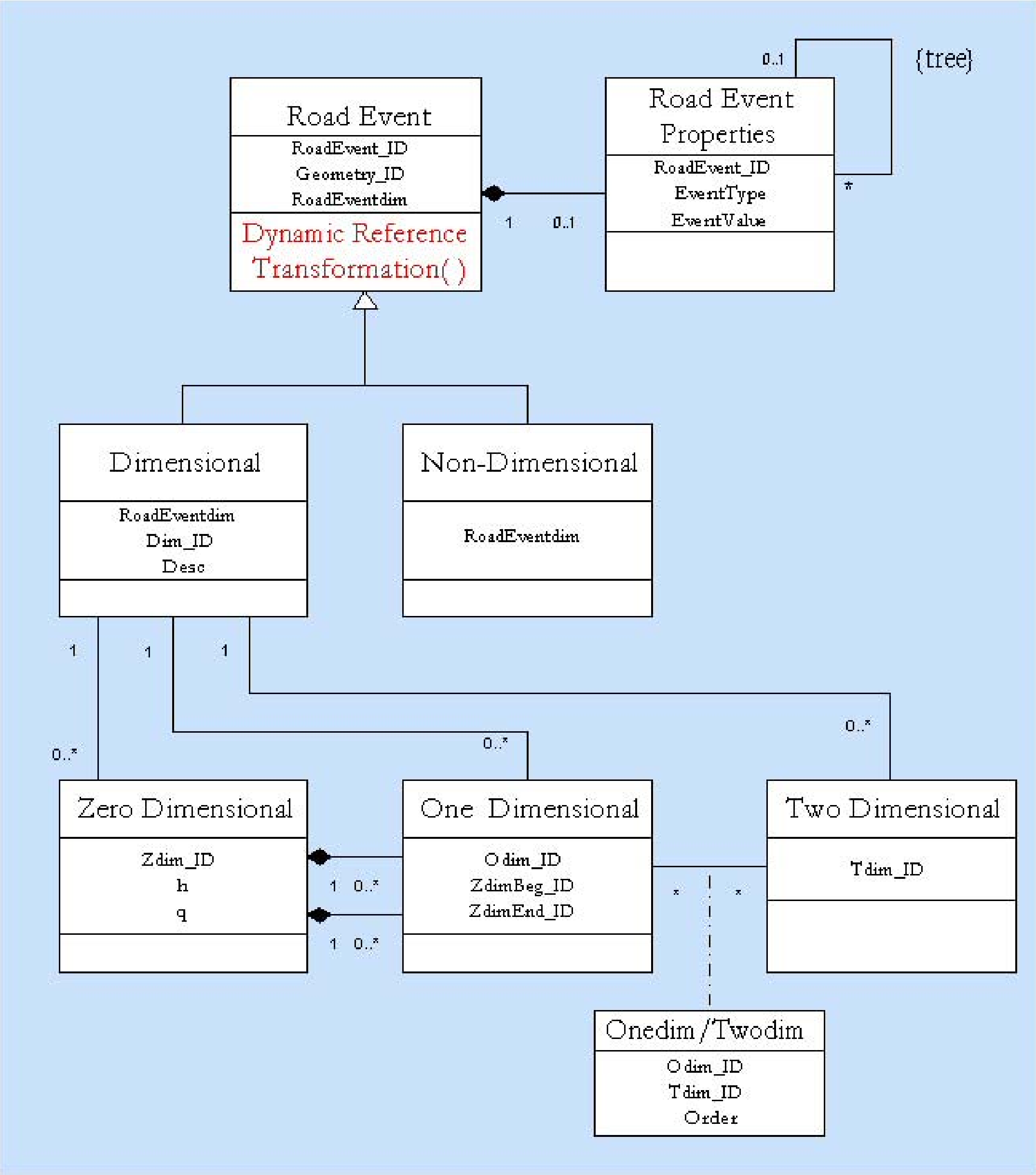


Figure 4-13: The Conceptual Model For Cross - Sectional Spatial Information

After having introduced the above objects to the conceptual data model, the reference systems described in the information analyses Section 3.2.2 can now be fully mapped. This is shown in Table 4-1 .The two empty boxes are obtained using other associations.

Table 4-1: The Matrix of Road Reference Systems and Geometry

	X, Y, Z	**h, l**	**h, q**
Point	☑	☐	☑
Line	☑	☑	☑
Area	☑	☐	☑

An overview of the proposed conceptual data model is shown in Figure 4-14. In order to simplify the diagram, only those specifically generated methods, ***Dynamic Reference Transformation*** and ***Detecting Alignment Elements,*** are shown in the diagram.

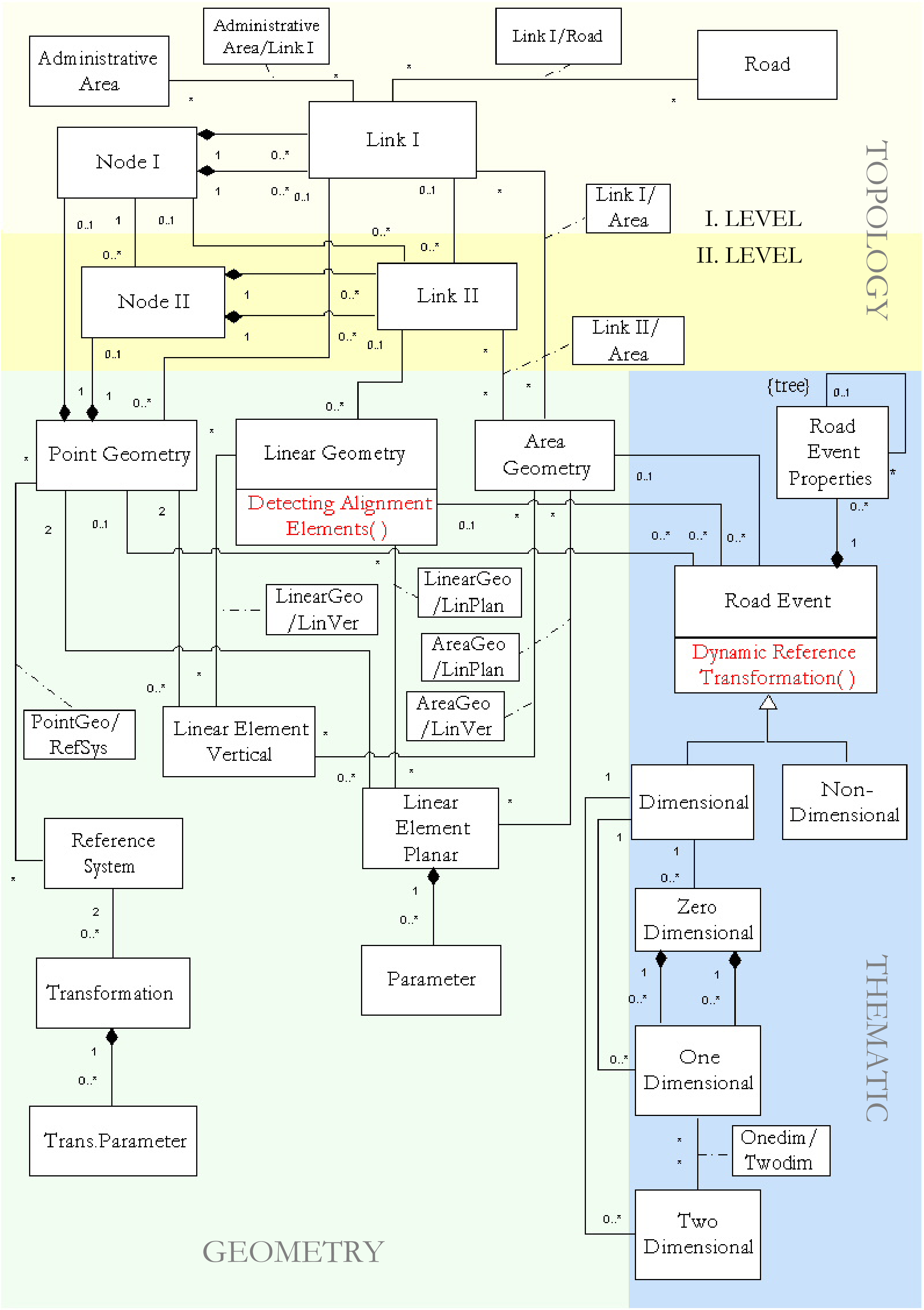


Figure 4-14: The Overview of the Proposed Conceptual Data Model

In the proposed conceptual data model; topology, geometry and thematic data are modeled independently. Using two abstraction topology levels, multiple topologic representations may be realized. In addition, with relationships assigned between objects, multiple geometry can be associated with a single topology. Although planar topology is simpler to both create and maintain, a non-planar topological model was implemented with newly introduced height information and required method in order to better model reality and to appropriately satisfy user requirements. Various spatial characteristics of road data, such as linear referencing systems, are non-redundantly modeled using the introduced methods. Agency business rules are introduced into the conceptual data model. Thematic information is modeled separately from the geometry and topology on the basis of a detailed examination of the spatial characteristics of the road data structure. The modeling process is carried out independently of software vendors proprietary systems.

By applying the proposed conceptual data model:

- All the varied highway agency views were implemented.
- Geometrical, topological and thematic information was modeled transparently.
- Spatial information was referenced using the geometry element: point.
- Multiple topological levels are designed.
- Non-planar topology is implemented.
- A modular structure was designed.
- A dynamic transformation between three dimensional and one-dimensional coordinate systems was defined.

4.2.4 Completing Conceptual Data Model

In order to simplify the conceptual data model diagram, the metadata components of; integrity constraints, history and quality will be considered separately.

4.2.4.1 Integrity Constraints

Several agency rules, for instance the relationship between **Administrative Area** and the topological element **Link**, have already been modeled in the conceptual data model, others have not been defined yet. In order to implement agency rules in databases, a variety of techniques are provided based on consistency conditions and integrity constraints.

Constraints are functional relationships between objects, object attributes and associations. *[RUMBAUGH, 1991]* In general, two types of integrity constraints can be distinguished. These are static integrity constraints which define the valid state of a database, and dynamic integrity constraints which are the conditions on the allowable transitions from one database state to another. *[MOLENAAR,1998]* While static integrity constraints validate completeness and correctness of the current data in the database, dynamic integrity constraints are used during local updates.

In GIS, due to redundancy, integrity constraints are required wherever geometry interacts. Due to the limitations and requirements of current GIS software for data modeling, many additional integrity constraints are required in order to validate geometry and topology. For example, defining geometrical elements via their parameters would avoid many redundancies. This is because only beginning point, ending point and parameter values are required to generate the linear element. However, these parameterized elements cannot be visualized in GIS. Since, every visible feature in the system has to be stored with its geometrical coordinates. In this case, for visualization, the point coordinates of the polygon need to be accessed to provide the geometry of the linear element. Therefore, additional methods need to be designed in order to satisfy user requirements and control the redundancy.

In Chapter 2, encapsulation concept was described, as permitting the design of spatial and non-spatial data as well as methods within each object and importance for GIS was highlighted. However, with relational and object-relational databases the encapsulation concept is not yet fully realized. As a result, these required methods need to be defined separately and controlled by the user.

Especially in the case of highway administrations, the maintenance of the spatial aspect of the data in the conceptual data model is difficult. This is due to the complex relationships between objects, the complex business rules and other prerequisites. The spatial aspects of information must be validated whenever objects are updated. This task is automated when the defined methods are available. Relationships between geometry and topology can be used to formulate consistency constraints for spatial databases.(Hadzilacos and Tryfona 1992, Kainz 1995 and Plümer 1996). *[MOLENAAR,1998] [GRÖGER,2000]* During the transaction process for topological updates, three actions are possible. These transactions are insertion, deletion of a link and changing the geometry which the link is assigned to.

In Figure 4-15 an example of a simple operation is shown. A node has been displaced due to data quality improvements or correction of errors. Therefore, the possible discontinuities and errors in the topological and geometrical elements must be controlled.

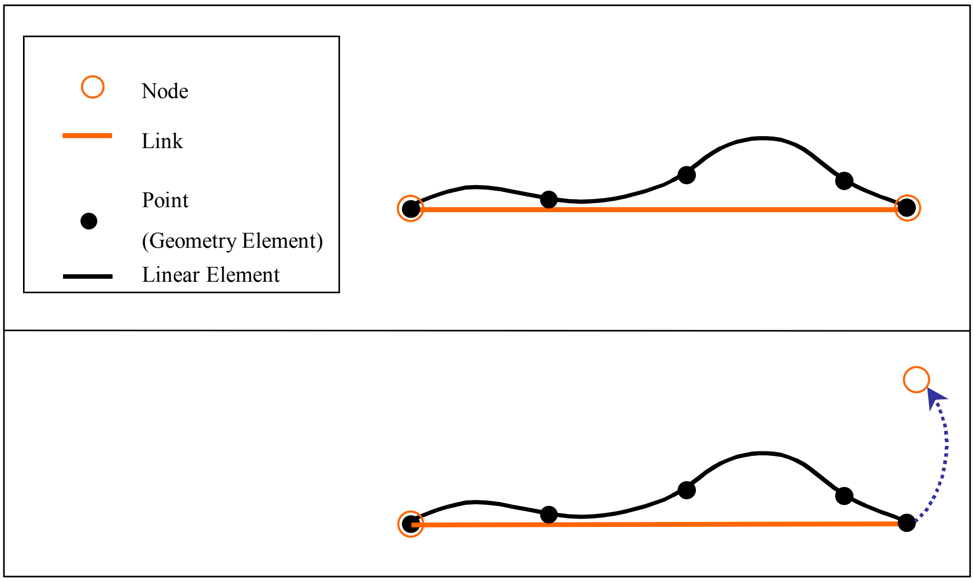


Figure 4-15: A Transaction With Sample Objects

The transaction of inserting a link is as follows:

Transaction: Insert a link

Conditions:

a. Two nodes, the beginning node and ending node, are required. According to the conceptual proposed data model it is not possible to insert a link without the prior existence of the two nodes.

b. The beginning point of the first geometrical element and the ending point of the last geometrical element should correspond with the beginning and ending nodes.

Action: The link is inserted as a logical connection between both the two nodes and the corresponding geometrical element points.

The transaction of deleting a link is as follows:

Transaction: Delete a link

Conditions:

a. Two nodes are required in order to define a link.
b. A node may be assigned to many links.
c. A link has an association between geometrical elements.

Action: Search for the beginning and the ending nodes of the link. If either of these nodes are assigned to only this link then delete the node in question. No action is taken for nodes which are assigned to more than just this link. Search for geometrical elements assigned to this link and delete them.

The transaction of changing the geometry which the link is assigned to;

Transaction: Change geometry of link

Conditions:

a. Topology is invariant of geometry.
b. The beginning point of the first geometric element and the ending point of the last geometric element must correspond with the beginning node and the ending node respectively.

Action: Identify in which element geometrical change takes place.

- If the geometrical displacement takes place at the first or the last element, determine whether the beginning point of the first element or the ending point of the last element is changed.
    i. In the case of a displacement which modifies a point assigned to a node, the user must be warned of the necessity of running other consistency checks, such as geometrical element consistency.
    ii. In case where the displacement is of a point which is not assigned to a node, no action is taken.
- If geometrical displacement is not in the first or in the last geometrical element take no action.

Some other examples of integrity constraints used in the implementation are provided in Chapter 5.

4.2.4.2 History

Highway administrations need to be able to track the changes in the system over time. This is important for the preparation of documents such as annual reports or yearly plans. It is additionally necessary to be able to reconstruct the state of the system at any specified point in time. The visualization of such “rolled-back” system states is essential.

In the conceptual data model, it is proposed to model history as an object. This means the history for all objects, including relationships between each other, are stored in one history object. The transaction log approach is adopted. *[DORSEY,1999]*. By using the history object it is possible to report or re-create a required transaction.

In order to track the object history in the conceptual data model, the additional objects; **History**, **Value Range**, **Event Properties** and **Event Type** were introduced. In order to illustrate the approach used, the **Link I** object from the conceptual data model is presented as an example.

The object **Link I** has three attributes:

- **Link I** identifier, data type declared as long and name LinkI_ID, since link identifier is not automatically generated, but externally identified.
- The beginning node identifier (**Node I**), which is assigned to **Link I**, data type declared as long and name as NodeI_IDBegin
- The ending node identifier (**Node I**), which is assigned to **Link I**, data type declared as long and name as NodeI_IDEnd

The **Link I** object is represented in Figure 4-16.

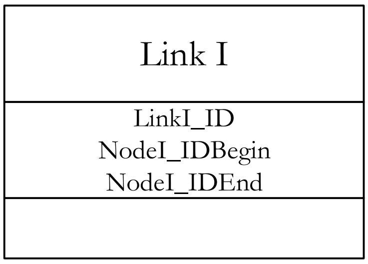


Figure 4-16: Attributes of the Object ***Link I***

Objects, **History**, **Value Range**, **Event Properties**, **Event Type** and their relationships are shown in Figure 4-17.

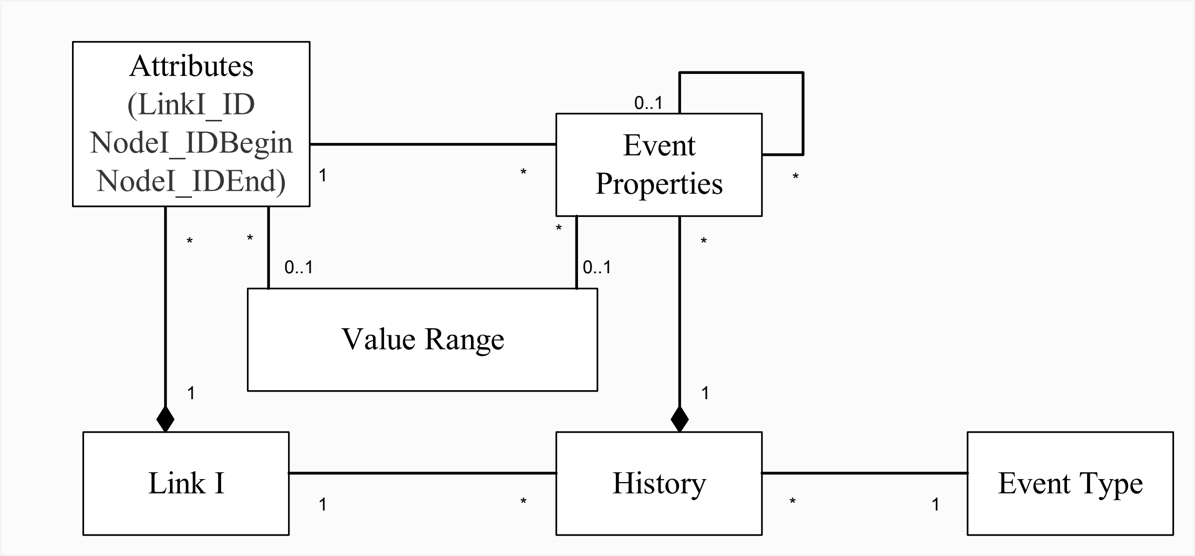


Figure 4-17: The History Component of the Conceptual Data Model

The **Link I** object has a **1:N** relationship between its attributes. A **1:N** relationship between **Link I** and its attributes (LinkI_ID, NodeI_IDBegin, NodeI_IDEnd) means that every object must have at least one or many attributes, and every attribute must belong to only one object. The **Value Range** object contains a pre-defined set of attributes. In the case of **Link I,** these are the set of identifier values defined by the highway agency. A **N:0..1** association exists between **Attributes** and **Value Range.** This means one attribute may have at most one pre-defined value range, such as link identifiers, and a value range can be defined for many attributes. The **Event Properties** object stores only the altered values. Additionally other required information such as operator details and time records, which are required in order to control sequence of history event and reconstruct it when is it necessary, is kept in this object.

The **Event Properties** object has a tree structure which can be expanded or collapsed as required. The association between **Range Values** and **Event Properties** are mapped as **N: 0..1**, the same as the association between **Attributes** and **Range Values.** This is because the character of the information has not changed, only the attribute value was altered. In the **History** object, the history identifier, which should be automatically uniquely generated, and the object identifier, in this case **Link I**, are stored as the attributes. The association between **History** and **Event Properties** is **1:N**. This means that a **History** has many event properties, and every **Event Property** must be assigned to only one **History**. In order to track the changes, the **Event Type** defines which type of action has been performed, such as updating or deleting. A **1:N** association exists between **Event Type** and **History.** One event type can have many history objects, for instance an operator can update many objects at one time (before the transaction), and every history object is associated with only one event type. All these processes are controlled by transaction rules.

4.2.4.3 Quality Aspect

Due to the ever increasing requirements for data integration and data exchange, data quality aspects have become more important. For all data and applications, a minimum required quality criteria exist, such as positional accuracy or specific semantic attributes. These criteria are generally dependent on the use of data and defined by the administrative regulations, data collection techniques, data structure characteristics as well as standards. However, during evaluation of the current situation in highway agencies, it was identified that information stored in the system can be inaccurate; in many cases attributes are missing or incorrect. There are many reasons for this situation, such as:

- Lack of data maintenance and an appropriate maintenance tools.
- The high economic cost for data acquisition.
- Information obtained from varied sources can differ when it has not been structured and integrated.
- Automation has not been fully applied, especially in the areas of data entry and acquisition.

In addition, due to variations in user requirements and in the generalization level of reality, the required data quality also differs considerably. Therefore, it is necessary to provide existing information quality, in order to evaluate the appropriateness of analyses and to guide other possible users. As a result, data quality certification is highly attractive to highway agencies, especially with respect to data integration and data exchange. This needs to be solved in an integrated GIS-T approach for each entire agency. Only by doing so will the user's expectations be fulfilled.

In the conceptual data model the quality aspect was implemented using member methods of the individual objects. The quality aspect was not modeled in the proposed conceptual data model in the same manner as was done for history. This is because it is very unlikely that errors or poor quality data needs to be regenerated. Using this approach the quality of the current data is reproducible at any time in the form of documents or tables. As an example, in the proposed conceptual data model, the standard deviation values ($\sigma_x, \sigma_y, \sigma_z$) were stored in object **Point** as attributive information and appropriate methods were provided in order to regenerate positional accuracy using adjustment techniques.

4.3 Special Aspects of Highway Information System

4.3.1 The Linear Referencing System

Depending on the required use, the real-world phenomena road is multi-dimensionally defined:

- one-dimensional, linear referencing system
- two dimensional, planar coordinates
- three dimensional, planar coordinates and height information
- four dimensional, time in the case of dynamic objects

However, in many cases this variety is not fully supported by GIS-T. This is due to the fundamental problems already highlighted including the lack of an adequate analysis of the road information structure or limitations in the GIS software. One of the main aspects discussed in this study is the integration of one-dimensional data with three-dimensional data.

Some of the main problems in the management of linearly referenced data include;

- The geometrical state of highways is subject to progressive refinement. During this process all linearly referenced data changes its reference due to the new geometry and it is possible for existing data connections to be lost. An example is illustrated in Figure 4-18. After realignment, the traffic accident at kilometer (2.2+33.314) would appear to have occurred along the new stretch of road. This is completely incorrect, and because the location of the accident is now not part of the road network, the road event must be deleted from the current system. Similarly, the traffic accident at kilometer (3.4+63.886) would "move" if its linear reference were not to be updated. It must therefore be corrected according to the new kilometer value. Additionally, other situations such as; corrections in length due to improved accuracy following re-survey, or topological modification also cause similar data maintenance complications.

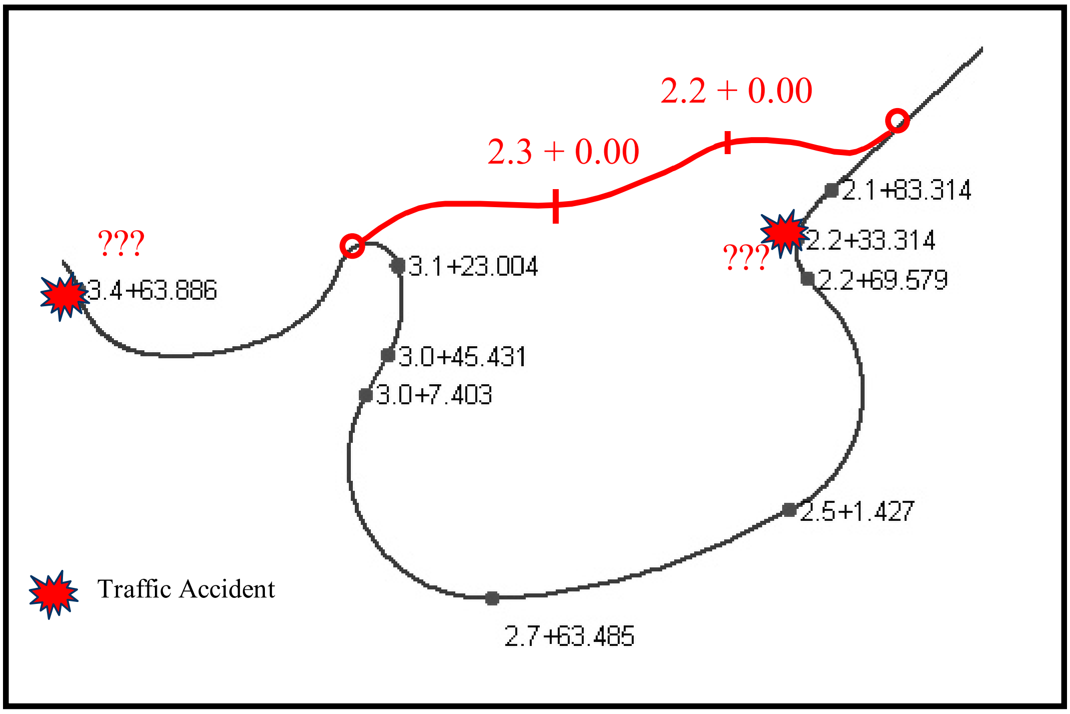


Figure 4-18: The Road Information After Re-alignment

- There is no agreement or common linear referencing methodology for road events. The linear referencing systems used varies from country to country, even from road department to road department. This results in communication and data exchange problems within highway agencies, due to not integrated various methods.
- Reference points in the field are often poorly maintained due to high costs.
- The more accurate, rapid and economic solutions provided by techniques such as; GPS and photogrammetry, cannot be easily combined with linear referencing methods due to their three-dimensional nature.
- Enormous effort and planning is required in order to keep the road data consistent.

Efficient usage of GIS-T can only be realized when the different methods applied for linear referencing are integrated. These include; route-mile-point, stationing, link-node, control-section, route-reference-post offset, and route-reference point-offset. Users unfamiliar with linear referencing systems have difficulties using the provided data, leading to many problems if data is planned to be shared with other organizations such as the ministry of health, national police, cadastre and the private sector. However, road administrations point out that even if other data acquisition techniques such as GPS and remote sensing become widespread, linearly referenced data collection will remain common. This is because the linear referencing method is the most "natural" way of collecting linear object road information.

Several solutions exist in order to overcome this situation. The current solutions include:

1. Generic road data models
    - Beginning with the initial NCHRP 20-27 project, the research evolved into the GIS-T Linear Referencing Pooled-Fund Study (Fletcher et al., 1995), and then to the current NCHRP 20-27(3) project which proposes guidelines for the implementation of multi-modal transportation location referencing systems.
2. Framework road data models
    - The NSDI framework is intended to enable integration of disparate data sets. A more formal model for classifying road features has been proposed by Dueker and Butler.
3. Location approach
    - An alternative approach can be characterized as a location (geometry) approach *[SUTTON, 2000]*. This approach embraces the use of two and three-dimensions, such as those that might be derived from a series of GPS-derived coordinates.

With reference to the solutions proposed above, GIS vendors have developed proprietary data models covering linear referencing issues. The two most well known models are ESRI's Arc/Info Dynamic Segmentation module, and Intergraph's MGE Segment Manager method. Both software products use different methods to manage and query linearly referenced data, reflecting their underlying data models for managing spatial data and topology. *[SUTTON, 2000]*. Additionally, Oracle have developed its own methods of linear referencing.

Dynamic segmentation associates network attribute databases that are linearly referenced with topologically structured spatial databases (network models) whose reference frameworks are coordinate-based. To avoid the need for explicit representation of all point features and segment boundaries within the spatial database, dynamic segmentation computes coordinates from linear references "on-the-fly". Dynamic segmentation and network overlay enable spatial analysis and integration of highway inventory databases and any other databases that are linearly referenced.*[VONDEROHE ,1993]* Using dynamic segmentation, the basic edge/node structure is preserved, but a structure is added above it. This represents the two types of discrete entities, zero dimensional and one-dimensional, located at arbitrary locations on the network.*[GOODCHILD,1996]* Although these methods provide partial solutions, some of problems remain unresolved, and additional requirements arise. These issues include

- All approaches, with the exception of the location approach, do not provide solutions for the integration of one-dimensional data with two or three-dimensional data. In these approaches, one linear referencing method is selected. The different linear referencing methodologies used by highway administrations are not considered. By standardizing on a single referencing method within an agency, the data held in the obsolete formats will either be lost or else must be converted.
- The linearly referenced data is stored as attributive data or in an explicit object without any spatial character ,such as operation kilometer or anchor section, in the database. Although dynamic segmentation is realized “on-the-fly”, for representation purposes, efforts of maintaining non-integrated spatial information sources do not lessen. Different data sources must be maintained parallel.
- In addition, the solutions require very high levels of data acquisition activity. During this research it was discovered that, at least in Turkey and Germany, many of the referencing points have been lost through time. The implementation of GIS-T with the proposed solutions can only be realized after the determination of every referencing point on the highway network. Application of this method will lead to extra cost and a delay in implementation.
- The location approach requires a full re-digitizing of the road features using GPS. This is needed to enable linkage by coordinate snapping of spatially accurate tracking, or events, to a spatially accurate map base. “However, this approach has not been formally stated or tested. Issues, such as repeatability of GPS positions, how to abstract networks, how to relate to other location referencing systems, and representation at smaller scales have not been adequately addressed”. *[DUEKER,2000]*
- The dynamic segmentation solutions offered by GIS vendors are dependent on software platforms. In order to use one of these solutions, the core modules of the appropriate GIS vendor’s software needs to be deployed at the agency. Additionally, GIS vendors have prerequisites with respect to the conceptual data model design, which can limit the data modeling scope. Existing databases therefore need to be modified to take into account the prerequisites of the software data model.

According to research carried out in the USA, there is on average a 50m error (800 m in the worst case) in translating a two dimensional coordinate into a linearly referenced system.*[NORONHA,1999]* This is caused by alignment errors and generalization in the databases. Due to disagreement in linear measurements, there is a further 50m error caused by communicating the location to a linear reference in a different database. In other words, the offset of a fixed point on the ground could be 800m with respect to one database, and 850m with respect to another.

Due to these problems, the usage of a three-phase transformation method is proposed for the integration of linearly referenced road data and three-dimensional coordinate system. *[GIELSDORF,1998]* This transformation was introduced to the conceptual data model as the member method ***Dynamic Reference Transformation*** of the **Road Event** object.

When using the ***Dynamic Reference Transformation*** methodology, spatial data, including linear referenced data, is integrated into the system in three dimensions. All data can be modeled without redundant storage of geometrical information. As all linear reference systems are based on the specification of the direction and distance from a known point to the unknown point, every linear reference system can be identified using (l, q), where (l) is the distance between the origin and the beginning point of the road event, and (q) is the normal to the linear element. Since highway administrations consider the use of linearly referenced methods to be inevitable, transformation from (x, y, z) to (l, q) systems is required. Another very important point is that during this transformation, linearly referenced data is not stored in one-dimension, but rather, in three-dimensions. Transformation needs to be realized dynamically, in order to prevent data inconsistencies and redundancy. In the proposed conceptual data model, the geometrical properties of a road axis from the (x, y) plane are described using the linear elements; line, arc and clothoid in three dimensions. For the transformation between the (l, q) system and (x', y'), the value of curvature (k) and the bearing angle (τ), which are both functions of the length, need to be defined.

The curvature (k) of the linear element is;

$$\text{Line: } k = 0 \tag{4.6}$$

$$\text{Arc: } k = \frac{1}{R} = constant \tag{4.7}$$

$$\text{Clothoid: } k = \frac{1}{R_A} + \left(\frac{1}{R_E} - \frac{1}{R_A} \right) \cdot \frac{1}{L} \tag{4.8}$$

Bearing angle (τ) for linear elements;

$$\text{Line: } \tau = 0 \tag{4.9}$$

$$\text{Arc: } \tau = k \cdot l \tag{4.10}$$

$$\text{Clothoid: } \tau\,(l) = k_A \cdot l + (k_E - k_A) \cdot \frac{l^2}{2L} \tag{4.11}$$

In Figure 4-19 the relationship between the three-dimensional and the linearly referenced coordinate systems is shown.

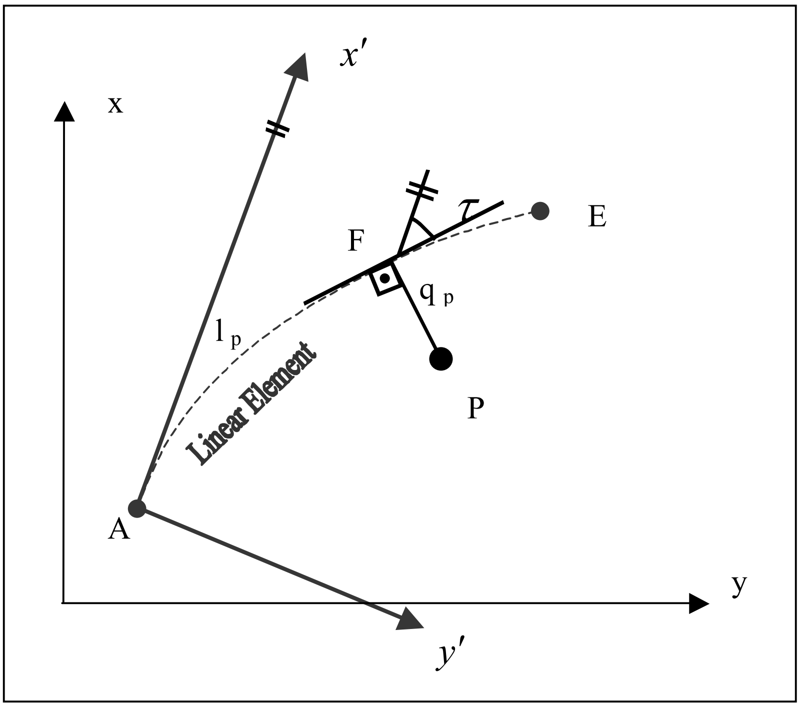


Figure 4-19: The Relation Between Coordinate Systems[GIELSDORF,1998]

The transformation can be realized in three steps:

STEP 1:

(*x, y*) point coordinates in a three-dimensional coordinate system are transformed into a local two-dimensional cartesian coordinate system (x', y') using a four parameter transformation *[BAUMANN,1995]*. The local origin is defined to be the linear element's beginning point. The x axis of the local system is defined to be parallel to the element's tangent at the origin. For transformation, the identical points in both systems are the element's beginning and ending points. With reference to Figure 4-19;

$$x_P = x_A + A.\,x'_P \tag{4.12}$$

$$x'_P = A^{-1}.\,(x_P - x_A) \tag{4.13}$$

Transformation matrix:
$$A = \begin{bmatrix} a & -o \\ o & a \end{bmatrix} = m.\begin{bmatrix} \cos\varepsilon & -\sin\varepsilon \\ \sin\varepsilon & \cos\varepsilon \end{bmatrix} \tag{4.14}$$

where,

The scale factor is
$$m = \sqrt{a^2 + o^2} \tag{4.15}$$

and the rotation is
$$\varepsilon = \arctan\left(\frac{o}{a}\right) \tag{4.16}$$

STEP 2:

Transformation of an (*l, q*) system into an (x', y') system or the retransformation of an (x', y') system into an (*l, q*) system.

a) In order to transform a linearly referenced system (*l, q*) into an (x', y') system, firstly the (x', y') coordinates of point F should be calculated. F is on the linear element and perpendicular to point P. The point P is defined by (*l, q*) coordinates with respect to the linear reference system.

$$d\bar{x} = \cos\tau(l)dl \quad \text{and} \quad d\bar{y} = \sin\tau(l)dl \tag{4.17}$$

The coordinates of $F\ (x', y')$ are given by:

$$\bar{X}_F = \int_0^{l_p} \cos\tau(l)\,dl \quad \text{and} \quad \bar{Y}_F = \int_0^{l_p} \sin\tau(l)\,dl \tag{4.18}$$

This is illustrated in Figure 4-20:

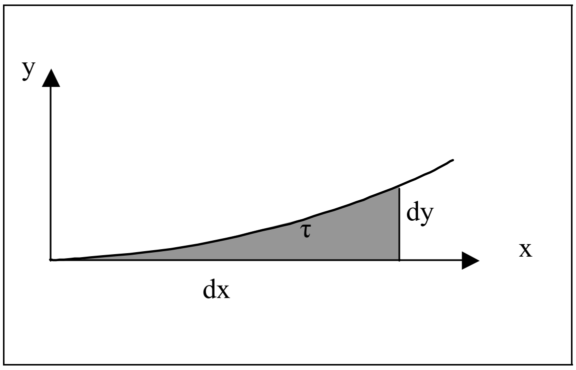


Figure 4-20: dx, dy Shown for a Linear Element Clothoid

Depending on the linear element type, integral solutions are:

For line
$$\bar{X}_F = l \quad \text{and} \quad \bar{Y}_F = 0 \tag{4.19}$$

For the arc $$\overline{X}_F = R.\sin\frac{l}{R} \quad \text{and} \quad \overline{Y}_F = R.\cos\frac{l}{R} \tag{4.20}$$

For the clothoid a direct solution is not available. Two alternative solutions exist.

i) Series expansion: *[MÜLLER,1984]*

$$\overline{X}_F = l - \frac{l^5}{40A^4} + \frac{l^9}{3456A^8} \mp ...... \quad \text{and} \quad \overline{Y}_F = \frac{l^3}{6A^2} - \frac{l^7}{336A^6} + \frac{l^{11}}{42240A^{10}} \mp .... \tag{4.21}$$

where $A^2 = R \cdot l$ and the origin is defined by $R_A = 0$.

ii) Numerical integration:

The approach taken here is to solve the problem using by discretising the continuous method $f(x)$ and directly integrating the discrete known methods thus:

$$\int_a^b f(x)\,dx \tag{4.22}$$

The most straightforward numerical integration technique uses the Newton-Cotes formulae (also called Quadrature formulae). Other similar techniques are the Trapezoidal Rule and Simpson's Rule. Numerical integration is more suitable for computational implementation.

The final part of *Step 2a* is to determine (x', y') coordinates for point P:

$$\overline{X}_p = \overline{X}_F - q_P \cdot \sin\tau(l) \quad \text{and} \quad \overline{Y}_p = \overline{Y}_F + q_P \cdot \cos\tau(l) \tag{4.23}$$

b) In order to transform (x', y') into a linear referencing system (l, q) of the linear element, The origin of (x', y') system is situated on the origin of the linear element. The l-coordinate corresponds to the distance of the point along the element and the q-coordinate represents normal to the linear element.

F is on the linear element and is perpendicular to P as illustrated in Figure 4-19. The distance between P and F must be minimized.

$$q(l) \overset{!}{=} \min \tag{4.24}$$

Therefore

$$q^2(l) = f(l) \overset{!}{=} \min \tag{4.25}$$

$$\frac{df}{dl} = f'(l) = 0 \tag{4.26}$$

Then,

$$q^2(l) = f = (\overline{X}_P - \overline{X}_F)^2 + (\overline{Y}_P - \overline{Y}_F)^2 \tag{4.27}$$

$$f' = \overline{X}_F \overline{X}'_F - \overline{X}_P \overline{X}'_F + \overline{Y}_F \overline{Y}'_F - \overline{Y}_P \overline{Y}'_F \tag{4.28}$$

In general

$$\overline{X}_F = \int_{l=0}^{l_P} \cos\tau(l)dl \quad \text{and} \quad \overline{Y}_F = \int_{l=0}^{l_P} \sin\tau(l)dl \tag{4.29}$$

$$\overline{X}'_F = \cos\tau(l_P) \quad \text{and} \quad \overline{Y}'_F = \sin\tau(l_P) \tag{4.30}$$

For linear elements;

Line:

$$l_P = \overline{X}_F \tag{4.31}$$

$$q_P = \overline{Y}_P \tag{4.32}$$

Arc:

$$l_P = R.\arctan\left(-\frac{\overline{X}_P}{\overline{Y}_P}\right) \tag{4.33}$$

$$q_P = R - \frac{\overline{X}_P}{\sin\frac{l_P}{R}} \tag{4.34}$$

For the clothoid, because a direct solution for the integral is not available, l_P can be solved using Newton's method. The general formula for Newton's method is

$$l_{m+1} = l_m - \frac{f'(l_m)}{f''(l_m)} \tag{4.35}$$

$$f'' = \overline{x}'^2_F + \overline{x}_F\overline{x}''_F - \overline{x}_P\overline{x}''_F + \overline{y}'^2_F + \overline{y}_F\overline{y}''_F - \overline{y}_P\overline{y}''_F \tag{4.36}$$

$$\overline{x}''_F = -k(l)\cdot\sin\tau(l) \tag{4.37}$$

$$\overline{y}''_F = k(l)\cdot\cos\tau(l) \tag{4.38}$$

Applying the differentials;

$$l_{m+1} = l_m - \frac{\cos\tau(l)\cdot(\overline{x}_F - \overline{x}_P) + \sin\tau(l)\cdot(\overline{y}_F - \overline{y}_P)}{1 + k(l)\cdot(\cos\tau(l)\cdot(\overline{y}_F - \overline{y}_P) - \sin\tau(l)\cdot(\overline{x}_F - \overline{x}_P))} \tag{4.39}$$

where m is the iteration number.

The distance q_p for a clothoid is given by;

$$q_P = \frac{\overline{x}_F - \overline{x}_P}{\sin\tau(l)} \tag{4.40}$$

STEP 3:

The l-coordinate in the linear element's system is then transformed to the linear referencing system. The origin is determined according to whatever linear referencing method is being used. For example, in the case of the link-node method; the link's beginning node.

In the proposed conceptual data model, the stochastic properties of the linear elements and the (l, q) parameters are determined from the standard deviations of the point coordinates and the significance test's parameter (q). The transformations need to be performed dynamically, which means without storage of linear referencing information. Therefore, stochastic properties are required in order to control consistency. By using the significance test the points in an (l, q) reference system can be checked to see if they belong to a specific linear element. Within the implementation of the conceptual data model, the method ***Dynamic***

Reference Transformation was implemented using a Visual Basic program in order to automate the process.

With the implemented ***Dynamic Reference Transformation*** method;

- Decomposition of spatial and non-spatial information is realized. No additional objects must be defined for linear referencing data. By using the existing objects proposed in the conceptual data model, information is modeled. This increases the stability of the GIS-T and simplify the data maintenance.
- Thematic data is independent of geometrical displacements such as; realignment and error corrections.
- Multi-dimensional road information can be mapped into the conceptual data model without redundancy.
- No pre-defined methodology is required, users are free to apply the most appropriate methodology from their point of view. This is because every one-dimensional reference system is transformed dynamically “on-the-fly”, and stored in a three-dimensional coordinate system.
- Re-transformation into linearly referenced system is modeled and supported by means of interfaces.
- Stochastic properties of linear elements and (l, q) parameters are available.
- Full integration is realized with other data acquisition techniques where information is referenced in three-dimensions, such as; photogrammetry and GPS
- Existing data with multiple referencing systems can be fully integrated.
- Due to the minimal data acquisition requirements, a more economical solution is provided compared with other techniques.
- The proposed technique is independent of any software vendor or platform.

Although this approach provides solutions to the problems identified earlier, there is currently a performance problem. It has been noted in the literature that extensive algorithmic processing for techniques to integrate linearly referenced data, is not desired. This is due to performance issues and is especially so in the case of ITS applications *[GOODWIN, 1996]*. However, ITS technology requirements differ from those of highway administrations with respect to the speed and relevancy of processing linearly referenced road information. Linearly referenced data is of interest to the ITS community, but it is essential for highway administrations. Additionally, with respect to the rate of developments in information technology the performance problem can be expected to diminish.

4.3.2 The Cross - Sectional Information

The final reference system to be considered is the (h, q) reference system, named in this study as the cross-sectional system. Other highway administration requirements can be fulfilled using the defined referencing system. These include; the maintenance of road facilities, and the determination of suitable advertising panel locations from a traffic safety point of view.

In one of Germany’s federal states, approximately 100,000 cross sectional data objects are maintained yearly. *[NWR,1995]* Although there is a high demand and a source of the spatial information, in the conceptual data models this information is modeled without consideration of its spatial structure. One of the reasons this spatial information is neglected in GIS-T is that other systems are used, such as road design software or CAD, to process this information. Unfortunately, such neglect results in the spatial character of the information being wasted.

In order to model this spatial information in GIS-T three approaches can be followed;

- Consider the information as thematic information
- Consider the information like any other information. Then, according to the proposed conceptual data model concept, decompose this into spatial and non-spatial components.
- Consider this information as thematic information with spatial character.

The first approach is the one mainly used in current GIS-T. The cross-sectional design information is linearly referenced in road databases. Therefore it is assigned as thematic information to proprietary GIS databases using road network links. This information can be easily queried and the results can be presented in tables, just like any other thematic information. Other systems, such as CAD, are required for the processing of geometrical information, due to the degree of detail of this information. As a result, current systems need to deal with a continuous data exchange between GIS-T and CAD, and avoid data redundancies in both systems. However, although interfaces between both systems exist, automation of the data exchange over time is not fully supported. Both systems are accordingly maintained in parallel. Additionally, CAD systems have limited capabilities for querying spatial information. Therefore using this approach will not be efficient.

The second approach would be a complete solution, but it cannot be realized. Cross sectional design data also has spatial and non-spatial character, and can be handled similarly to other road events. The information provided using this reference system could be decomposed into geometrical and thematic parts, where geometry is stored in the geometrical component of the proposed conceptual data model, and non-spatial data in the thematic part. However, GIS architectures currently only support at most two and a half dimensions with roads represented by lines or polygons. As a result, reference system objects could not be visualized in GIS, due to limited abstraction levels and required information detail degree in cross-sectional design.

In the conceptual data model, in order not to lose the spatial character the cross sectional reference system was modeled as a road event with a dimension, following the third approach. This was described in Section 4.2.3.3. With the development of object-oriented databases in the future it will be possible to model such situations in an optimized way. This is due to;

- The layer philosophy will be superceded by the adoption of object-oriented concepts.
- Coordinates and reference systems can be considered as attributes of objects.
- An object can be presented at various abstraction levels.
- Encapsulation of methods and data can be realized, such as reference systems and transformations.
- Objects can be identified without their coordinates.
- Objects may have various geometries.

4.3.3 Geometrical Data Integration

Data integration is highly desired due to the decreasing long-term costs of obtaining and maintaining data, as well as its beneficial effects on data consistency and accuracy within agencies. Paradoxically, complete and integrated databases are very rare. It is very common to encounter situations where multiple sources of the same information exist. In addition, as user requirements vary considerably, information is also of varying quality.

Since an important aim of GIS-T is the integration of data sources, in particular spatial data, adjustment techniques have proven to be the most effective tool. This is due to the technique's provision of consistent data and the ability to determine accuracy.

During this study the data integration approach was widely used, especially for the detection of alignment elements in the horizontal and vertical planes. Non-planar topology was also implemented using the following objects and techniques;

- **Linear Element Vertical**
- **Linear Geometry** member method ***Detecting Alignment Elements***
- Implemented relative height information

These were described in Section 4.2.3.2. In order to implement non-planar topology, height information is required. Since relative height information is sufficient to fulfill the requirements of highway administrations, during the implementation of the proposed model relative height information was used. This is because the relative height information can provide concrete height information for geometrical elements and is accurate enough for applications such as; freight management or ITS technology.

However, in some cases relative height information is incomplete or absent (at least in digital form). It is very common in the conceptual data models that this information was considered as thematic information. As a result highway linear vertical elements can not be introduced into the system. In this case, there are other information sources available were height information can be determined by highway administrations.

Examples of such sources are;

- The road gradient and road inclination values.
- Geometrical design regulations, for example minimum overpass height, driving dynamics and safety regulations.
- Digital Terrain Models (DTM).

In order to obtain non-planarity, for the entire highway network, concrete height information for geometrical elements is required. These are only fully available in Digital Terrain Models among the introduced sources. However, several problems are apparent. Firstly, the required accuracy in applications which need non-planar topology, such as ITS and freight management, can not be achieved economically. Because DTM accuracy is tightly correlated with economical aspects which depend on the data collection method and map scale. Costs increase as the provided accuracy increases. Additionally, having high accuracy DTM data will not alone fulfill this requirement. This is because DTM data does not match with road structures such as; bridges, tunnels and overpasses in a one-to-one manner.

In order to solve this problem, other data sources, such as; geometrical design regulations, driving dynamics and traffic safety regulations, needed to be introduced to the system. However, in this case the solution is not unique, leading to redundancy. By means of adjustment techniques, the required accuracy can be achieved and the redundancy can be controlled.

The adjustment theory is an established optimization technique used to determine unknown parameters based on given observations. It provides a straightforward solution to the above described problems. The aim of least squares adjustment is to optimize the solution of a functional model by minimizing the residuals of the observations.

$$\sum v_i^2 \, P_i = \min \tag{4.41}$$

where, v are the residuals and **P** is the weight matrix containing values corresponding to the observation accuracy.

The unknown parameters $\bar{x}$ can be solved according to the following equation:

$$\bar{x} = (A^T\ P\ A)^{-1} A^T\ P\,(\,l - f(x_0)\,) \tag{4.42}$$

where A is the Jacobean matrix of the function derivatives with respect to the unknowns, l are the observations and $f(x_0)$ is the value of the function calculated with approximate values.

This optimization problem can be solved in one of two ways; direct and indirect. The direct approach is introduced into the system using conditional equations. The indirect approach is generally preferred due to its better suitability for computation and error estimation. With the indirect approach, two options are available;

- Introduce conditional equations between the unknowns.
- Enforce specific observations as being more accurate in the stochastic model.

Since conditional equations produce large normal equation systems, the second approach is preferable.

In order to help clarify the proposed second solution approach, an example illustrating the interpolated DTM height information and one of the constraints, minimum overpass height, is presented. This information is shown in Figure 4-21.

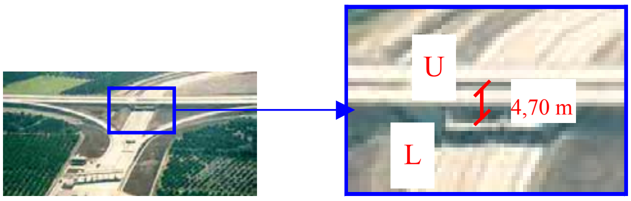


Figure 4-21: One of the Sample Constraints, Overpass-Height

The unknowns in this example are the relative height differences (Δh)

$$\Delta h = H_A - H_B \tag{4.43}$$

The interpolated height observations, representing the relative height differences between points along a road, can be defined in the functional model as;

$$\Delta h + v_{\Delta h} = H_A - H_B \tag{4.44}$$

Constraints, in this case overpass height, is introduced into the system, using the same functional model;

$$\Delta h + v_{\Delta h} = H_U - H_L = 4.70\,m \tag{4.45}$$

In the stochastic model, each observation's corresponding weight is stored in the diagonal elements of **P** matrix as:

$$P = \frac{\sigma_0^2}{\sigma_{\Delta h}^2} \tag{4.46}$$

For this example, the standard deviation of point heights obtained from the interpolated DTM is assumed to be $\sigma_{\Delta h} = \pm 5\,m$, and the observation variance is $\sigma_0^2 = \pm 1\,m$. The overpass-height observation is introduced with a standard deviation of $\sigma_{\Delta h} = \pm 1\,cm$. These introduced source, which is acquired more accurately, define conditions in the stochastic model. Since $\sum v_i^2 \; P_i$ must be a minimum and every observation must completely fulfil the conditions, the introduced overpass-height constraint enforces the model in order to fulfill its condition. Since the standard deviation information is "fixed", other parameters must change, including the observations. This process is performed iteratively, until $\sum v_i^2 \; P_i$ is minimum and all conditions are fulfilled. Consequently, when the proposed method is applied to a system where the height differences are interpolated from a DTM of lower accuracy, the non-planar topology can be optimized for the entire highway network. Therefore, high accuracy expectations are fulfilled at low cost. However, with this approach there is a risk of introducing very "strong" constraints, which results in undesired deformations of other observations in the system. This is due to, an adjustment approach allows for a change of all parameters while simultaneously enforcing constraints together with the tolerance of the constraining points. This problem can be solved by loosening the "strong" constraints until the appropriate solution is achieved.

4.3.4 Unique Identifier (Unique_id)

The necessity for data integration highlights another problem, being identification of the object in the system. Although object identifiers could be defined "temporary or persistent, locally or globally unique " *[BISHR,1999]*, GIS feature/object identifiers lose their meaning during data integration and data exchange between systems if they are locally unique and temporary. In GIS, a feature or an object can not exist without an identifier. Since every spatial and non-spatial information object is assigned to this identifier, the object identifier in GIS is the only information assigned which should be kept unchanged during the object's life-span in the system. It must therefore be persistent.

There are two types of object identification approaches;

- System automatic identifier generation.
- User defined identifiers.

The system automatic identifier generation is the main approach followed by GIS software vendors. Using this approach objects in GIS are uniquely defined by identifiers which are automatically generated by the system. This identified object is only unique in the defined system, database or table, not outwith the system. The automatic generated keys solution will have problems if there is a need for data integration and data exchange with other systems. If information is exported from a system, processed in the target system and re-imported to the original system with additional information, the use of automatically generated identifiers will result in two different and independent objects. This will definitely lead to redundancy and an effective loss of information from the database.

Secondly, the system automatic identifiers are assigned to the planar coordinates (x, y) of the object. Due to this, persistency of the system cannot be ensured. Object coordinates can easily change during the life-span of an object. For example, data obtained by means of a more accurate data acquisition technique or integrating additional information will lead to a change in these coordinates. Additionally, these assigned coordinates are two-dimensional while highway information is multidimensional.

The limitations in using system generated identifiers promote the use of user defined identifiers in GIS. When choosing an appropriate object identifier for highway agencies, there are many candidates including; road numbering systems and administrative area codes. However, persistency and uniqueness can not be ensured using these as identifiers due to involved spatial frame.

At first sight highway numbering systems, which are mainly used to provide guidance for drivers and by the technical applications of highway administrations, can be used as identifiers. This is because they are complete and uniquely identified for all road networks. However, during this study it was discovered that these identifiers are not permanent. This is due to geometrical frames being contained in the numbering systems, Therefore they cannot fulfil this requirement.

The highway numbering system of Germany is a good example in order to show the usage of spatial information as a frame. In this system, nodes are named uniquely and every net node consists of seven units. The first four digits are defined using the name of the 1:25000 scaled Gauss -Krüger projection map. The remaining three digits are determined from the node numbers for example [….499]. Every net point is numbered by the highway agency without consideration of which road they belong to (national, district..etc).

With this approach, since the first four digits are taken from the 1:25000 map, the numbering system is framed with spatial information, in this case the map projection. If another map projection needs to be used, these nodes change their spatial frame, concluding their name and cannot be found at their original locations. Indeed they may be in a completely different quadrangle. As a result their uniquely identified name will begin with a different number. This is shown in Figure 4-22.

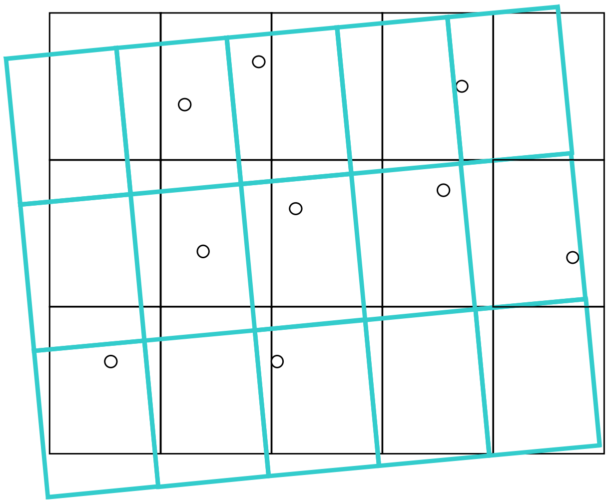


Figure 4-22: Geodetic Datum Transformation

Additionally, it is possible to introduce new naming systems, as in the case of the United States where all nodes names were newly defined according to an other naming regulation.

Another proposal is to use jurisdiction or administrative areas as the reference frame, and to identify features uniquely within this frame. This will also not provide the required solution. It is again a geometrical property which can change over time due to governmental or organizational requirements.

Consequently, the candidate user-defined object identifiers should ensure;

- Uniqueness
- Persistency
- Independence of any spatial information and frame
- Independence of any hardware or software vendor
- Simplicity of implementation

In addition to these requirements, there is currently an on-going discussion on the establishment of “intelligent (compound) identifiers” in order to identify data source or the owner of information in the case of data exchange between departments or other organizations.*[EAN,2001]*, *[OPENGROUP,2001] [ANSI,2001] [INTERLIS,2001]*

For highway agencies, the proposed solution is based on the naming of links, which have no geometric character. This is because the topological element link is a logical connection and invariant with respect to geometry. It is therefore independent of transformations, scale and other factors which can affect identifier permanency. All other identifiers such as, name, number and code need to be stored as thematic data in the database. The use of these should be avoided as they are in many cases redundant, incomplete, non-persistent and non-unique.

The required persistent and unique link identifiers can be named according to the following existing naming regulation steps. This is true even in the case of maps. These frames can be used in order to assist highway agencies, but including this information in the identifier’s name must be avoided. The appropriate naming regulation for link identifiers must be independent of any spatial frame including, geometrical structure, location, coordinate system, map projection or administrative area.

The unique identifier’s structure should be defined as follows. The number of digits should be determined according to the length of the road network and with consideration of the database and GIS capacities. If it is decided by highway administrations to establish intelligent identifiers, country codes, which are internationally unique, can be integrated into the identifier, in order to exchange data at international level. Additionally, a code identifying the national highway organization can be included in the system, in other to avoid data exchange problems with other national organizations such as; national police and health ministry. Other codes, including administrative area, highway administration division or any related to software or hardware peripherals should be omitted in order to preserve persistency.

Chapter 5 : Implementation of the Proposed Conceptual Model

5.1 Overview

During the conceptual data modeling stage, implementation details and system requirements were not considered. These topics will be considered in the context of the physical implementation of the data model. The physical storage proposals, as well as constraints and triggers, which needed be applied in order to realize methods were explained in this chapter. The developed data model and concepts were implemented on a sample project in order to recognize gaps and requirements which are not fulfilled during conceptual data modeling.

Before implementing the proposed concepts in a pilot project, it is necessary to decide which approach is the most appropriate to be followed. From the currently available approaches, the integrated architecture approach is selected, due to the expectation that this approach will promote the success of implementation. The benefits of the integrated approach were described in Section 3.3.1.

In order to support the usage of this approach, an appropriate software needed to be selected. To date a large number of geographic information system software packages exist. However, the GIS systems available on the market are not specifically tailored for road administration organizations, nor for the proposed conceptual data model.

Currently available GIS and database software was examined with respect to the developed conceptual data model in order to determine physical data modeling requirements, potential problems and limitations. The appropriate software selection criteria can be listed as follows;

- Open architecture.
- Integration possibilities between databases.
- Spatial data support.
- Storing spatial and non-spatial information in the same database.
- Ability to introduce user defined types.
- Ability to introduce user defined methods.

5.1.1 Software Used

5.1.1.1 Database Management System

During this study, Oracle8i, object-relational database management system (DBMS) with its Spatial Data Option extension was used for implementation. The Oracle8i provides a high level of functionality with user administration, multiple user access, distributed data retention, backup/recovery and replication.*[ORACLE8i, 2001]* The system provides many benefits for GIS architectures, especially through its use of Structured Query Language (SQL), spatial data storage and processing capabilities, as well as support of client-server environments.

With Oracle Spatial it is possible to store and process spatial data in databases through the use of indexing mechanisms. Storage of spatial and non-spatial data in a common database with integrated administration is possible. Currently only two-dimensional data is fully supported. There is however the possibility of calculating indices for higher-dimension points and for storing these in partitioned tables. Oracle Spatial defines three different two-dimensional elementary graphical objects being: point and point cluster, n-point polygons, as well as line strings and arc line strings.

The data model of the Oracle Spatial is oriented towards the simple features standardization proposal of the OpenGIS. Oracle8i extends the database model to include an SDO_Geometry object.

In Oracle Spatial, a spatial index provides a mechanism to limit searches based on spatial criteria. Three types of indexing mechanisms are used; R-tree indexing (the default), Quadtree indexing or Quadtree Hybrid indexing, which is a combination of both R-tree and Quadtree. An R-tree index approximates each geometry by a single rectangle that minimally encloses the geometry, which is called the minimum bounding rectangle (MBR). An R-tree index consists of an hierarchical index with respect to MBR´s tree structure. In the linear quadtree indexing scheme, the coordinate space is decomposed into tessellations, completely defining tiles for every stored geometry, until the termination criteria is reached. This is illustrated in Figure 5-1. Fixed or variable-sized tiles can be used to cover the geometry. Success of the query depends on the level and size of the defined tiles. If very small tiles are used the number of tiles needed to cover large areas would be very large. If large tiles are used, index selectivity will be unsuccessful due to the inability of approximating small geometries. Quadtree Hybrid indexing uses a combination of fixed-size and variable-sized tiles for spatially indexing a layer.

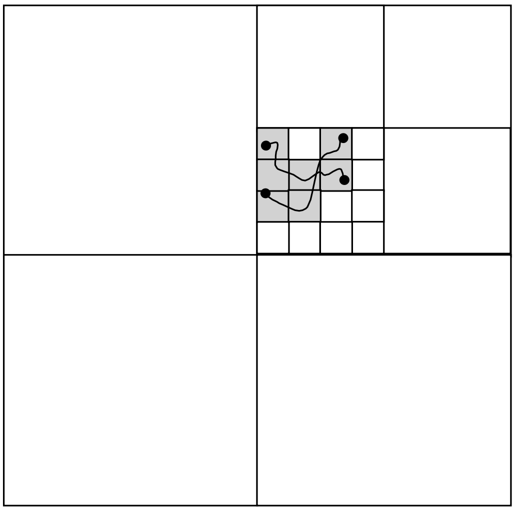


Figure 5-1: Linear Quad-tree Indexing Scheme

A two-tier query model with two distinct operations is used to resolve spatial queries and spatial joins. These are the primary and secondary filter operations. With the primary filter pre-selection is realized, these results are then transferred to the secondary filter. Oracle Spatial uses the secondary filters to determine the spatial relationship between entities in the database. The most common spatial relations are based on topology and distance. Filter operations provide spatial sorting and operations for objects, such as; locate, fence locate, spatial join, touch, equal or any interact.

The importance of using the Oracle Spatial Cartridge during implementation of the proposed data model include;

- One database management system including SQL
- Non-proprietary database.
- Integrated data management of spatial and non-spatial data.
- Storage of vector geometries.
- Query spatial relationships between geometries.
- Extensible component of network computing architecture, flexible application support.

- Combining standard components, lower cost of implementation, and standard protocols SQL, CORBA, HTTP.

5.1.1.2 GIS Software

The proposed conceptual data model was implemented using the Geomedia Professional 3.0 and ArcInfo8^{TM}. Both software are fulfilling above listed criteria for implementation purposes with their concepts of distributed databases, open architecture, possibility to introduce user-defined objects and methods.

Intergraph's is a software solution for the entry, administration, modeling, analysis and output of spatial data. With the software, it is possible to use the standard Microsoft database application Access, or to connect to another distributed database. Read/Write database access for Oracle is already provided by the software vendor. GeoMedia Professional 3.0 also provides extensive methods for data acquisition and processing, such as raster data registration and processing and geometry analyses. GeoMedia Professional 3.0 does not provide geometric network possibilities for the storage of topological information, but another concept is introduced, called "On the Fly" topology. This concept treats points, which are situated within a certain distance from each other as topological nodes. The consequence of this is that with geometrical manipulations these points are shifted together. GeoMedia requires that features stored in databases have a primary key, in order to realize the linkage between geometry components and the database. System recommends a maximum of eight pairs of coordinates for the Oracle Spatial module. Within GeoMedia Professional, it is possible to introduce user-defined objects since it is based on the Component Object Model (COM) technology developed by Microsoft, which allows components to be reused at binary level.

Second software used during implementation was ArcInfo8^{TM}. It is a powerful GIS software developed by Environmental Systems Research Institute (ESRI), provides solutions for the entry, administration, representation, analysis and output of geographical and thematic data. ArcInfo8^{TM} is based on three applications; ArcMap, ArcCatalog and ArcTool . The ArcMap is the environment providing the graphical interface for working with map data. The ArcCatalog is the data manager, which enables the administration of local directories and the relational data bases available in the system network. ArcTool is used for performing operations, such as data conversion and geographic datum transformations. Implementation of the data model is provided by means of the Computer Aided Software Engineering (CASE) tools. CASE tools also support the creation of COM classes. ArcInfo8^{TM} supplies various possibilities for the data management and integrity such as; attributes domains, relationship objects with relationship rules, geometric network with connectivity rules, user-defined objects and user-defined rules for objects.

ArcInfo8^{TM} is ArcSDE, which manages the physical storage of feature geometries in order to use simple standard data types within the host RDBMS. *[ESRI,1999]* Similarities of ArcSDE with the above described Oracle Spatial Cartridge can be found in the way that both of them manage geometrical storage of geometrical features in relational databases. When using Oracle Spatial Cartridge with ArcSDE, which was the case in the pilot project, geometry, spatial index and spatial searches are performed by Oracle Spatial.

The ArcSDE gateway enables the ArcInfo8 Geodatabase to leverage Oracle Spatial to store and manage feature geometry. With the ArcSDE, it is possible to realize;

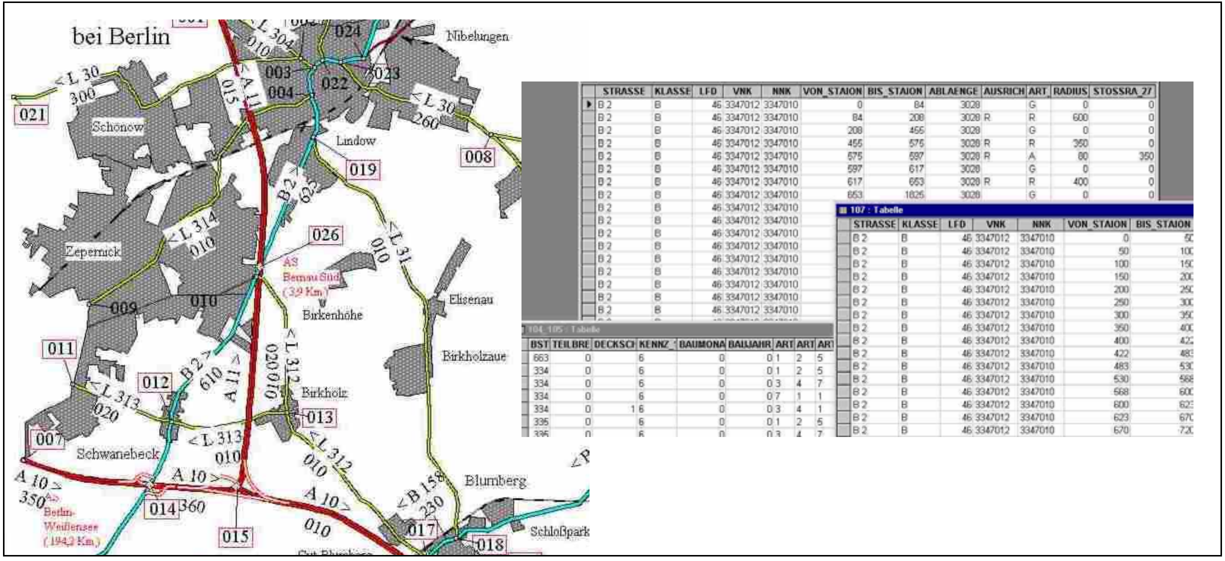
- Versioning ,which is long transaction editing
- Gateways to the ArcIMS, the internet web-server of ESRI,

- Gateways to the Dynamic Segmentation module, which covers linear referencing issues

With the ArcObjects object-component model, users can extend the data model using exactly the same COM technology which ESRI used to build ArcInfo8^{TM}. ObjectExtensions are an ActiveX DLL from ESRI which equips a feature object with additional behavior. With the implementation of the IobjectObjectValidation interface, it is possible to program user-defined consistency conditions.

5.1.2 Sample Data

The Brandenburg State Office for Traffic and Roads (BLVS) of Germany provided data for the pilot project. The provided data is located in an area to the northeast of Berlin around road B2. It was received in MapInfo (*.mif) and .dbf format shown in Figure 5-2. The data is created according to ASB regulations.



STRASSE	KLASSE	LFD	VNK	NNK	VON_STAION	BIS_STAION	ABLAENGE	AUSRICH	ART	RADIUS	STOSSRA_27
B 2	B	46	3347012	3347010	0	84	3028		G	0	0
B 2	B	46	3347012	3347010	84	208	3028	R	R	600	0
B 2	B	46	3347012	3347010	208	456	3028		G	0	0
B 2	B	46	3347012	3347010	455	575	3028	R	R	350	0
B 2	B	46	3347012	3347010	575	597	3028	R	A	80	350
B 2	B	46	3347012	3347010	597	617	3028		G	0	0
B 2	B	46	3347012	3347010	617	653	3028	R	R	400	0
B 2	B	46	3347012	3347010	653	1026	3028		G	0	0
B 2	B	46	3347012	3347010							
B 2	B	46	3347012	3347010							
B 2	B	46	3347012	3347010							
B 2	B	46	3347012	3347010							
B 2	B	46	3347012	3347010							
B 2	B	46	3347012	3347010							
B 2	B	46	3347012	3347010							
B 2	B	46	3347012	3347010							
B 2	B	46	3347012	3347010							

107 : Tabelle

STRASSE	KLASSE	LFD	VNK	NNK	VON_STAION	BIS_STAION
B 2	B	46	3347012	3347010	0	5
B 2	B	46	3347012	3347010	50	10
B 2	B	46	3347012	3347010	100	15
B 2	B	46	3347012	3347010	150	20
B 2	B	46	3347012	3347010	200	25
B 2	B	46	3347012	3347010	250	30
B 2	B	46	3347012	3347010	300	35
B 2	B	46	3347012	3347010	350	40
B 2	B	46	3347012	3347010	400	42
B 2	B	46	3347012	3347010	422	48
B 2	B	46	3347012	3347010	483	53
B 2	B	46	3347012	3347010	530	56
B 2	B	46	3347012	3347010	568	60
B 2	B	46	3347012	3347010	600	62
B 2	B	46	3347012	3347010	623	67
B 2	B	46	3347012	3347010	670	72

104_105 : Tabelle

BST	TEILBRE	DECKSCH	KENNZ_	BAUMONA	BAUJAHR	ART	ART	ART
663	0		6	0	0	1	2	5
334	0		6	0	0	1	2	5
334	0		6	0	0	3	4	7
334	0		6	0	0	7	1	1
334	0	1	6	0	0	3	4	1
335	0		6	0	0	1	2	5

Figure 5-2 Sample Data from the Brandenburg Highway Administration

Both data sources were successfully imported into the system according to below defined physical storage descriptions.

5.2 The Realized Physical Storage

After determination of implementation environment the physical storage was performed. The proposed data model has been implemented according to the provided tables. The tables topology, geometry and thematic components were structured according to the proposed conceptual data model concepts. Complete implementation proposal tables can be found at Annex D.

In the conceptual data model, the topological component consists of 12 objects with corresponding relationships between each other. Object were uniquely identified by their identifiers. In Table 5-1 the physical storage for several first level topological objects are shown. Second level topology can be implemented in the same way for the specified objects.

Table 5-1: Implementation of the First Abstraction Level of Topology

Object	Attribute	Data type	Description
Node I	NodeI_ID	Long	Identifier of Node
Link I	LinkI_ID	Long	Identifier of Link
	NodeI_IDBegin	Long	Identifier of Beginning Node of Link
	NodeI_IDEnd	Long	Identifier of Ending Node of Link
Road	Road_ID	Long	Identifier of Road
	RoadName	String	Name of Road
LinkI/Road	LinkI_ID	Long	Identifier of Link
	RoadName	String	Name of Road
	LinkIDOrder	Integer	Order of Links for Road

The geometrical component consists of 13 objects. In Table 5-2 and Table 5-3 the geometrical elements **Point Geometry**, **Reference System** and **PointGeo/RefSys** is provided according to the conceptual data model design.

Table 5-2: Implementation of the ***Point Geometry***

Object	Attribute	Data type	Description
Point Geometry	Point_ID	Long	Identifier of Point
	NodeI_ID	Long	Identifier of Node I
	NodeII_ID	Long	Identifier of Node II
	LinkI_ID	Long	Identifier of Link I
	LinkII_ID	Long	Identifier of Link II

Table 5-3: Implementation of the ***Reference System*** *and the Relation Between* ***PointGeo / RefSys***

Object	Attribute	Data type	Description
Reference System	RefSys_ID	Integer	Identifier of Reference System
	Sysname	String	Reference System Name
	Projection Algorithm	String	Projection Algorithm Name
	Projection Parameters	Long	Projection Parameters Storage Center, Resolution
	Geodetic Datum	String	Geodetic Datum
	Reference Ellipsoid	String	Reference Ellipsoid Name
	Ellipsoid Parameters	Long	Ellipsoid Parameters
PointGeo/ RefSys	Point_ID	Long	Identifier of Link
	RefSys_ID	Long	Identifier of Beginning Node of Link
	X	Long	X – Coordinate
	Y	Long	Y – Coordinate
	Z	Long	Z – Coordinate
	SigmaX	Double	Standard Deviation of X Value
	SigmaY	Double	Standard Deviation of Y Value

The implementation recommendation as described above overcomes any redundancy, which can occur if different referencing systems are used. Unfortunately it is not possible to realize this practically with conventional GIS software. With relational or object-relational GIS, each geometrical feature or object is assigned to its geometrical planar coordinates (x, y) with feature identifiers for the visualization of these objects. If the point geometry does not have (x, y) coordinates and an object identifier assigned to its coordinates, it is not possible to visualize point geometry with GIS. Therefore, the **Point Geometry** object was reconsidered in order to implement the proposed conceptual data model, as follows;

Table 5-4: Implementation of the ***Point Geometry***

Object	Attribute	Data type	Description
Point Geometry	Point_ID	Long	Identifier of Point
	NodeI_ID	Long	Identifier of Node I
	NodeII_ID	Long	Identifier of Node II
	LinkI_ID	Long	Identifier of Link I
	LinkII_ID	Long	Identifier of Link II
	X	Long	X – Coordinate
	Y	Long	Y – Coordinate
	Height	Long	Z – Coordinate
	SigmaX	Double	Standard Deviation of X Value
	SigmaY	Double	Standard Deviation of Y Value

Clearly, after this revision X, Y, Z, SigmaX and SigmaY should be deleted from the **PointGeo / RefSys** object.

Road Event, the third component of the conceptual data model was implemented in the following way, illustrated in Table 5-5;

Table 5-5: ***Road Event*** *Implementation*

Object	Attribute	Data type	Description
Road Event	Roadevent_ID	Long	Identifier of Road Event
	Geometry_ID	Long	Identifier of Geometrical Type of Road Event

Road event objects are assigned to one of the highway administration pre-defined geometrical elements through the geometry identifier. Road event identifiers are listed and assigned to road event objects.

5.3 Implementation

Within the pilot project, the first abstraction level topology, geometry and thematic components were implemented as presented in Chapter 4. Implementation was realized using SQL scripts directly in Oracle8i. The implementation was realized according to the above described physical storage. Using the SQL scripts, implementation of the proposed conceptual data model was simplified and automated. A sample SQL script generated for the linear geometry object is illustrated in Figure 5-3.

```
………
CREATE TABLE "OSC"."LINEARGEO"(
"LINEARGEO_ID" NUMBER(10) NOT NULL CONSTRAINT plineargeo PRIMARY KEY,
"POINTGEO_BEGIN_ID" NUMBER(10) NOT NULL CONSTRAINT flineargeo_1
REFERENCES "OSC"."POINTGEO"("POINTGEO_ID"),
"POINTGEO_END_ID" NUMBER(10) NOT NULL CONSTRAINT flineargeo_2
REFERENCES "OSC"."POINTGEO"("POINTGEO_ID"),
"LINK1_ID" NUMBER(10) NOT NULL CONSTRAINT flineargeo_3
REFERENCES "OSC"."LINK1"("LINK1_ID"),
"GDO_GID" NUMBER(38)
)
TABLESPACE "USERS"
…….
```

Figure 5-3:A Sample SQL Script for Creating Objects [PFANNMÖLLER,2001].

With Oracle's enchanted spatial indexing concept it was possible to store spatial and non-spatial information in one database. The SQL scripts were directly implemented in Oracle and access to GIS software was realized with software vendor provided interfaces.

In GeoMedia Professional 3.0, the conceptual data model's thematic and geometry components was fully realized. Two methods, which were introduced in Chapter 4, the ***Detecting Alignment Elements*** and ***Dynamic Reference Transformation*** were successfully implemented into system. In order to implement ***Detecting Alignment Elements*** AXTRAN software was used. The ***Dynamic Reference Transformation*** implementation was realized using a Visual Basic program.*[PFANNMÖLLER,2001]* Using these methods, the parameterized approach of the proposed conceptual data model was realized and automated. These methods were implemented into GeoMedia Professional 3.0.

Within the proposed conceptual data model, it was assumed that linear elements are defined by means of their parameters and the beginning and ending point and then generated without storage of any other geometrical information. However, although it was possible to parameterize defined geometric elements and introduce them to the system;

- It was not possible to generate these geometrical elements automatically using their parameters, without the storage of any other geometrical information. Because, geometrical features must be identified by their planar coordinates in the software.
- It was not possible to visualize the clothoid.

The first mentioned problem was solved by implementing the geometrical elements redundantly and controlling the redundancy with the developed methods. The visualization of clothoid was realized with the help of other geometrical elements, since it was not possible to define geometrical objects in the system.

Unfortunately the topology component can not be implemented in the manner proposed in the conceptual data model. The main problem encountered was the lack of user-defined types, meaning implementation opportunities were limited due to software vendor defined features, specifically spatial features. Topology information is considered differently in GeoMedia. Only in the "Maintain Coincidence Mode" ("On-the-Fly" topology instruction), displacement of one point has the consequence that, points which are situated within a determined distance will follow this movement. Additionally, it is not considered whether these points are assigned to topological nodes or not. In Figure 5-4 displacement of a node by means of "maintain coincidence mode" is shown.

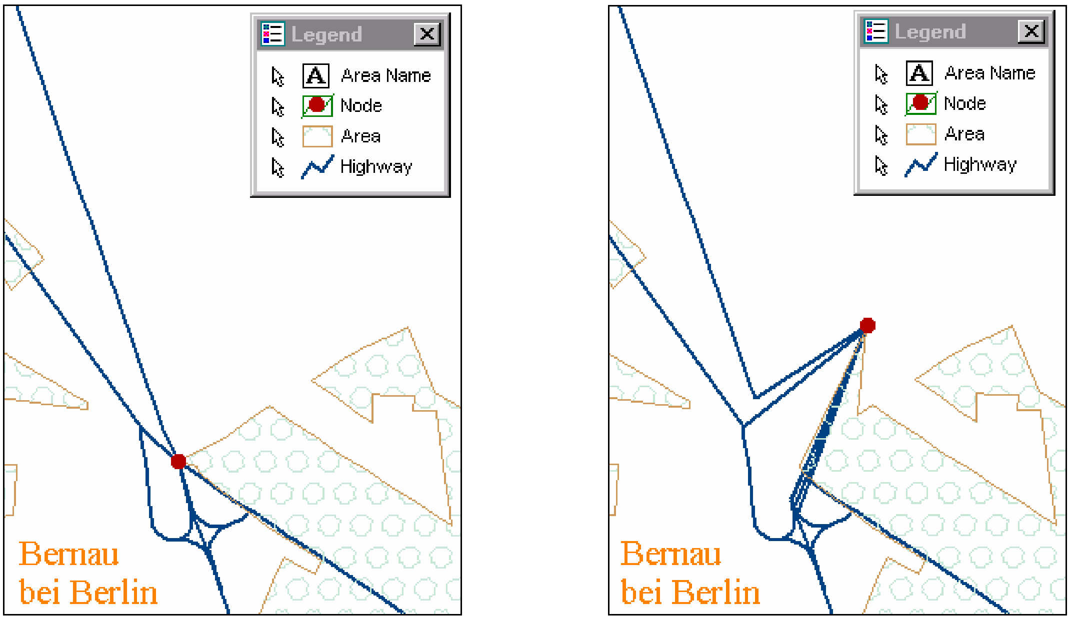


Figure 5-4: Editing the Node Using Maintain Coincidence Mode

There are some additional reasons, which affects the implementation of topology. The available mechanisms were insufficient in controlling the redundancies. The versioning concept is unavailable. A transaction is always automatically terminated in the system after a new object is created. This requires consistency checks to be realized immediately. However, in the conceptual data model consistency checks are pre-required for some entries, such as **Link** and **Node**, which were discussed in Section 4.2.4.1. There are also situations where this is reversed, such as determining **Point Geometry** and **Node** relations. In this case control mechanisms did not provide expected results. Additionally, it is not certain whether the quantity of the consistency conditions required are realizable using triggers. A further problem results from the fact that the user does not have influence on the end of a transaction *[PFANNMÖLLER,2001]*.

The second selected software ArcInfo8^{TM} has a very important benefit. Users can define feature types and are not limited to the software vendor's concepts and definitions due to provided ArcObject concept. With the "geometry network" possibility, topological elements can be realized according to the descriptions in the conceptual data model. In ArcInfo8^{TM} integrity constraints called "validation rules" are also available and can be defined by the user. Additionally, the versioning concept is available. The conceptual data model was successfully implemented through ArcInfo8^{TM} , although during the implementation some problems needed to be solved. These are as follows;

1. *Geometrical elements can not be generated using their parameters.*

   The main concept is similar to that used by GeoMedia Professional, where geometrical features are assigned to their planar coordinates for visualization purposes. In order to solve this issue, geometrical features and parameters were stored by using the user-defined objects and methods. The occurred redundancies was controlled using user-defined methods and validation rules.

2. *Features having geometrical characteristics, must be assigned to a pre-defined geometrical feature of ArcInfo8^{TM}.*

   Regardless whether an object is defined as a geometrical component or not, every object which should be visualized in GIS must have planar coordinates assigned to it, including the topological elements: node and link. During the implementation, an additional column was added to the physical data model table in order to define the geometry of feature. Consequently, objects such as linear elements were stored redundantly in the shape column with their geometrical information, in this case the coordinates of the polygon’s points. These redundancies were controlled by consistency rules.

3. *The linear element clothoid is not supported in ArcInfo8^{TM}*

   Although clothoids were stored with their parameters in the database, automatic generation of parameterized elements was not possible. In order to be able to visualize them, clothoids were simplified as line and arc objects during visualization.

4. *Object can only be characterized by their object identifier (Object_ID), which are automatically assigned by the software. Object_ID´s, with respect to object- identifier principal requirements, are not unique throughout the system, only unique within a table.*

   With respect to the unique identifier problem mentioned in Section 4.3.4, a significant problem became apparent. An individual column or a combination of columns can not be uniquely defined in ArcInfo8^{TM}. However, it is possible to implement unique attributes which can be realized through user programmed ObjectExtensions. This uniqueness of attributes can only be realized using the ArcMap^{TM} Editor tool, Validate Selection. Additionally, this can only be executed manually by the user. In this case, the importance of establishing regulations for unique identifiers and the automation of this process gains more importance.

5. *By using the “geometric network” of ArcInfo8^{TM}, topological relations can only be modeled between geometrical objects.*

   With the “geometric network” of ArcInfo8^{TM}, topological objects and relationships between each other could be realized. However, due to non-separation between geometry and topology, this realization can only be made between geometrical objects. Due to this limitation, implementation of a **0..1:1** relationship between **Node** and **Point** could not be realized.

In order to implement the proposed conceptual model, several modifications were made with respect to fifth item. The geometric element point was subdivided into three parts being; node, element point and intermediate point. **Node** was preserved as node. From the other new defined objects **Element Point** defines the beginning or ending point of a linear element. The **Intermediate Point** was introduced in order to indicate any location along a link, which is neither the beginning nor the ending point of the linear element. The modifications are presented in Figure 5-5.

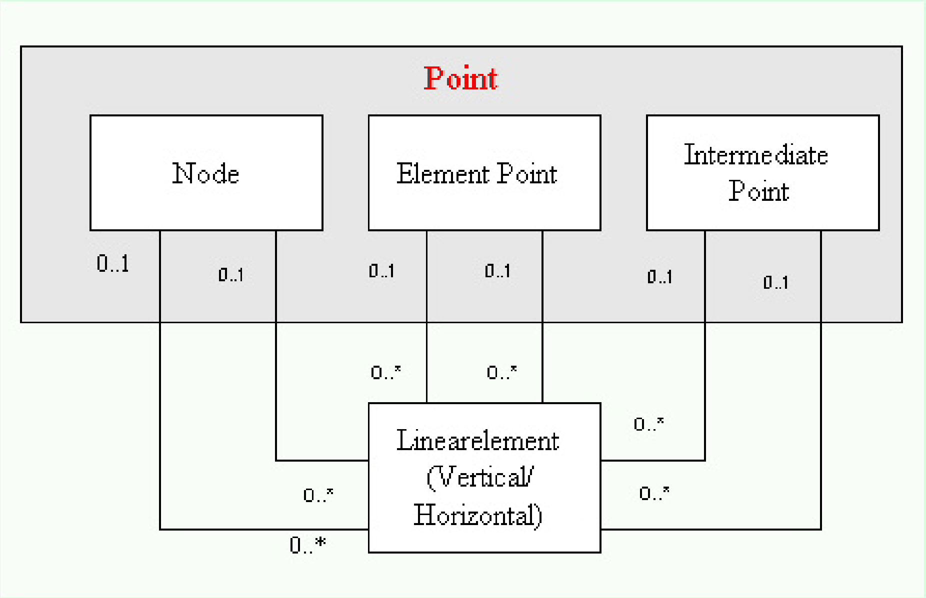


Figure 5-5: Modifications in the Conceptual Data Model

After these, implementation of the conceptual data model in ArcInfo8^{TM} was realized. A sample implementation for **Point** object is provided below in Table 5-6.

*Table 5-6: Sample Implementation for the **Point** Object*

Object Class	Object	Data type	Description
Point	ObjectID	Long	ArcInfo8^{TM} assigned object Identifier
	Shape	Geometry (Point)	Geometry type for all features assigned in Point
	S_Type	Integer	Division of Node, Element point, intermediate point
	LinkI_OID	Long	Identifier of Link I (ArcInfo8^{TM} assigned)
	LinkII_OID	Long	Identifier of Link II (ArcInfo8^{TM} assigned)
	Point_ID	Long	Identifier of Point
	NodeI_ID	Long	Identifier of Node I
	NodeII_ID	Long	Identifier of Node II
	Height	Long	Z - Coordinate
	SigmaX	Double	Standard deviation of X
	SigmaY	Double	Standard deviation of Y

In the table Table 5-6, S_Type column indicates the feature geometry type. In addition to the above mentioned **Point** object, consistency conditions were implemented in order to satisfy realization of the required rules, which are illustrated in Table 5-7.

Table 5-7: Consistency Conditions for Object Point

Attribute	Condition	Relation	
		Referenced primary key column	Cardinality
Object_ID	Positive integer, NOT NULL, Unique	__	__
S_Type	NOT NULL, 0 - 2	__	__
LinkI_OID	NOT NULL, Positive integer	Link I.Object_ID	0..1 to 0..*
LinkII_OID	NOT NULL, Positive integer	Link II.Object_ID	0..1 to 0..*

The conceptual data model thematic component **Road Event** was also implemented subject to similar modifications.

With the defined method ***Detecting Alignment Elements*** of **Linear Geometry** geometrical component of the proposed conceptual data model was generated. Topology component was implemented as proposed in the conceptual data model, which is shown in Figure 5-6.

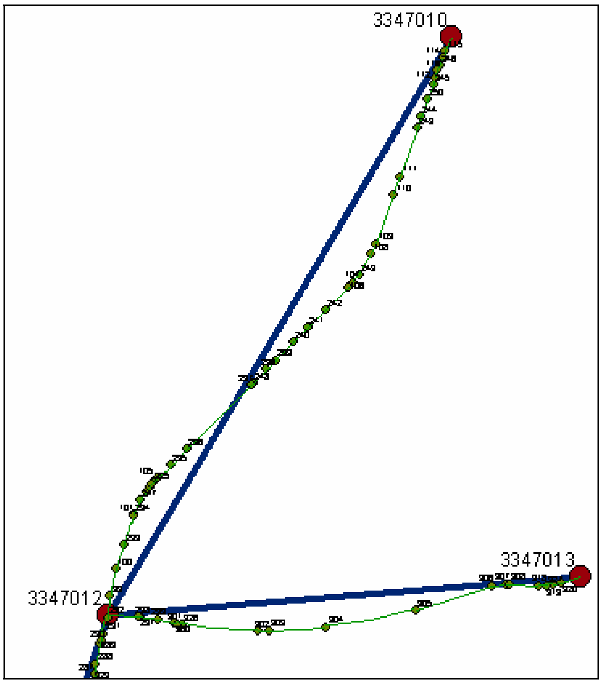
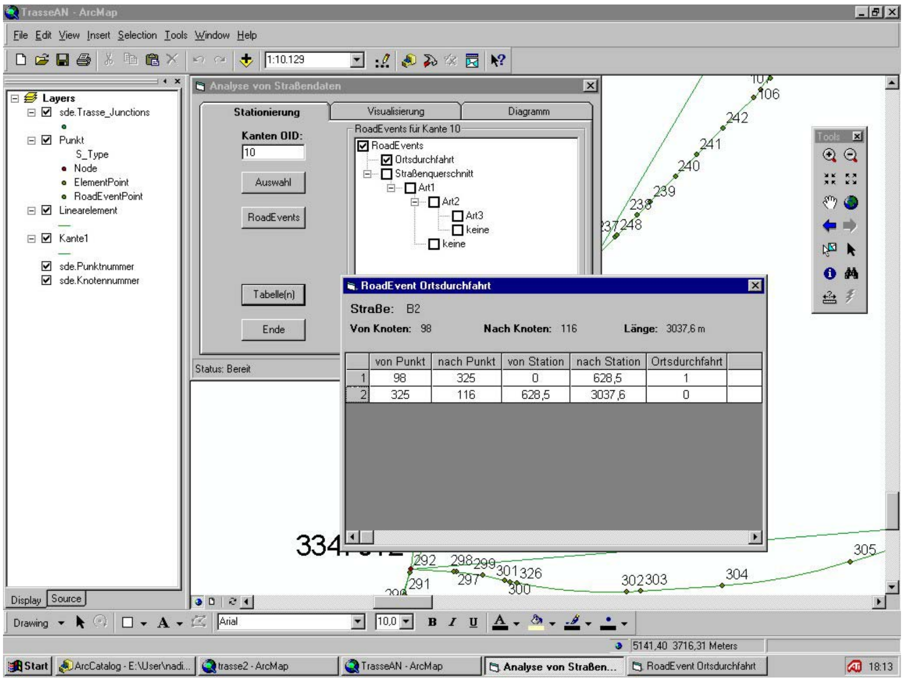


Figure 5-6: Implemented Geometry and Topology Components

The defined method ***Dynamic Reference Transformation*** of **Road Event** was generated using ArcObjects. Through compatible programming languages such as Visual Basic, Visual C++ and Delphi, it is possible to create or add user-defined objects to a database, which operate directly with ArcObjects. As the sample data for this pilot project was dependent on the link-node linear referencing system as proposed by the ASB, this methodology was used for re-transformation during designed interfaces. **Road Events**, described in the conceptual data model, were obtained using the program and existing interfaces shown in Figure 5-7. *[PFANNMÖLLER,2001]*


RoadEvents für Kante 10

- ☑ RoadEvents
  - ☑ Ortsdurchfahrt
  - ☐ Straßenquerschnitt
    - ☐ Art1
      - ☐ Art2
        - ☐ Art3
        - ☐ keine
      - ☐ keine

RoadEvent Ortsdurchfahrt

Straße: B2

Von Knoten: 98 **Nach Knoten:** 116 **Länge:** 3037,6 m

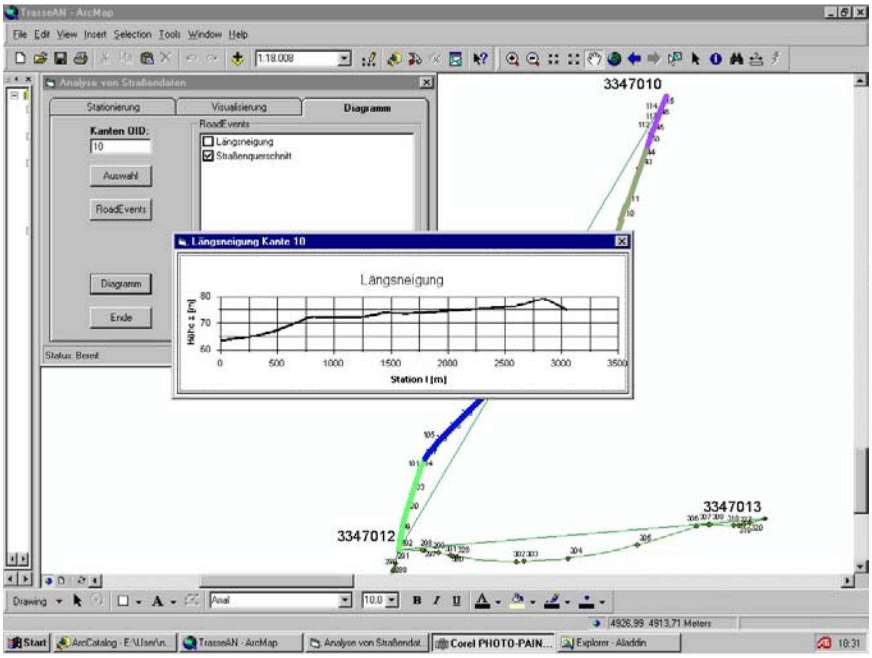
	von Punkt	nach Punkt	von Station	nach Station	Ortsdurchfahrt
1	98	325	0	628,5	1
2	325	116	628,5	3037,6	0

Kanten OID: 10

Status: Bereit

Figure 5-7: The Implemented **Road Event** Component of Sample Data

In addition, **Road Event** object information identified at vertical sections were also realized using the methods "***Detecting Alignment Element***" and "***Dynamic Reference Transformation***". The height information was already existed in the provided data set. This was linearly referenced using the link-node method, which is illustrated in Figure 5-8.


Kanten OID: 10

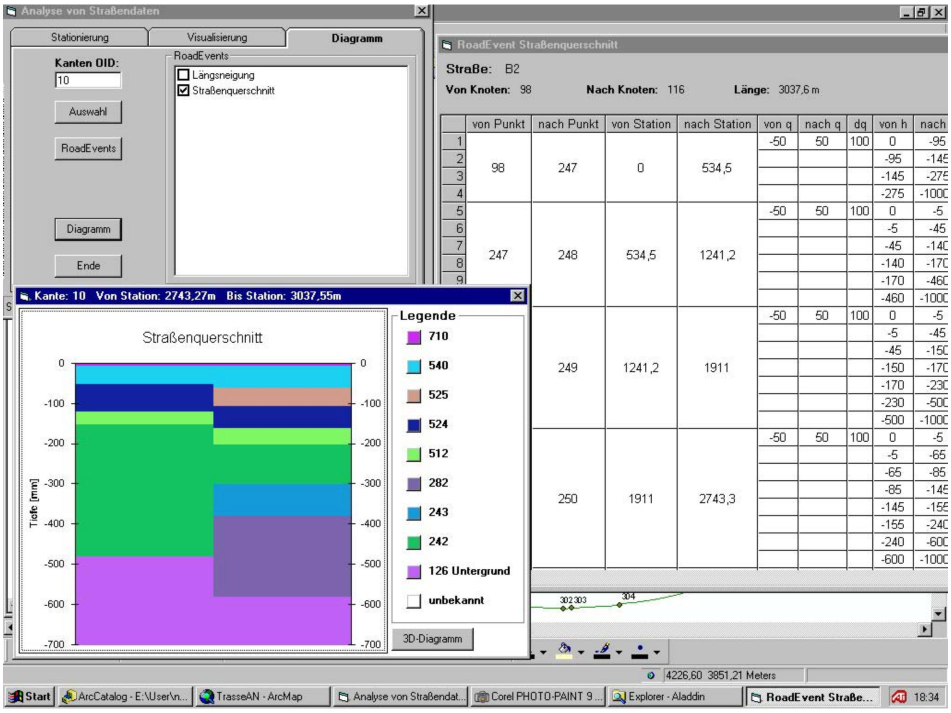
RoadEvents

- ☐ Längsneigung
- ☑ Straßenquerschnitt

Status: Bereit

Figure 5-8: The Vertical Linear Elements Sample

The cross-sectional design information, which was predefined as the (q , h) reference system, was also realized. Within this example, the stationing data of the road layer structure was connected to various types of layers, coded by the entries for Art1, with the coordinates of the (q, h) coordinate system describing the layers, such as Art2, Art3 *[ASB 1998]*. The cross-sectional design reference system origin was identified using the linear-referencing system data from the given node. The sample example is shown in Figure 5-9.



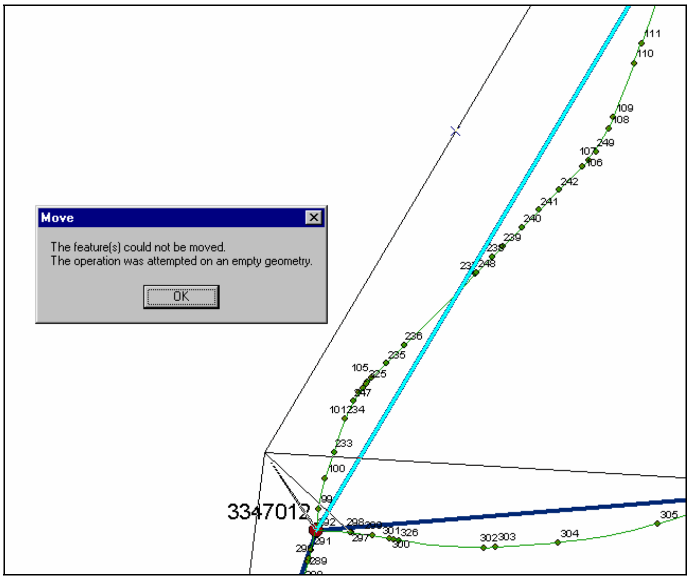
RoadEvent Straßenquerschnitt

Straße: B2
Von Knoten: 98 **Nach Knoten:** 116 **Länge:** 3037,6 m

	von Punkt	nach Punkt	von Station	nach Station	von q	nach q	dq	von h	nach
1	98	247	0	534,5	-50	50	100	0	-95
2								-95	-145
3								-145	-275
4								-275	-1000
5	247	248	534,5	1241,2	-50	50	100	0	-5
6								-5	-45
7								-45	-140
8								-140	-170
9								-170	-460
								-460	-1000
		249	1241,2	1911	-50	50	100	0	-5
								-5	-45
								-45	-150
								-150	-170
								-170	-230
								-230	-500
								-500	-1000
		250	1911	2743,3	-50	50	100	0	-5
								-5	-65
								-65	-85
								-85	-145
								-145	-155
								-155	-240
								-240	-600
								-600	-1000

Figure 5-9: The Cross-Sectional Reference System and Analyzed Road Data

Some of the defined integrity constraints in the proposed conceptual data model was performed using existing “validation rules” in “geometry network” of ArcInfo8^{TM}. Others were introduced into the system by means of triggers, which will be introduced below. An example of the validation rules used in the system is given in Figure 5-10.



The feature(s) could not be moved.
The operation was attempted on an empty geometry.

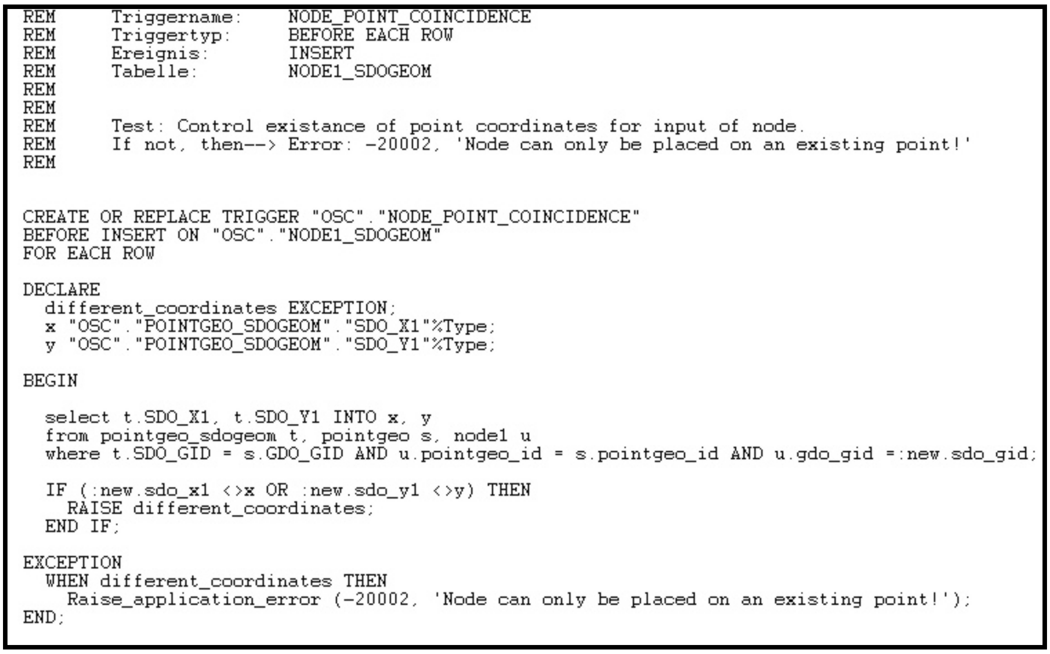
Figure 5-10: An Example of Validation Rule Used in Topology Component

The maintenance of complex relationships and validation of complex rules are often needed to be defined externally, due to lack of realization of the encapsulation concept. In order to define these, triggers are in the Oracle database management system. A trigger is a method which is invoked whenever a specified object or attribute is inserted, updated or deleted. Ideally, it should be possible to invoke the full range of GIS methods within a trigger and it should be possible to cause the current transaction to be rolled back if an invalid condition is found within a trigger. Procedures are started explicitly by the user, by an application or also a trigger. These are procedures written in PL/SQL, Java, or C that execute ("fire") implicitly. With the releasing event it concerns one or more Data Manipulation Language (DML) operations (insert, update and delete), illustrated in Table 5-8.

Table 5-8: Trigger Types

Category	Value	Description
Command	INSERT,DELETE, UPDATE	Defines which DML command fires the trigger
Occurrence	BEFORE or AFTER	Defines when trigger is going to be fired.
Scope	Row or Application	Row-scoped trigger is going to be fired for every concerned row. This will happen once, before or after a specified time interval.

The connectivity rules, redundancy controls, integrity constraints were realized using this method. In Figure 5-11 below, an example from the pilot project was illustrated, which was generated in order to control the consistency of the conceptual data model objects point and node.



```
REM     Triggername:     NODE_POINT_COINCIDENCE
REM     Triggertyp:      BEFORE EACH ROW
REM     Ereignis:        INSERT
REM     Tabelle:         NODE1_SDOGEOM
REM
REM
REM     Test: Control existance of point coordinates for input of node.
REM     If not, then--> Error: -20002, 'Node can only be placed on an existing point!'
REM


CREATE OR REPLACE TRIGGER "OSC"."NODE_POINT_COINCIDENCE"
BEFORE INSERT ON "OSC"."NODE1_SDOGEOM"
FOR EACH ROW

DECLARE
  different_coordinates EXCEPTION;
  x "OSC"."POINTGEO_SDOGEOM"."SDO_X1"%Type;
  y "OSC"."POINTGEO_SDOGEOM"."SDO_Y1"%Type;

BEGIN

  select t.SDO_X1, t.SDO_Y1 INTO x, y
  from pointgeo_sdogeom t, pointgeo s, node1 u
  where t.SDO_GID = s.GDO_GID AND u.pointgeo_id = s.pointgeo_id AND u.gdo_gid =:new.sdo_gid;

  IF (:new.sdo_x1 <>x OR :new.sdo_y1 <>y) THEN
    RAISE different_coordinates;
  END IF;

EXCEPTION
  WHEN different_coordinates THEN
    Raise_application_error (-20002, 'Node can only be placed on an existing point!');
END;
```

Figure 5-11: The Example of Available Triggers in System.

Chapter 6 : Conclusion

6.1 Overview

In highway agencies many of the benefits of GIS-T including, integration of data and methods, enforcing rules and standards, cost reduction and quality improvement was not fully realized and efficiency of GIS-T is generally under estimated. The relative success of implemented system is not clear without a detailed information analysis and a data model, which rely on formal data model design methodologies. A structured approach to database design permits to adequacy and efficiency in meeting the demands of organizations. In order to increase the efficiency and highlight the benefits of GIS-T, this study considered a progressive approach appropriate to the conceptual data modeling requirements of an entire highway agency.

Existing problematic areas in the modeling of spatial highway information is determined from user assessments and by the evaluation of conceptual data models. It is concluded that many problems existing within GIS-T systems are due to not recognizing the spatial characteristic of road information. It was also determined during this study that several demands of highway administrations were not responded by means of current systems. These topics can be summarized under several categorizes being;

- Relationships among geometry, topology and thematic information
- Analyzing multi-dimensional road information and realization of transformations between these dimensions
- Modeling highway administrations business rules and integration into the system.
- Realization of multiple topological representations, with a non-planar topology
- Integration of existing information sources and methods.
- Metadata, such as consistency rules, quality specifications and history information needed to be incorporated into the system.
- Permanent, non-spatial and a unique object identifier is required.

In order to fulfill these requirements, the main approach of the proposed data model was abstraction and decomposition of geometry, topology and non-spatial (thematic) data in order to provide a efficient data management within highway administrations. In order to achieve data integration, control of redundancy and optimization of data maintenance a parameterized approach was mainly adapted.

The basic component of the data model is geometry, which is introduced into the system using geometric elements. All road related information was mapped in an integrated manner for a complete enterprise. Problems identified during the analysis of current GIS-T, particularly with respect to the spatial nature of road information, were solved in the conceptual data model by means of designing appropriate objects and relationships between each other. Two methods were introduced into the system being; ***Detecting Alignment Element*** and ***Dynamic Reference Transformation.*** With the help of these methodologies geometry elements were parameterized and linearly referenced data was integrated into the system with its three dimensional coordinates. Consequently transformation between road multi-dimensions were performed. Topology, being non-planar and having two abstraction levels, is modeled as a logical connector abstracted from geometry. Non-spatial information was introduced into the system as descriptive information. Cross-sectional design information was also successfully integrated into the system with respect to its spatial information. During the proposed data model, the quality control, error trapping, data consistency checks and acceptance tests were designed and implemented into the system. This will definitely increase acceptance and the level of success of the system in highway administrations.

The main features of the proposed conceptual data model can be summarized as follows:

- Topology, geometry and thematic information is conceptually independent
- Multiple topologic representations, supporting different abstraction levels, are realized with two abstraction levels of topology, in order support diverse applications in highway agencies.
- With the incorporation of height information and the designed objects in the conceptual data model, non-planar topological model is achieved. In order to achieve non-planar topology, other techniques were also proposed including, introducing constraints by means of adjustment techniques.
- Road information, such as data collected through linear referencing systems or cross-sectional design information, is modeled with decomposition of spatial and non-spatial characteristics.
- Highway business-rules are modeled using integrity constraints, user defined methods and triggers.
- Existing road information was be integrated into the system without redundancy through defined methods and adjustment techniques.
- Metadata, including history information was modeled.
- Quality specifications, including accuracy certification, were defined in the conceptual data model for objects and introduced methods.
- For highway administrations, a proposal for a permanent non-spatial unique object identifier was made.
- The conceptual data model is designed to be independent of software implementation details.

The proposed conceptual data model was successfully implemented after a criteria list was constituted in order to select the appropriate software supporting integrated GIS implementation. During logical and physical data model design, an object-relational approach was applied. The proposed conceptual data model was automatically implemented by means of SQL scripts. Using the integrated approach spatial and non-spatial information was implemented in one object-relational database.

However, during the implementation some problems were encountered, especially due to the parameterized approach proposed in the conceptual data model. In order to solve this problem, both parameterized approach and pre-defined geometrical feature supported by software vendors were implemented redundantly. By means of user defined methods, control of this redundancy were realized. In order to facilitate system maintenance, many user defined methods and constrains were defined and implemented.

These have adversely impacted the system’s query response performance. This problem was predicted before the conceptual data model design was begun, due to the inherent nature of GIS software, such as the lack of separation between geometry and topology. However, during the conceptual data model design, this decision was taken due to the longer life span of conceptual data models and data compared to software. Additionally, keeping data consistent and non-redundant provides advantages when maintaining data and for further usage of the system. This issue will be less significant in the longer term, due to an increase at computational hardware technology.

6.2 Perspective

Three topics have a great potential in GIS-T with respect to increasing efficiency of the entire highway agency GIS-T.

It can be expected that in the near future software vendors will provide software based more on object oriented technology, especially providing solutions using parameterized approach. The concept of encapsulation has not yet been extensively applied to databases and GIS software, therefore; conventional layer approach is mainly applied. Additionally, user defined data types and methods was not in an extended content supported by software vendors. For highway administration GIS-T purposes, with the development of encapsulation concept and other concepts of object oriented technology, the proposed conceptual data model will be highly adequate and effective with respect to data maintenance and spatial queries.

A second topic, with a great potential in GIS-T is the so-called “Geoportal”. This is an internet site or service providing spatially referenced information. Use of this technology by highway agencies has many benefits including; organization-wide spatial information distribution, increased public usage of road information as well as efficient data integration and data exchange with other organizations. Due to platform independence of this service, the contribution of many relevant parties is achieved. This collaboration also improves the level of data integration and quality of GIS-T. This concept will support the system architectures in highway administrations and distribute the spatial information efficiently and rapidly. Currently, the geoportal interfaces provided by software vendors are highly proprietary. Consequently, usage of most sites is limited to owners of the vendor specific systems. Spatial queries are also extremely limited. If these problems are solved by means of new techniques and approaches, the efficiency of GIS-T systems will highly increase, promoting the integrated approach.

The third emerging topic is standardization of GIS-T, accelerated due to the expectations of highway agencies for data integration and quality. Currently problems have been arising due to the cartographic representation of road information, positional accuracy and lack of existing conceptual data model and/or documentation. The formal data modeling approaches will at least solve such current problems and will prepare basics for standardization in GIS-T.

Chapter 7 : Literature

[ADAMS,1999]	Adams T., Vonderohe A., Butler J., Koncz N.: Model Considerations for a multi-modal, multi-dimensional location referencing system http://www.bts.gov/programs/gis/BTSWEB/GIS-T_99/Session_21/211.html
[ALTAN, 1996]	Altan M. O., Toz G., Külür S., Seker D.Z., Iyidiker A, Iyidiker H., Duran Z.: Information System Applications in ITU, Altan - Gründig (Eds) First Turkish-German Joint Geodetic Days, ITU Civil Engineering Faculty Printing Office, 1996
[ANSI,2001]	www.ansi.org/
[BAHNDORF, 1994]	Bahndorf J., Gielsdorf F., Gründig L. : The Design and Management of the Three Dimensional Alignment of Rails and Switches, FIG XX. International Congress, Melbourne, Australia, 1994
[BALAKRISHNAN,2000]	Balakrishnan R., Ranganathan K.: A Textbook of Graph Theory, Springer - Verlag, ISBN 0-387-98859-9, New York, 2000, pg 2-6
[BATINI,1992]	Batini C., Ceri S., Navathe S.B.: Conceptual Database Design: an Entity-Relationship Approach, The Benjamin/Cummings Publishing Company Inc, ISBN: 0-8053-0244-1, California, USA,1992, pg 231-243
[BAUMANN,1995]	Baumann E.: Vermessungskunde, Band 2: Punktbestimmung nach Lage und Höhe, Dümmler Verlag, Bonn, 1995
[BENDER, 1999]	Bender P., Bosworth M., Dueker K.: ORBIT, The Oregon Road base Information Team Draft Summary Report II, Research Project Report 111, Center for Urban Studies, Portland State University, Oregon, USA, 1999 www.upa.pdx.edu/cus/
[BILL-I,1999]	Bill R.: Grundlagen der Geo-Informationssysteme, Band I Analysen, Anwendungen und neue Entwicklungen Wichmann, ISBN 3-87907-326-0, Heidelberg, 1999, pg 167-172
[BILL-II,1999]	Bill R.: Grundlagen der Geo-Informationssysteme, Band II Analysen, Anwendungen und neue Entwicklungen Wichmann, ISBN 3-87907-326-0, Heidelberg, 1999, pg 385-391
[BISHR,1999]	Bishr Y.A.: A Global Unique Persistent Object ID for Geospatial information Sharing http://gipsie.uni-muenster.de/publication/10_march99_yaser.htm,1999
[BMVBW,1998]	Bundesministerium für Verkehr Bau- Und Wohnungswesen Abteilung Strassenbau, Strassenverkehr Anweisung Strasseninformationsbank (ASB), ASB92, Verkehrsblatt-Sammlung, Nr S 15011998, 1998, pg 1-3

[BOOCH,1994] Booch G: Object-Oriented Analysis and Design with Applications, Addison Wesley Longman Inc., ISBN0-8053-5340-2, USA, 1994

[BOOCH,1999] Booch G., Rumbaugh J., Jacobson I., Addision Aus dem Amerikan. Von Bernd Kahlbrandt und Dorothea Reder: Das UML- Benutzerhandbuch, Addison – Wesley – Longman, ISBN 3-8237-1486-0, Bonn,1999, pg. 49-80

[CASPARY,1998] Caspary W. , Joos G.: Quality Criteria and Control for GIS Databases, Kahmen, H. ;Brückl, E. ; Wunderlich, T. (Hrsg.) Geodesy for Geotechnical and Structural Engineering. Eisenstadt, 1998, S. 436-441

[CEN,2001] http://www.nen.nl/cen278/

[CHURCH et al,1994] Church R., Coughlan D., Cova T., Goodchild M., Gottsegen J., Lemberg D.: Final Report for Caltrans Agreement 65T155 (MOU 1) National Center for Geographic Information and analysis University of California, Santa Barbara CA, http://www.ncgia.ucsb.edu/Publications/Tech_Reports/94/94-6.PDF , USA, 1994, pg 44-46

[CONNOLLY,1999] Connolly T.M., Begg C.E.: Database Systems, A practical approach to design, implementation and management, Wesley Longman Limited, ISBN: 0-201-34287-1, USA,1999,pg

[DIGEST,2001] http://www.diffuse.org/gis.html#DIGEST

[DORSEY,1999] Dorsey P., Hudicka J.: Oracle 8 Design Using UML Object Modeling, The Ultimate Guide to Designing with Oracle8's Databases, Osborne /The McGraw-Hill Companies, Inc, ISBN: 0-07-882474-5, California, USA,1999, pg. 376-380

[DUEKER,1997] Dueker J.K., Butler J.A.: GIS-T Enterprise Data Model with Suggested Implementation Choices, http://www.upa.pdx.edu/CUS/, 1997

[DUEKER,2000] Dueker J.K., Butler J.A: A geographic information system framework for transportation data sharing, Transportation Research Part C 8, www.elsevier.com/locate/trc,2000, pg 13-36

[EAN,2001] European Article Numbering http://www.ean-int.org/index800.html

[ERTICO,1996] GDF 3.0 Documentation, 1996 http://www.ertico.com/links/gdf/gdfdoc/gdfdoc.htm

[ESRI,1999] Understanding ArcSDE: Integrating GIS and RDBMS, Environmental Systems Research Institute. Inc., United States of America,1999.

[EU-APAS,1996] European Commission Transport Research-APAS: Databases and Scenarios for European transport, Luxembourg: Office for Official Publications of The European Communities, ISBN 92-827-7418-X, 1996

[GDF,1996] GDF 3.0. Documentation , Part 2 http://www.ertico.com/links/gdf/gdfdoc/gdfdoc.htm,pg78-92

[GIELSDORF,1998] Gielsdorf F.: Datenmodellierung für Gleisnetze und Schaffung eines einheitlichen Raumbezugssystems, Deutsche Geodätische Kommission ,ISBN:3-7696-9531-3, München, 1998, pg.21-25

[GIS-T/ISTEA,2001] GIS-T/ISTEA Pooled Fund Study http://www.sandia.gov/Trans_Tech/gistea.html#GIS

[GOODCHILD,1996] Goodchild M. F.: Geographic Information Systems and Disaggregate Transportation Modeling, NCGIA Symposium I-10, http://www.ncgia.ucsb.edu/vital/research /pubs/mfg9607.pdf, 1996

[GOODWIN,1996] Goodwin C.H.W., Siegel D., Gordon S.R.: Location Reference Message Specification, Revision A, Spatial Data Interoperability Protocol for ITS Project, http://itsdeployment.ed.ornl.gov/spatial/document/html/lrmspec3.htm, 1996, pg 8-10

[GRÖGER,2000] Gröger G.: Modellierung raumbezogener Objekte und Datenintegrität in GIS, Wichmann, ISBN:3-87907-354-6, Heidelberg, 2000, pg.71-87

[GRÜNDIG-I,1988] Gründig L.: Datenverwaltungskonzepte für typiche Aufgaben aus der Ingenieurgeodäsie, Deutsche Geodätische Kommission, Reihe C.Nr.336, 1988,pg.38-43

[GRÜNDIG-II,1988] Gründig L., Bahndorf J.: Ein rechnergestütztes Entwurfsystem zur drei dimensionalen Trassierung, X. Internationaler Kurs für Ingenieurvermessung, München, 1988

[GRÜNDIG,1989] Gründig L.: Benchmark Test Database Working Group 1 Documents,1989

[GRÜNDIG,1992] Gründig L., Bahndorf J., Singer P.: The Design and Management of the three Dimensional Alignment of Rails and Switches using RAILCAT, Comp' Rail 92 Proceedings of an International Symposium, Washington, 1992

[HELOKUNNAS,1995] Helokunnas, T.: Object-Oriented Approaches Applied to GIS Development, Acta Polytechnica Scandinavica, Mathematics and Computing in Engineering Series No:75, Helsinki, 1995,pg. 8-13

[INTERLIS,2001] www.interlis.ch/

[ISO,2001] http://www.iso.ch/iso/en/stdsdevelopment/

[ISTEA,1995] Geographic Information System-Transportation ISTEA Management Systems Server net Prototype Pooled Fund

Study: Proposal and Project Plan, 1995, http://www.bts.gov/ntl/DOCS/pfs2.html

[JAMES,1976] James G.: Mathematics Dictionary, Van Nostrand Reinhold Company,ISBN:0-442-24091-0, New York, 1976, pg. 169

[KUHLMANN,1996] Kuhlmann H.: Ein Beitrag zur Überwachung von Brückenbauwerken mit kontinuierlich registrierten Messungen, Nr.218 Wissenschaftliche Arbeiten der Fachrichtung Vermessungswesen der Universität Hannover, ISSN 0174-1454, Hannover, 1996

[LAURINI,1994] Laurini R., Thompson D.: Fundamentals of Spatial Information Systems, The APIC Series, ISBN:0-12-438380-7, USA,1994, pg. 362,394-395

[MOLENAAR,1991] Molenaar M.: Linking Thematic Data to positional Data(Geometric)Status and problems of geographical information systems. The necessity of a geoinformation theory, ISPRS Journal of Photogrammetry and Remote Sensing, 46, Elsevier Science Publishers B.V., Amsterdam, 1991, pg. 85-103

[MOLENAAR,1998] Molenaar M.: An introduction to the theory of Spatial Object Modelling for GIS, Taylor & Francis Ltd, ISBN 0-7484-0774-X(paper), London,1998, pg 121-136

[MÜLLER,1984] Müller G.: Ingenieurgeodäsie Verkehrsbau Grundlagen, Verlag für Bauwesen, Berlin,1984

[MÜLLER,2000] Müller G.: Eisenbahnbau, Möser M. ...(Hrsg).-Heidelberg: Wichmann, ISBN: 3-87907-297-3, 2000

[NCHRP,1997] National Cooperative Highway Research Program: Development of System and Application Architectures for Geographic Information Systems in Transportation Research Results Digest, Transportation Research Board, National Research Council, Number 218, USA, 1997, pg 1

[NCHRP,1998] National Cooperative Highway Research Program: Development of System and Application Architectures for Geographic Information Systems in Transportation Research Results Digest, Transportation Research Board, National Research Council, Number 221, USA, 1998

[NORONHA,1999] Noronha V.: Towards ITS Map Database Interoperability-Database Error and Rectification, International Workshop on GIS-T and ITS, Chinese University of Hong Kong http://www.ncgia.ucsb.edu/vital/research/reports.html,

1999

[NWR,1995] Verkehrstelematik in Nordhein-Westfalen, Bericht über die Verkehrsdatenerfassung und Verkehrsbeeinflussung auf den Autobahnen in Nordhein-Westfalen, Ministerium für Stadtentwicklung und Verkehr des Landes Nordhein-Westfalen, Düsseldorf, 1995

[NWSIB,1998] Straßeninformationsbank Nordrhein-Westfalen (NWSIB) Ministerium für Wirtschaft und Mittelstand, Technologie und Verkehr des Landes Nordrhein-Westfalen Düsseldorf, Referat Öffentlichkeitsarbeit http://www.mwmtv.nrw.de/cgi-bin/mwmev/custom/pub/visit.cgi, 1998

[OKSTRA,2000] OKSTRA, Standardisierung graphischer Daten im Straßen- und Verkehrswesen Teil 2 - Realisierung Teilbericht B: Ergebnisse der Teilprojekte Forschungsbericht FE-Nr. 09.092 G95D, http://www.okstra.de/docs.html, 2000

[OMG,1999] OMG Unified Modeling Language Specification, Version 1.3 Rational, Software Corporation, http://www.omg.org/library/issuerpt.htm, 1999

[OPENGROUP,2001] Open Group Publications http://www.opengroup.org/ public/pubs/catalog/web.htm

[ORACLE,2001] Oracle ® Spatial User's Guide and Reference Release Part No. A88805-01 http://technet.oracle.com/docs/products/spatial/content.html , 2001

[ORACLE8i,2001] Oracle8i Concepts Release 2 (8.1.6) A76965-01, http://technet.oracle.com/doc/oracle8i_816/server.816/a76965/c01intro.htm#12981

[PAPAZOGLOU, 1989] Papazoglou M., Valder W.: Relational database management : a system programming, Prentice-Hall international series in computer science, New York, ISBN: 0-13-771866-7, 1989

[PFANNMÖLLER,2001] Pfannmöller N., Oheim F.: Realisierung eines GIS-Datenmodelles für Strassendaten unter besonderer Berücksichtigung von Linearen Referenzsystemen, Diplomarbeit Technische Universität Berlin, 2001

[PORTELE,2000] Portele C.: OKSTRA- Standardisierung von Informationen im Strassen und Verkehrswesen und die Umsetzung in die Praxis, Zagel B.(Hrsg): GIS in Verkehr und Transport, Heidelberg Wichmann, ISBN:3-87907-342-2, Karlsruhe, 2000, pg. 138-146

[PORTELE,2001] Portele C.: Stand und Weiterentwicklung des OKSTRA, Arbeitsausschuss 9.7. „ grundstazfragen der Datenverarbeitung“ Arbeitsgruppe 9 „Sonderaufgaben“, Forschungsgesellschaft für Strassen- und Verkehrwesen (FGSV), 2001

[ROLLAND,1992] Rolland C., Cauvet C.: Trends and Perspectives in Conceptual Modeling, http://citeseer.nj.nec.com/rolland92trend.html, 1992

[RUMBAUGH,1991] Rumbaugh J., Blaha M., Premerlani W., Eddy F., Lorensen W. : Object-Oriented Modeling And Design, Prentice-Hall Inc., ISBN 0-13-630054-5, USA, 1991, pg.73-75

[SCHEU,1995] Scheu M.: Erfassung und Rekonstruktion raumbezogener Objekte in Vektorform, Deutsche Geodätische Kommission, ISBN:3-7696-9494-5, München, 1995, pg.19

[SCHILCHER,2000] Schilcher M., Hosse K.: Tutorial: Geoinformationssyteme in der Ingenieurpraxis Einführung und Grundlagen am Beispiel des GIS Projektes Eichenau, TU München Fortbildungsseminar Geoinformationssysteme, 2000

[SUTTON, 2000] Sutton J.C., Wymann M. M.: Dynamic location: an iconic model to synchronize temporal and spatial transportation data, Transportation Research Part C 8, www.elsevier.com/locate/trc, 2000, pg 37-52

[TRB,1992] Transportation Research Board: Data for decisions: Requirements for national transportation policy making. (Special report; 234), Transportation--United States--Information services, ISBN 0-309-05 156-8, http://ntl.bts.gov/DOCS/DFD.html, USA, 1992

[VIS,1998] The Road Sector Information System, VIS- organisationen v/ Vejdirektoratet, Vejdatacentret, ISBN: 87 7491 660 2, Kobenhavn, 1998

[VONDEROHE,1993] Vonderohe A. P., Travis L., Smith R. L., .Tsai V.: Adaptation of Geographic Information Systems for Transportation, NCHRP Report 359, National Academy Press, Washington D.C., 1993

[VONDEROHE,1994] Vonderohe A.: On The Results of a Workshop on Generic Data Model for Linear Referencing Systems, http://www.bts.gov/gis/reference/report.htm, Wisconsin, USA, 1994

[VONDEROHE,1997] Vonderohe A., Chou C. et al: A Generic Data Model for Linear Referencing Systems, NCHRP, Subject Area: IA Planning and Administration, Number 218, USA, 1997

[WALTER,1997] Walter V.: Zuordnung von raumbezogenen Daten -am Beispiel der Datenmodelle ATKIS und GDF, Deutsche Geodätische Kommission, ISBN: 3769695208, München 1997

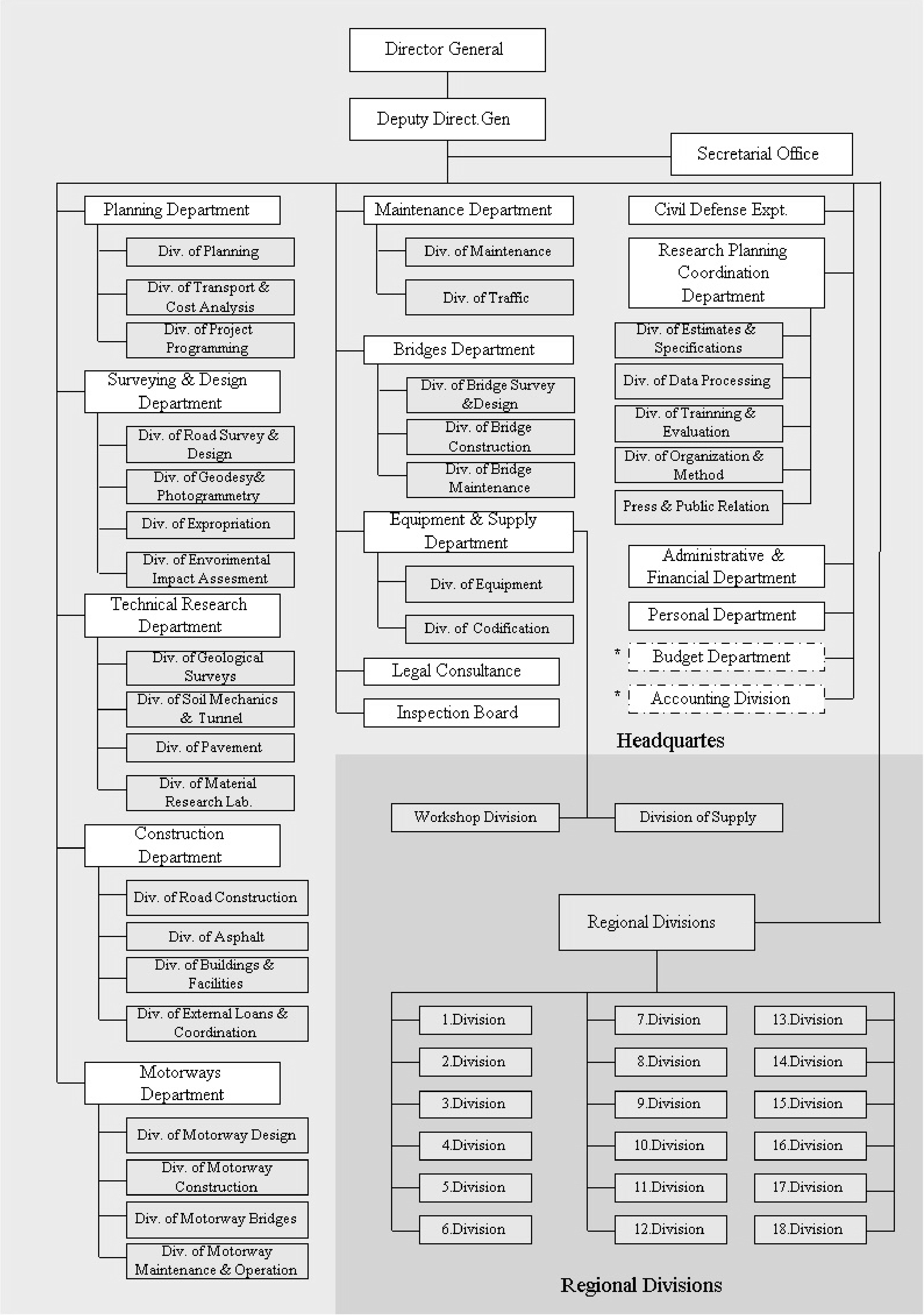
[WEBOPEDIA,2001] http://www.webopedia.com/

[WEISSTEIN,1999] Weisstein E.W.: The CRC Concise Encyclopaedia of Mathematics, CRC Press LLC, ISBN: 0-8493-9640-9, Florida, USA, 1999, pg.734

[WIDMANN,2000] Widmann, N.: Datenqualität in der Verkehrstelematik- Traum und Wirklichkeit, Zagel B.(Hrsg): GIS in Verkehr und Transport, Heidelberg Wichmann, ISBN:3-87907-342-2, Karlsruhe, 2000, pg. 225

Annex Index:

Annex A: The Organization Chart of General Directorate of Highway, Turkey



(*) Units of the Ministry of Finance

Annex B: Turkish Road Administration Departments Requirements

1. Department of Survey and Design:
    - a. Design information
    - b. Owner inventory
    - c. Environmental impact assessment information
    - d. Archaeological site, natural site
    - e. Budget, bidding and payment certificate information
2. Department of Technical Research:
    - 2.1. Soil Mechanics and Tunneling Division:
        - a. Soil survey information
        - b. Landslides areas determination and surveys
        - c. Tunnel inventory
        - d. Reinforced earth structures and geosyntetics
    - 2.2. Pavement Division:
        - a. Properties of hot mix asphalt layers
        - b. Pavement layers thickness
        - c. Serviceability data
            - i. Roughness (IRI, Ride number)
            - ii. Skid number
            - iii. Deflection
            - iv. Priority lists for pavement management system
            - v. Surface defects
        - d. Data for existing pavement layer
            - i. Thickness
            - ii. Properties of existing pavement layer
            - iii. Properties of sub_grade
    - 2.3. Geological Services Division
        - a. Geological-geotechnical research survey
        - b. Research information
        - c. Boreholes
        - d. Geophysical survey
        - e. Seismic
        - f. Resistivity
3. Department of Maintenance
    - 3.1. Division of Traffic
        - a. Accident analyses, data would be transferred via on-line connection from the National Police.
        - b. Existing data should be integrated
        - c. Location, property and the all information on the permission licenses
        - d. Advertisement signs inventory
        - e. Locations of the vehicle inspection stations

- f. Accident black spot inventories

3.2. Division of Maintenance

- a. Coordinates of bridges and intersections shall be determined.
- b. Characteristics of highways is going to be stored parallel with national catalogue codes.
- c. Existing inventory of 20,000 km. of highways should be integrated into the system.
- d. Inventory of damages related to avalanche and flood should be stored. (as point or area at a control section)
- e. Road Condition Report shall be followed up and immediately transferred to internet.
- f. The locations of salt quarries and related expenditures should be stored.
- g. Road facilities maintenance analysis should be possible
- h. Road geometric elements should be stored

4. Department of Bridges
   - a. Bridge inventory
   - b. Information about historical bridges
5. Department of Planning
   - a. Road network information, control Section numbers
   - b. International road definitions, numbers will also included.
   - c. Availability analyses
   - d. Simulation
   - e. Capacity and congestion analysis
   - f. The information related 1,2,3,4,5 numbered forms, which are for construction projects tracing.
   - g. Any information about allocation will be entered.
   - h. Integration of the computerized information
   - i. Traffic Counts of last five year, last year and forecast information of next 10 year
   - j. Existing road conditions
   - k. The information provided by modems from Automatic Counting Equipment Stations (WIM)
   - l. Axle Load Survey (axle equivalent) data will be entered
   - m. The Location of O-D Survey will be entered.
   - n. Intersection traffic counts will be entered.
     - i. The information of entrance-exit of customs
     - ii. The freight information concerning all the modes of transportation.(Vehicle-Ton-Passenger/Km.)
     - iii. Total construction costs of completed works
     - iv. Total expenditure of working places
6. Department of Construction
   - a. Information about the following up special agreements works. (Construction, Asphalt)
   - b. Statistical data of the works that are completed in the previous years
   - c. The place, type (i.e. asphalt chipping, base, sub-base) and amount of the aggregate stocks
   - d. Conditions of the bitumen stocks (i.e. place, type, amount)

   e. Asphalt force account works (work prosecution reports in terms of divisions and provinces)

7. Department of Machine and Supply
   a. Machinery Fleet movements shall be followed.
   b. Running, consumption and field control information shall be registered for each machine.
   c. Information related to each machine shall be followed.
8. Department of Motorways
   a. The information pertinent to motorways the design and construction
   b. Existing motorway routes will be recorded on 1/25000 scaled maps which comply with the projects at the final acceptance projects.
   c. Inventory information on motorways will be included in the system after determined by the related division.
   d. Operation information will be followed (income, input-output monthly)
   e. Motorways maintenance, operation and service facilities information
   f. Motorways and Bosphorous Bridges toll information

Annex C: The Questionnaire

Organization

- Organization profile?
- Structure of organization
- Activities, responsibilities of highway administrations
- Which analyzing techniques was used of determining the needs?
- Future plans

Data

- Types of data required by division and their relevancy
    - Accident reports
    - Slope, volume
    - Hourly directional traffic volumes
    - Classification of vehicles
    - Average traffic speed
    - Road geometry
    - Hilliness
    - Road related structures(buildings, toll areas)
    - Constructions
    - Traffic relevant information
    - Signpost information
    - Road types
    - Traffic flows
    - Vehicle weights
    - Variables influencing travelling behavior
    - Traffic accident information
    - Maintenance
- How frequent maps are used in daily tasks?
- Scales of used maps
- Which projection is used, cartographic representation?
- How often geometrical elements, cross-sectional designs are used?
- Which linear referencing system is used for obtaining road inventory?

- How this information is used?
- In which detail the topology information is required?
- Analyzing techniques for data
- For planning purposes which facilities should be added (DTM, visualization, cameras)
- Existing data sources in digital format.
- Which systems, software are existing? What are the purposes?
- The relevant lowest-volume data, level of accuracy
- Required quality of data
- How data accuracy and quality is being controlled?
- Resolution of data

Data Acquisition Techniques

- Which techniques are used for acquiring digital information?
- In data acquisition which techniques are mostly used ?Reason for this
- How far automation is used?
- Frequency of acquisition of each data type (annually, seasonally, special needs)
- Is it more efficient to acquire new data or existing data should be included?
- How is the collection of information coordinated?
- How data is updated , In which format data is used and stored?

Business Functions

- Daily tasks
- Correspondence with other specific divisions
- Hardware and software (existing , required)
- Is the any user-developed programs in usage?
- How is it possible to increase the efficiency of daily tasks?

Data Model

- The data model and structure
- Time requirement for modeling,
- Which object classes/entities are used?
- In data model which layers, features, attributes is being used?
- Which spatial information is modeled?
- How the existing data combined into the system?

- How topology is modeled?
- Is history information considered?
- How data model is documented?

Database System Management

- Which database system is used, reasons for choosing this system?
- System architecture preferred in the organization
- How updates are performed?
- How far new techniques is used, such as internet options for distributing data ?
- Which data formats are available?
- Any shared database?
- Data management (distributing, sharing, security)

GIS

- Goals and objectives of the GIS
- Who are the users of GIS?
- Current demands fulfilled
- Cope with changing demands(growth)
- Is there any data share between other organizations?
- Which procedures are pursued of establishing a GIS?
- Basic problems faced during the establishment
- Efficiency of the system
- Total investment for establishing the system
- Training of the personnel
- Which standards had been taken into consideration?
- How system is maintained?
- How base-maps are acquired? Is there a necessity to have divided highway information?
- Cartographic representation
- Generalization

Annex D: Completing the Psychical Data Model

The second topological abstraction level is the same as the level one. The physical data model of relationships between first level and second level abstractions are shown in Table A -1.

Table A -1 Transformation and Transformation Parameter

Object Class	Object	Data type	Description
Transformation	Trans_ID	Double	Identifier of Transformation
	Refsys_IDCur	Integer	Identifier of Current Reference System
	Refsys_IDTar	Integer	Identifier of Target Reference System
	Transtype	Integer	Code of Transformation Type
Transformation Parameter	Trans_ID	Double	Identifier of Transformation
	TansParaType	Integer	Transformation Parameter Type
	TransParavalue	Integer	Transformation Parameter Value

*Table A -2: **Link II***

Object Class	Object	Data type	Description
Link II	LinkII_ID	Long	Identifier of Link
	NodeII_IDBegin	Long	Identifier of Beginning Node of Link
	NodeII_IDEnd	Long	Identifier of Ending Node of Link
	Node I_ID	Long	Identifier of the Second Abstraction Level
	Link I_ID	Long	Identifier of the Second Abstraction Level

Table A -3 ***Node II***

Object Class	Object	Data type	Description
Node II	NodeII_ID	Long	Identifier of Node
	Node I_ID	Long	Identifier of the Second Abstraction Level

Other basic elements of the geometrical component of data model are **Area Geometry** and **Linear Element**.

Table A-4: ***Area Geometry, LinkI/Area, LinkII/Area****,* ***AreaGeo/LinPlan*** *and* ***AreaGeo/LinVer*** *association classes*

Object Class	Object	Data type	Description
Area Geometry	AreaGeo_ID	Long	Identifier of Area Geometry
LinkI/Area	LinkI_ID	Long	Identifier of Link I
	AreaGeo_ID	Long	Identifier of Area Geometry
	LinkI_IDOrd	Integer	Order of Link I
LinkII/Area	LinkII_ID	Long	Identifier of Link II
	AreaGeo_ID	Long	Identifier of Area Geometry
	LinkII_IDOrd	Integer	Order of Link II
AreaGeo/LinPlan	AreaGeo_ID	Long	Identifier of Area Geometry
	LinElP_ID	Long	Identifier of Linear Element Planar
	LinElH_IDOrd	Integer	Order of Linear Elements
AreaGeo/LinVer	AreaGeo_ID	Long	Identifier of Area Geometry
	LinElV_ID	Long	Identifier of Linear Element Vertical
	LinElV_IDOrd	Integer	Order of Linear Elements

Table A-5: ***Linear Geometry***

Object Class	Object	Data type	Description
Linear Geometry	LinGeo_ID	Long	Identifier of Linear Geometry
	LinkI_ID	Long	Identifier of Link I
	LinkII_ID	Integer	Identifier of Link II

Table A-6: ***Linear Element Vertical*** *and the relation between* ***Linear Geometry*** *and* ***Linear Element Vertical***

Object Class	Object	Data type	Description
Linear Element Vertical	LinElV_ID	Long	Identifier of Linear Element Vertical
	Point_IDBeg	Long	Identifier of Beginning Point of Linear Element
	Point_IDEnd	Long	Identifier of Ending Point of Linear Element
	Length	Double	Measured Length of Linear Element
	Curvature	Double	Curvature of Linear Element
LinearGeo/ LinVer	LinGeo_ID	Long	Identifier of Linear Geometry
	LinElV_ID	Long	Identifier of Linear Element Vertical
	LinElVOrder	Integer	Order of Linear Elements

Table A-7: ***Linear Element Planar** and **Parameter***

Object Class	Object	Data type	Description
Linear Element Planar	LinElH_ID	Long	Identifier of Linear Element Planar
	Point_IDBeg	Long	Identifier of Beginning Point of Linear Element
	Point_IDEnd	Long	Identifier of Ending Point of Linear Element
LinearGeo/ LinPlan	LinGeo_ID	Long	Identifier of Linear Geometry
	LinElH_ID	Long	Identifier of Linear Element Planar
	LinElHOrder	Integer	Order of Linear Elements
Parameter	LinElH_ID	Long	Identifier of Linear Element Planar
	ParameterTyp	String	Description of Parameter
	ParameterValue	Double	Value of Parameter
	SigmaParaVal	Double	Standard Deviation of Parameter

The additional **Road Event** classes are:

Table A-8: **Road Event Properties**

Object Class	Object	Data type	Description
Road Event Properties	RoadEvent_ID	Long	Identifier of Road Event
	EventType	String	Road Event Description
	EventValue	Double	Road Event Value

Table A-9 ***Dimensional, Zero Dimensional, One Dimensional, Two Dimensional*** *and* ***Onedim/Twodim***

Object Class	Object	Data type	Description
Dimensional	RoadEventdim	Integer	Identifies Dimension of Road Event
	Dim_ID	Integer	Identifier of Dimensional System
	Desc	String	System Description
	Origin_h	Long	System Origin h
	Origin_q	Long	System Origin q
Zero Dimensional	Zdim_ID	Long	Identifier of Zero Dimension
	h	Long	h-Coordinate
	q	Long	q-Coordinate
One Dimensional	Odim_ID	Long	Identifier of One Dimension
	ZdimBeg_ID	Long	Identifier of Beginning of Zero Dimension
	ZdimEnd_ID	Long	Identifier of Ending of Zero Dimension
Two Dimensional	Tdim_ID	Long	Identifier of Two Dimension
Onedim/Twodim	Odim_ID	Long	Identifier of One Dimension
	Tdim_ID	Long	Identifier of Two Dimension
	Order	Integer	Order of One Dimensional Elements in Two Dimension

Curriculum Vitae:

Surname, Name:	Demirel, Hande	
Date of Birth:	29.01.1974	
Date of Place:	Ankara, Turkey	
Education:	1981 – 1991	T.E.D. Ankara College
	1991 – 1996	Istanbul Technical University, Faculty of Civil Engineering Department of Geodesy and Photogrammetry Engineering
	1996 – 1998	Istanbul Technical University, Institute of Science and Technology, Geodesy and Photogrammetry Engineering, M. Sc. Eng.
	1998 – 2002	Technical University of Berlin, Faculty of Civil Engineering and Applied Geosciences, Division of Geodesy and Adjustment Techniques, Ph. D.
Work Experience:	Since 1997	Research Assistant at Istanbul Technical University, Faculty of Civil Engineering, Photogrammetry Division