ABSTRACT

Researches and experiments in the field of transportation studies have proved that an effective and robust urban transportation system can favour sustainable urban development and optimisation of transportation facilities. Modelling urban transportation information is an engineering and research issue that has attracted the design of many conceptual and logical data models. However, the application of these models to urban systems is a challenging task especially when public transportation systems are made of several transit networks. This often leads to the development of specialised data models that do not favour the development of integrated approaches. This paper introduces an existing public transportation data model as currently developed in the city of Guangzhou in China, and a novel approach that integrates multi-modal transit network data, and multiple representations of network components. The main goal of our research presented is to develop a conceptual integrated GIS data model for a multi-modal transportation system. The resulting model is specified using an UML-based, object-oriented visual modelling tool integrated with spatial plug-in for visual languages (PVLs).

INTRODUCTION

As a typical metropolis in China, Guangzhou evolves towards a complex and multi-modal public transportation system composed by bus and metro transit networks. A survey of travelling behaviours made in 2005 showed that nearly 34.3% of local residents used public transportation services for their urban displacements. In order to meet the transportation demand, there are about 8000 buses running on more than 300 lines, and 4 metro lines to provide public transportation services, currently. However, there is still a massive imbalance between demand and supply of public transportation services after two decades of rocketing development of the city of Guangzhou. At the national level, bottlenecks in public transportation development in Chinese cities during the 90s were involved in not only an insufficiency of public transportation infrastructures, but also in an inadequate integration of public transportation services including policies, planning and management (Chen & Tan,
In the context of the city of Guangzhou, primary limitations of the public transportation system are as follows: (1) initial development was mainly oriented to bus-based services; (2) bus and metro services are not fully integrated into a concept of integrated transportation services; (3) public transportation services cannot match the travel demands of a large number of commuters whose main transportation mode is still pedestrian; (4) effectiveness of the existent transportation model need to be enhanced and (5) diffusion and manipulation of transportation data still imply to consider and deliver GIS-based information.

The research presented in this paper introduces a generic transportation data model whose objective is to reinforce integration of the public transportation services. Geographic Information System (GIS) appears as an ideal candidate system to design an integrated transportation service due to its capabilities in management, storage, manipulation, analysis and visualization of geographical information (Rodrigue et al., 2006). A specific domain of GIS, namely Geographic Information Systems for Transportation (GIS-T), encompasses the principles and methodologies related to the application of geographical information techniques to solve transportation problems (Curtin et al., 2003; Miller & Show, 2002; Claramunt et al., 2005). Nevertheless, successful development of GIS-T still implies and requires the design of appropriate data models to represent thematic and spatial relationships, and the underlying logical and structural properties of the transportation network systems. Properly defining the semantics, geometry, temporality, and integrity constrains of static and dynamic entities and relationships to be included in an integrated schema is an essential part of the data model design. However, transportation models are not easy to design as transportation-related data not only is supplied by diverse sources, but also encompasses multiple spatial, temporal and multi-modal characteristics. This implies to retain modelling approach which should provide an explicit and rigorous definition of spatio-temporal entities of the urban world to be modelled, as well as logical connectivity and topological relationships involved in the multi-modal transportation system. Recent proposals have made some noticeable progress in the search for a common spatial data infrastructures, such as the INSPIRE initiative (INSPIRE Drafting team, 2007), which introduce a generic conceptual model based on the conceptual schema language UML (Unified Modelling Language). The objective is to make a bridge between the conceptual and implementation levels in order to rely on automatic processes able to derive, for instance, UML application schema.

We develop an integrated GIS data model for transportation system. The modelling approach combines a UML-based design, and an object-oriented visual modelling based on spatial plug-in for visual languages (PVLs) whose objective is to support a conceptual data model for a multi-modal transportation system. The resulting data model is specified in a high-level UML class model schema extended with spatial and temporal stereotypes (Rumbaugh et al., 1999). This approach supports the design of GIS-T model with a special emphasis on spatial, temporal and spatio-temporal components provided by GML extensions.

RELATED WORK

Public transportation is an activity that is based on stable and well-defined directional routes along or intersected with the roadways. An integrated GIS data model for public transportation applications should explicitly represent topological relationships, not only between entities (e.g. static geospatial objects, spatio-temporal objects, temporal service schedules and etc.) within the transit network, but also between transit networks, road networks and surrounding non-network information (such as land uses and landmarks).

Preliminary work has been already initiated towards an integrated GIS data model for public transportation system in the city of Guangzhou (Chen & Tan, 2004). This public transportation GIS has been developed as a preliminary framework for the development of intelligent public transportation management system. Public trip guidance has been an area of initial interest for the delivery of services to the users. Such services include a variety of specific applications ranging from simple route planning to interactive trip planning.
Passenger information system integrating data on directional routes, fares and schedules, as well as related dynamic information has been also addressed for pre-trip planning (Casey et al., 2000). However, current representations of public transportations entities such as bus lines and stops are simplified in the form of non-directional lines and single stops (Huang, 2003). These representations need to be expanded for better system performances.

Elsewhere, several efforts have been made to validate a detailed spatial data representation for bus lines and stops. A pilot project was conducted in the Fairfax County, Virginia, to demonstrate applications of GIS in transit planning, operations and marketing (Jia & Ford, 1999). The authors introduced the development of a detailed transport-oriented database containing information on bus stops with references to roadways, roadway intersections, landmarks and surrounding areas (Sarasua et al., 1997). But multi-directional routes and multi-site stops are not taken into account. Another GIS-based data model that introduces bus multi-directional routes and multi-site stops for transportation data has been introduced (Huang, 2003). This model identifies stop locations, bi-directional bus lines, as well as relationships among stops, directional routes, lines and road networks. The advantage of this approach is that the bus transit network can be easily generated or extracted from the detailed definition to meet the needs of applications at different levels of representation.

While considering a single level of data representation and topologic relations with the road network, these proposals favour the design of mono-scale data models that inevitably cause negative effects on many specific applications, including transport planning, pre-trip planning and real-time travel guidance. In order to address the modelling of a transportation network at different levels and modalities, a multi-scale GIS-based approach should be considered.

TOWARDS AN INTEGRATED GIS-T DATA MODEL

Transportation entities are characterised by their spatial and temporal attributes whose dimensions should be considered as a prerequisite for the design of a spatial data model (Huang, 2003). Moreover, a multi-modal network system inevitably compasses complex spatial relationships, at different logical, temporal and spatial scales. This is reinforced by the fact that a multi-modal transportation system encompasses many different jurisdictions and stakeholders. Therefore, there is often a need to not only represent the topological relationships existing between spatial entities, but between networks, and between network and non-network data.

This leads us to retain a transportation data model that includes different scales and temporal granularities. For instance, a macro-oriented representation considers the transportation network based on a logical view that should rely on an arc-node model. At a lower level of abstraction, a micro-oriented representation consider a transportation network where the objects along the roadways are modelled not only as logical features but also as geometrical entities such as lane-based features (Malaikisanachalee & Adams, 2005). Integration of the logical and geometrical views should also be considered: regarding the modelling of the bus transportation system, bus lines and stops must be correlated to the existing road network structure. This should be done at the appropriate granularity level take into account various elements ranging from traffic restrictions to travel behaviours.

A remarkable advantage of an integrated data model is to provide an essential set of objects, features, and relationships that are integrated into a seamless and logical data model. It will favour the manipulation, analysis, visualisation and simulation of a transportation system (Chen et al., 2006). This implies to identify a modelling framework, supported by an appropriate methodology for that support the development of integrated transportation data model where spatial and spatio-temporal data properties should be analysed at the microscopic and macroscopic abstraction levels.

Moreover, with the dramatic development of positioning and telecommunication technologies, lots of dynamic data (such as trajectory data for moving buses) can be derived and collected
in real-time (Li & Lin, 2006). This makes geographic data integration, in the case of multi-modal transportation systems, a heavy task as data volumes are very large. These dynamic transportation data can be applied to multiple purposes, such as navigation, buses dispatching, passenger and traffic flows analysis, etc. And one fundamental requirement of these applications is to represent multi-modal transportation system at appropriated granularity levels. Therefore, an efficient management of these large volumes of dynamic transportation data have also to be addressed and should rely on filtering algorithms able to retain only relevant information (Li et al, 2006).

**UML-BASED VISUAL MODELLING OF A TRANSPORTATION SYSTEM**

UML is an object-oriented standard language developed by the Object Management Group (OMG). UML is used to specify, visualize, construct, and document software systems (Object Management Group, 2005). The object-oriented technology not only favours software analysis and design, but also facilitates the development of GIS-based data models. Spatial UML, an object-oriented modelling language for geographical information applications, is based on UML built-in extension mechanisms (i.e. stereotypes, tagged-values and constraints) which are used to define and illustrate geospatial object classes and their inheritances and relationships by using class model diagrams (Liu et al, 2004). Bédard and his colleagues developed a spatial modelling CASE (Computer-Assistant Software Engineering) tool, namely *Percepotry* (Bédard & Proulx, 2006) that plugs Spatial PVLs into UML class model. This CASE tool extends UML by using “stereotype” constructs for high-level conceptual class modelling of spatial-temporal databases with multimedia capabilities (Bédard, 1999).

![Figure 1: Basic pictograms of PVLs](image)

Spatial PVLs is a graphical-based language that represents geometric, temporal and visual properties of objects and attributes (Brodeur et al, 2004). It offers several components to define the geometry chosen for the spatial elements of a class model. The main objective is to explore a high-level abstraction to facilitate analysis and conceptual spatial-temporal modelling without considerations of implementation issues.

The spatial and temporal PVLs use five basic constructs represented by pictograms as shown in figure 1. In order to properly represent the spatial and temporal characteristic of objects and built-in attributes, these basic pictograms can be combined or extended with cardinalities in view of the standards of object-oriented visual modelling languages. Figure 2 illustrates several (but not exhaustive) variations of these pictograms for visually representing the complex, alternative, and multiple geometry or temporality.

### Simple geometry

0-dimensional (0-D) geometry, point

1-dimensional (1-D) geometry, line

2-dimensional (2-D) geometry, area

### Simple temporality
A spatial evolution can be informally defined as a change of characteristics of a given object in a geographical space (Etches et al., 1998). Two reference temporalities, i.e. instant time (●) and interval time (□), might be applied on attributes or geometries to define the spatial evolution or existence of a given object. This interval pictogram, denoted by □, defines the “age” of an object (entity or attribute) and implicitly includes the definition of “birth” and “death” as bounds of time interval (Bédard & Proulx, 2006). Such a pictogram can be applied for instance to metro stop schedule which is defined by valid period and dates bounding this period (e.g. holidays). Nevertheless, the spatial or spatio-temporal evolution of an object might be represented by a combination of geometry and temporality pictograms. For instance the combination □ depicts the capability of an object to have a temporal value.

We retain a formal description by applying an UML-based conceptual schema with a special emphasis on spatial, temporal and spatial-temporal components provided by PVLs extensions. Moreover, in order to facilitate the comprehension of stereotypes, we define, in accordance with the definitions of Perceptory, a representation of this concept into a UML class diagram: (1) a spatial stereotype of a class is placed on the left side of the name of object class; (2) a spatial stereotype of an attribute is placed directly nearby the defined attribute; (3) a temporal (existence) stereotype of a class is placed on the right side of object class; (4) a spatial evolution of a class is placed on the left side of the name of object classes instead of the spatial stereotype; (5) a descriptive evolution stereotype of an attribute is placed directly beside the defined attribute; (6) a spatial evolution of an attribute is placed nearby the defined attribute.

Figure 3 illustrates a class diagram where the spatial PVL’s pictogram (□) represents that a metro line is composed by two 1-D lines with opposite directions. The attribute “status” is defined by a “date of opening” and a “date of close”.

**Figure 2: Pictograms for spatial and temporal information**

<table>
<thead>
<tr>
<th>Pictograms:</th>
<th>Example:</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ ]</td>
<td></td>
</tr>
<tr>
<td>[ ]</td>
<td></td>
</tr>
</tbody>
</table>

**Complex geometry**

Hydrographic networks composed of 1-D lines.

**Alternative geometry**

Buildings having a C-D shape if< 1 hect.

**Multiple geometry**

Pictogram of multiple geometry for an instance:

Polygonal municipalities having a non-

**Pictogram of quantity of geometries for an instance:**

(A cardinality of 1:0 means a facultative shape, 1:1 means a necessary shape, 1:* means a specific shape.

**Example:**

“n” – Road networks are composed of several edges.

**Multiple temporality**

A bus has a durable time when it stops at a bus stop, an arbitrary time of a purualtural time when it passes a bus stop.
A metro line may include one or more branches, passengers can transfer between arterial line and its branch at the junction.

Moreover, a spatial-temporal definition of the attribute “construction” indicates that a metro line might be under construction and can be mapped progressively while allowing the user to record the new geometry and position as well as the corresponding period.

A GIS CONCEPTUAL DATA MODEL FOR MULTI-MODAL PUBLIC TRANSPORTATION SYSTEM

A public transit network encompasses complex spatial topological relationships and logical connectivity among and between the entities. Without considering topology and connectivity, the transit network is simply a collection of lines with severe drawbacks to the transportation GIS community (Li & Lin, 2006). The proposed GIS-based multi-modal transportation system is designed using a formal object-oriented approach relying on the conceptual schema language UML. This graphical modelling is based on the PVLs model which brings a visual description of the spatio-temporal components and the properties through specific pictograms.

The UML class model schema in figure 4 describes the spatio-temporal characteristics, topological relationships, logical connectivity and location references of the multi-modal public transportation system. At the higher abstraction level, the schema describes a Bus Line which is decomposed into two Bus Routes. Each Bus Route is defined by a directional path and is represented as an oriented path on the road network. Additionally, an individual stop site on a Bus Route is defined by a Bus Route Stop generally represented by a simple point data type. The relationship between two Bus Stops is defined as a Route Segment. A Bus Platform allows passenger flow connections and is typically composed by at least two Bus Stops sites across a roadway. The Bus Platform includes two spatial data types, i.e. point and polygon that represent the logical and geometrical views, respectively. The point data type is derived from the Bus Route spatial data type. A Bus Route acts as key location reference for the management of trajectory data of moving bus and the application of network analysis. The Bus Stop object is specialized to connected Entrance/Exit object to model the logical connectivity between the bus and metro networks.

Similarly to bus transportation system, the metro network rely, at a macro level, on Metro Line entities decomposed into two Metro Ways where each way is represented by a single directional line data type. An individual stop site on a Metro Way is defined as a Metro Way Stop and is represented by a simple point data type. Entrances and exits on the metro network are related to a Metro Station that includes at least one Metro Platform giving access to the line of metro through one of the two Metro Way Stops (one per direction).

Let us consider a passenger information system that provides pre-trip guidance analysis across different public transportation modes, the model presented in figure 4 shows that the road network acts as a physical connection between different transit networks via entrance/exit entities.
The analysis of the logical and physical connectivity, topological structure and multiple-representation of network system should therefore rely on the urban road network. Figure 5 shows the respective roles of the physical and logical entities of the metro transit network. These entities are modelled as UML classes, and additional topological relationships that relate them (Figure 4).
The UML schema of the multi-modal public transportation system provides an explicit representation of spatio-temporal objects and their interrelationships. This favours a visual, understanding, and comprehensive representation of the logical connectivity and topological relationships between the transit network objects. Moreover, this conceptual schema provides a graphic view of the transportation spatial-temporal data model. Such a medium offers an efficient communication mean between data model designers and transportation experts (Etches et al., 1998). In fact this schema is the result of an interactive design process that has implied discussions, exchanges, and at least a common understanding of the transportation application semantics.

CONCLUSION

The rapid development of the whole aspects of modern cities requires to model urban transportation data at appropriate abstraction levels. The main goal of the research presented in this paper is to provide an integrated data model for the multi-modal public transportation system at appropriate scales and levels of granularity. Such a multi-modal network system provides an essential and fundamental support to enhance integration of public transportation services within integrated GIS.

Combination of GIS-T and UML are increasingly used to model and represent transportation network system. Modelling public transportation data within GIS for a variety of transportation applications is a domain of crucial interest for transportation studies. We introduced a UML-based, object-oriented visual modelling tool integrated with Spatial PVLs. This facilitates the development of a reference model for visualizing and representing the conceptual data model that will be the first integrated multi-modal GIS-T data model for the city of Guangzhou.

The design of our transportation system model currently includes the modelling of bus and metro network as well as first steps of urban network modelling. Remaining work implies to (1) refine the multi-scale component of the road network data model; (2) to integrate transit networks and urban data models; (3) to apply and validate the modelling principles to the city of Guangzhou through the conception of an integrated GIS-T, whose objective will be to improve multi-modal transportation planning and facilitate sustainable urban development.
REFERENCES