BIOGRAPHICAL INFORMATION

Xavier R. Lopez, Ph.D.
Director, Spatial and Location Technologies
Oracle Corporation
1 Oracle Drive
Nashua, NH 03062

SPECIFIC RESPONSIBILITIES

Xavier Lopez joined Oracle in March 1998 and directs product management for the Spatial Technologies group. Xavier leads Oracle's efforts to incorporate location technologies across Oracle's database, application server, e-Business technologies. He has fifteen years of experience in the area of GIS and spatial databases.

PAST EXPERIENCE

Mr. Lopez has extensive background in the deployment of spatial technology with the Massachusetts Water Resources Authority and Massachusetts Department of Environmental Protection. Mr. Lopez has been active in numerous academic and National Academy of Sciences sponsored studies on spatial information. He is the author of a book on government spatial information dissemination practices and has published numerous research papers and book chapters on spatial technologies. Xavier served in U.S. Peace Corps Nepal during 1986-89.

EDUCATION

1998 University of California, Berkeley, (Chancellor's Postdoctoral Fellow)
1996 University of Maine, Ph.D, Spatial Information Engineering
1992 Birkbeck College, Univ. of London, Dept. of Geography (Fulbright Fellow)
1991 Massachusetts Institute of Technology, Master of City Planning
1984 University of California, BA Geography, Davis, BA

PROFESSIONAL MEMBERSHIPS

Geographic Information Technology Association (GITA), Board of Directors

International Geographic Information Foundation (IGIF), Board of Directors

Journal of Urban and Regional Information Systems (URISA), Editorial Board

THE CONNECTED DATABASE:
MANAGING NETWORK DATA MODELS IN AN RDBMS

Xavier R. Lopez
Oracle Corporation
1 Oracle Drive, Nashua, NH 03062

ABSTRACT:

Network modeling, management, and analysis are common tasks for Geographic Information Systems (GIS). This paper introduces an case study of how a DBMS centric network data model is used by British Telecom to manage its wireline assets. This is followed by a description of the germane network analysis capabilities, including routing (path computation) in transportation networks, tracing (accessibility) in utility networks, and resource allocation in decision-making and customer relationship management (CRM) applications. In this paper we present a network data model, that is a standard feature of database technology that lets users model and analyze networks. This feature simplifies network modeling, analysis, and management so that users can focus on application logic. The network data model provides an open, generic data model with many common GIS analysis capabilities. GIS applications based on the Oracle Spatial network data model are discussed in this paper.

INTRODUCTION

There are two common types of GIS applications: applications that deal with spatial proximity, and applications that deal with both spatial proximity and *connectivity*. For many applications queries on spatial proximity, which mainly consider metrics like Euclidean or geodetic distances, are sufficient. However, there are instances when connectivity needs to be taken into account, such as in finding the shortest distance between two locations in a road network. Many GIS applications are networks that require modeling and analysis on object relationships (connectivity). Figure 1 shows a road network of New York City this is comprised of 60384 nodes, 151962 links.

Oracle Corp. and IBM have supported spatial objects over the last eight years. A complete set of indexes, operators, and functions is available for managing information based on spatial proximity. Oracle recently introduced a network data model (NDM) that extends DBMS functionality to deal with connectivity. Currently there are many GIS vendors offering network solutions; however, their solutions may have the following issues that can be overcome through a network data model that is managed natively in a DBMS:

- Their data model is stored in proprietary file formats, and cannot be integrated with their database.
- The data model and analysis capabilities cannot be extended.
- Application information cannot be separated from connectivity information.
- Spatial information management and hierarchical relationships are not directly supported.

To address these issues, native DBMS network data model does the following:

- Provides an open and persistent network data model

  The network data model is stored as relational tables in the database and is managed through SQL.

- Simplifies network data management and analysis

  PL/SQL and Java APIs are provided for network data management and analysis.

- Separates connectivity and application information in the data model

  Connectivity information can be separated from application information. Both application information and connectivity information are managed in the database. However, only connectivity information is required for network analysis.

- Allows the extension of data model and analysis capabilities

  The Java API provides representations of network elements as Java Interfaces. Users can define their own network elements by extending these interfaces. As a result, users can implement their own user-defined representations and analysis functions.

- Integrates with native DBMS Spatial types for spatial information management

This paper is organized as follows: It introduces an case study of how a DBMS centric network data model is used by British Telecom to manage its wireline assets. This is followed by a detailed technical overview on the network data model database schema, APIs and architecture. Next is a discussion on how GIS application developers can use the data model and properly apply network constraints to enhance analysis capabilities.

CASE STUDY: APPLYING NDM AT BRITISH TELECOM

BT serves over 20 million business and residential customers with more than 29 million exchange lines, as well as providing network services to other licensed operators. The Cardiff, UK GIS group is leading the Piper Project whose main function is to provide GeoSpatial connectivity details of BT's networks (i.e. Copper, Fibre etc.) to support the national roll out of this service. The process involves accepting numerous data sources and then by employing a mixture of Spatial techniques to refine the underlying networks and the running of a rules based routing engine the required pathways through the various networks are determined. The data sources currently accepted into the process range from non-spatial legacy systems e.g. Customer Record details, to spatial sources such as Ordinance Survey Address Location info i.e. Post Code (Zip Code) and Ordinance Survey Centre Line of Road as well as Vectorized Raster Images produced from Drawing Office diagrams of Plant and Cable network

Network Analysis Routing Engine

A major component of the overall Piper process is the Network Analysis Routing Engine which facilitates the dynamic loading and analysis of requested networks. Essentially the Routing Engine is designed to provide results, in the form of paths, based upon the application of rule sets run against the requested networks. The engine is extremely flexible in its configurability, resulting in the ability to apply any set of rules to any supplied network as desired.

The Oracle Spatial Network Data Model software forms the core of the Routing Engine where extensive use is made of the NDM API functionality to provide network access, constraint application, path generation and storage of results.

Network Merging Utility

An interesting recent development in BT's use of the Oracle Spatial Network Data Model software has been the production of a utility to provide a comparison between two physically different but spatially similar networks e.g. Plant and Duct Networks. Due to the varied nature and quality of the data sources being fed into the Piper process a requirement arose where in order to provide improved connectivity information the ability to compare and cross-reference between different networks was deemed necessary. Once again the Oracle Spatial Network Data Model software has been used at the heart of this utility, providing the routing capability and network analysis functionality that underpins the application of the merging rules.

Technical Underpinnings of a DBMS-centric Network Data Model

The network data model consists of two parts: a network schema and network APIs. The network schema is the persistent data storage used to store network information. The network APIs contain a PL/SQL package for data management in the database and a Java API for data management and analysis on the client-side (via Java JDBC drivers).

Network Data Model Schema

A network contains network metadata, a node table, and a link table. In addition, path information (path table and path-link table) can be stored with the network if desired. Figure 2 shows a schematic view of a network in the database. Note that only connectivity information is stored. Additional application information can be stored in the network tables, or in other tables and referenced by foreign keys.

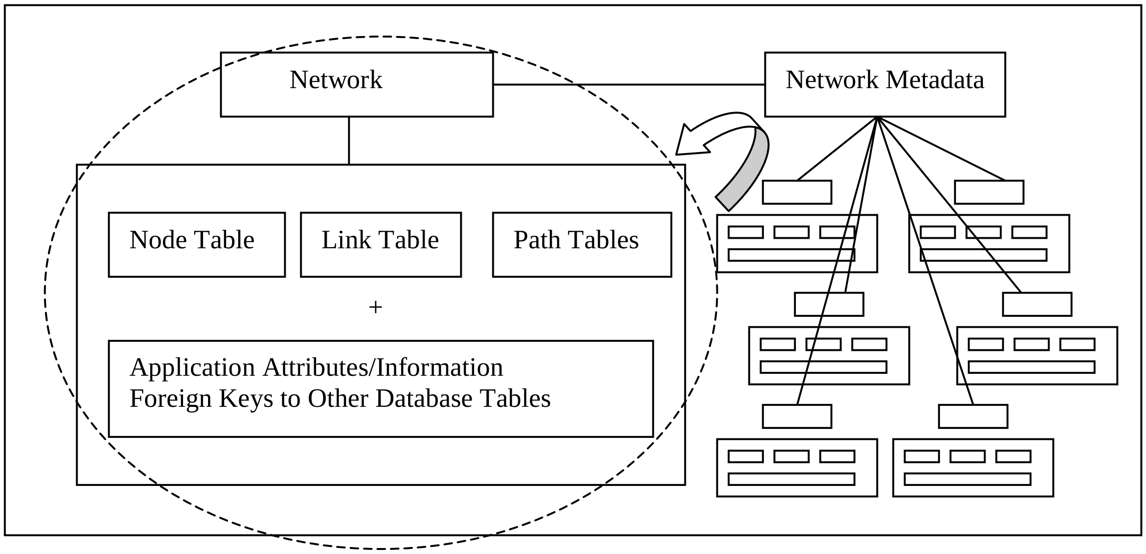


Figure 2. DBMS Network Data Model (Schematic View)

<u>Network Metadata</u>

Network metadata provides general information about networks. It includes the following network details:

- Directed or undirected
- Logical or spatial
- Hierarchical or flat
- Information about nodes, links, and paths
- Geometry information -- for spatial networks

<u>Network Tables</u>

A DBMS centric network contains two tables: a node table and a link table. A path table can be added if needed. Figure 2 shows the schema for the network data model, which includes these tables. The schema represents the information necessary for network management and analysis. Application attributes can be added to these tables or referenced from other application tables (through foreign keys). Note that the network data model is also capable of handling geometry information. That is, the network data model can represent both logical and spatial network applications. Adding geometric data to a logical network will allow the logical network to be displayed.

Network Data Model APIs

The network data model provides a PL/SQL API and a Java API for network management on the database and client sides. The Java API can also be used for network analysis.

Network Data Model Analysis Capabilities

The following analyses are supported in the network data model:

- Shortest Path: The shortest path from node A to node B
- Accessibility Analysis: Is node A accessible from node B?
- Minimum-Cost Spanning Tree: What is the minimum-cost tree that connects all nodes?
- Within-Cost Analysis: What nodes are within a given cost from a given node?
- Nearest Neighbors: What are the *N* nearest neighbors of a given node?
- K Shortest Paths: What are the *K* shortest paths from node A to node B?
- Connected Components Analysis: Label connected components with IDs.
- Graph Operations: Graph union, intersection, and difference.
- Traveling Salesman Problem: What is the minimum-cost tour that visits a set of cities?

MODELING NETWORK APPLICATIONS

The network data model takes a generic approach to solving network problems, by separating connectivity information from application-specific information. Figure 3 shows how a typical network application can be modeled and analyzed. First the network connectivity information (node connections and link cost) is extracted and separated from the application-specific information. Application-specific attributes are stored, if needed, with the connectivity information or separately. Once the connectivity information is extracted, network analysis is conducted on the generic model. Additional network constraints can also be considered. The final result is then mapped to application-related attributes, and displayed. This approach avoids customized network solutions and simplifies the data management of connectivity and application-specific information.

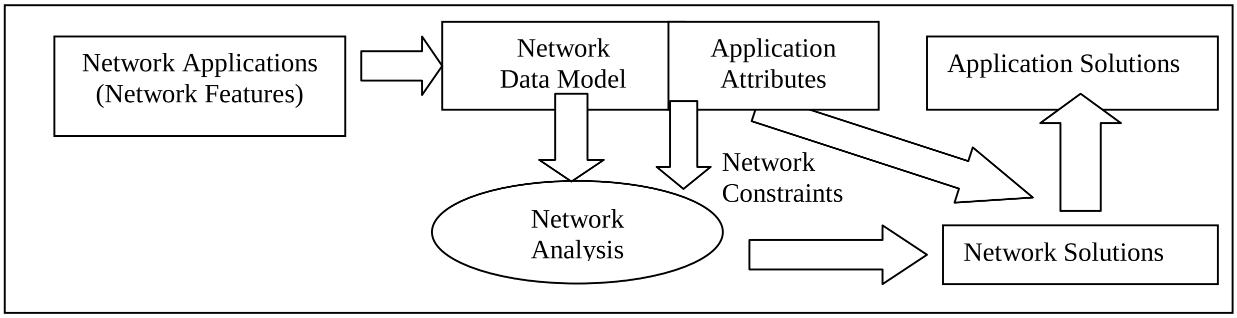


Figure 3. Network Modeling and Analysis Using the Network Data Model

Network Constraints

The network data model introduces the concept of network constraints, which provides a mechanism to guide network analysis. For example, you may want to compute the shortest path that passes through network links of a specific type. With network constraints, applications can easily incorporate application-specific logic into the network data model analysis engine without knowing how the engine works. Other constraints, such as path depth, or cost, can also be included in analysis. The network constraint is a Java interface and must be implemented by the application.

Java Representations of Network Elements

The Java network representations (network, nodes, links, and paths) are defined as Java interfaces and can therefore be extended. These interfaces specify the necessary behaviors for the network and its elements. In addition to these interfaces, user-defined analysis functions can be used by applications, allowing the modeling and analysis capabilities of the network data model to be extended.

Software Requirements

The Spatial network data model is now supported in standard DBMS technology. The PL/SQL package is pre-loaded in the database and required Java .jar files are provided; the Java API supports JDK (or JRE) version 1.3 or later. The network editor is also included as a utility tool for the network data model.

USING A NETWORK DATA MODEL

This section explains the usage of the network data model. There are three major steps.

Create a Network

1. Create and populate network tables and add metadata to the database.

2. Create a Java network object using the Java API and save it to the database.

Visualize or Edit the Network

1. Load a network from the database or an XML representation.

2. Visualize or edit the Java network object using the network data model editor. Store the network in the database, if needed.

Analyze the Network

1. Load a network into a Java network object.

2. Conduct network analysis. Save the path results, if needed.

GIS APPLICATIONS USING NETWORK DATA MODEL

GIS network analysis may include network tracing, network routing, and network allocation.

Network Tracing Applications

Tracing applications deal with queries like *Is node A reachable from node B?* or *What are the nodes that are reachable or can be reached from a given node?* Such queries are common in water or utility networks. Another type of tracing analysis is to find out how many connected components are in a network. Figure 5 shows some such queries.

Network Routing Applications

Routing analysis or path computation, probably the most studied topic in network applications, is divided into the following categories:

- Shortest Path or Fastest Path (transitive closure problem).

- K Shortest Paths: Find k shortest paths from a start node to a destination node.

- Traveling Salesman Problem: Find a minimum-cost tour that passes a set of locations.

Network Allocation Applications

Allocation analysis deals with designating destination points within a network. It provides information on a service area or coverage for points of interest. The network data model supports the following allocation analyses.

- Within Cost: Find all points of interest within a certain distance from a designated location.

- Nearest Neighbors: Find the *N* nearest restaurants to a designated location.
- Minimum-Cost Spanning Tree: Find the cheapest way to connect all nodes.

<u>Using Network Constraints in Analysis</u>

Constraints are conditions to be satisfied during analysis. The network data model supports network constraints so that applications can impose application-specific conditions on the network during analysis. The Java interface *NetworkConstraint* can be implemented by the user, and passed into any network data model analysis function. Figure 4 shows analysis information that is available for users to implement their network constraints.

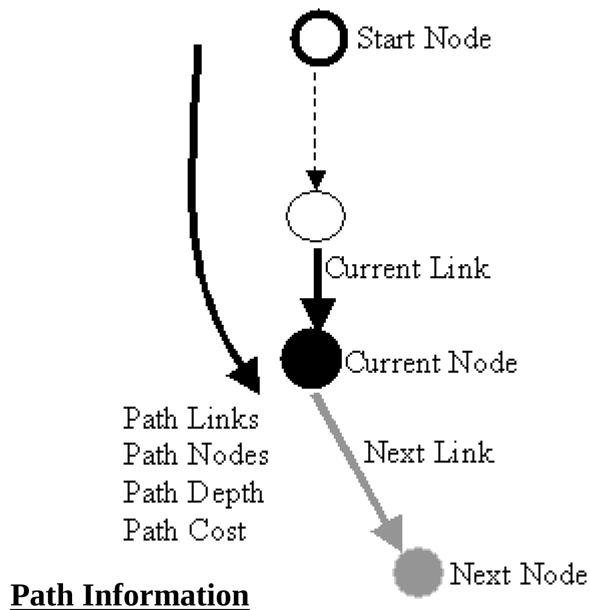


Figure 4. Analysis Information for Network Constraints

The following are some examples of network constraints:

- Depth (number of links), cost, and MBR constraints

  Network analysis can be limited based on the depth of the search path, the cost limit, or the area (minimum bounding rectangle) where the analysis occurs. These constraints can be used to specify a preferred subset of possible solutions. The network data model provides a *SystemConstraint* class (which implements the *NetworkConstraint* class) for these common network constraints. Users can create an instance of *SystemConstraint* and use it in analysis.

- Temporarily inactivated nodes or links

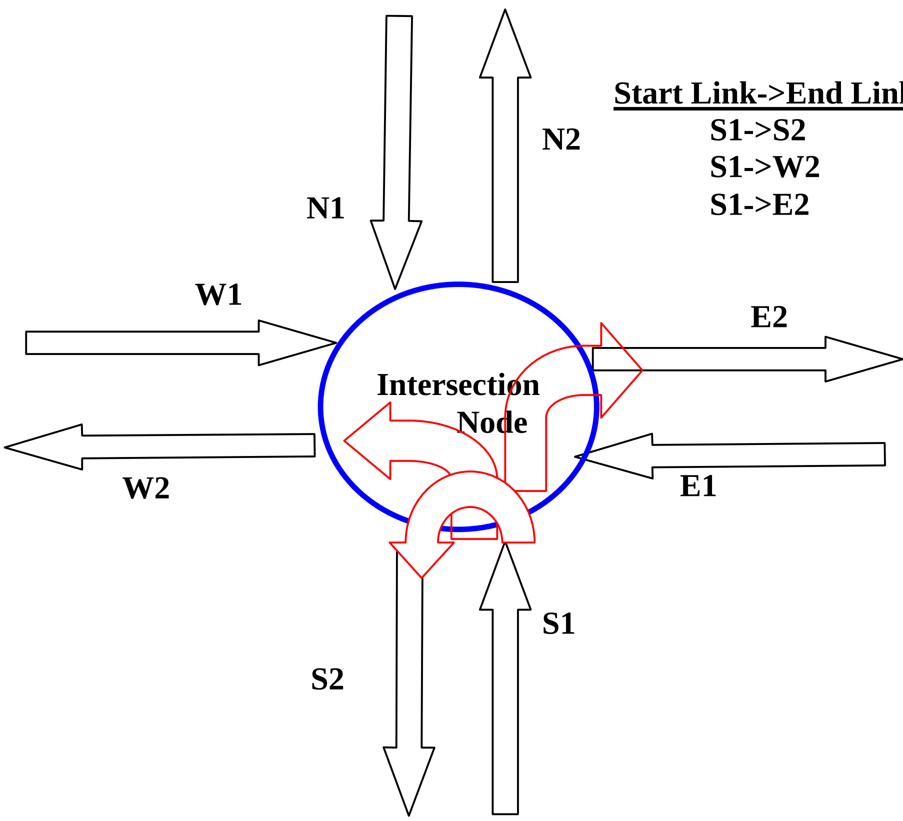
  Sometimes nodes or links must be temporarily turned off before analysis begins, for example, road segments (links) under construction in a road network, or water valves (nodes) shut down for repair in a water network. You can make a node or link inactive by setting its state to *false*. Network elements that are inactive will not be considered during analysis. Note that changing the state of nodes and links does not affect the persistent data model.

- Routing with specific types of links and nodes

  Sometimes network analysis must only be conducted through nodes and links of specific types or with specific requirements.

- Turn restrictions

  Turn restrictions are constraints involving two links. They are common in routing for transportation networks. In the following example, a prohibited turn is represented by a start link and an end link (see Figure 5). For intersections with turn restrictions, such as no "U" turn or no left turn, if the search encounters the start link of a prohibited turn, the search does not continue through the end link of that prohibited turn. This type of restriction can be easily modeled using *NetworkConstraint*, since information on the current link and next link is made available to users.



Start Link->End Link	Turn Type
S1->S2	U Turn
S1->W2	Left Turn
S1->E2	Right Turn

Figure 5. Turn Modeling for Road Networks

CONCLUSION

A native DBMS network data model is a generic network (graph) modeling and analysis environment useful for developing robust, scalable and secure network applications. GIS applications for routing, tracing, and allocation have been discussed. The network data model provides an open, easy-to-use, scalable, efficient, and manageable approach to developing GIS network applications. We are currently working with our customers and partners to extend the modeling and analysis capabilities of the network data model.

REFERENCES

*Oracle Spatial Topology and Network Data Models***,** Oracle Corporation.

*Oracle Spatial User's Guide and Reference***,** Oracle Corporation.

Oracle Linear Referencing System: A Technical White Paper, Oracle Corporation.

Oracle Spatial Network Data Model: A Technical White Paper, Oracle Corporation.