GTF Specification Overview

Over and synthesis of the

“Generalised Transportation-data Format” Specification

Authors

Dr B. Mandel, MKmetric GmbH, Karlsruhe, Germany

Prof. O.A. Nielsen, CTT (Centre for Traffic and Transport) at the Technical University of Denmark (DTU), Denmark

E. Ruffert, MKmetric GmbH, Karlsruhe, Germany

Contents

1 INTRODUCTION

The contents of this document was extracted from the paper submitted by Prof Otto Anker Nielsen and Eduard Ruffert to the PTRC for the "European Transport Conference" ETC, 10-12th September 2001.

Exchanging data and information on the data (meta-data) between transport models, as well as between transport models and other software, e.g. GIS, is always a very tedious, if even possible, task. There is often the problem of loss of information because the exchanged data only seemingly contains the information required. And there is also often the problem of inhomogeneous and proprietary data formats forcing the users of the data to re-format and re-combine the data from scratch every time.

This is both due to "low-level" differences in data formats, and due to more fundamental "high-level" differences in the conceptual models, e.g. for network topologies. Examples of the latter are the differences in describing a terminal by transfer tables versus by a sub-network, or a public transport network by timetables referring to the same line, versus by parallel arcs for each departure.

The solution to these problems is that not only data needs to be transferred, but also the precise meaning of the data (meta-data), including the underlying conceptual model. The "Generalised Transportation-data Format" GTF, based on the original work in Mandel & Ruffert E. (1999 & 2000) was developed to meet these demands (Note that the name GTF, especially the "Format" part, stems from its origin trying to find a common format. This subsequently evolved to a specification of a conceptual model, yet the name GTF was retained).

GTF is a proposed conceptual model (covering the most widely used objects in transport modelling), an exchange format (GTF-XML) based on standard XML, and an interchange language to run transportation models and retrieve results. This allows software applications, "GTF Translators", to exchange information and data between transportation models and other software.

The work started in the EU-research project BRIDGES where a survey of different conceptual models and formats was carried out (Nielsen *et al*, 1998). This lead to the first version of GTF (Mandel & Ruffert, 1999). The work is continued and refined in the thematic network: SPOTLIGHTS under EU's 5th framework programme, where further surveys, reviews and user input are carried out.

As SPOTLIGHTS is funded by the EU, it is the ambition that GTF eventually will become a EU-standard for the exchange of transport modelling data. This will provide a strong platform for utilising earlier work and transportation models when building new transportation models, as well as a tool to compare transportation models that cover the same geographic area. Both aims will be very useful for research as well as practice in the field of transport modelling.

2 CURRENT SITUATION AND PROBLEMS

The usual use of strategic transportation models is to define changes in the input data for each scenario to be analysed. The Input defines "Policy

Scenarios", like economic, demographic and spatial developments as well as network changes and changes in prices and fares for the use of transport supply (Eurostat, 1996).

2.1 Software and transportation model issues

Currently, the numerous software applications and databases used in practice are often inhomogeneous and largely incompatible with each other. This leads frequently to problems when comparing results from scenarios based on different software applications and databases.

Transportation modelling at the European level usually requests data from many sources and transportation models – often at least from each country. But often also from different sectors, e.g. road administrations, rail authorities, bus operators, ferry companies, airline systems, etc. This also applies to national transportation models. In some circumstances even further data from non-national sources are needed, e.g. from counties or even municipalities concerning road network data.

As an example, the national road administrations may only maintain databases of the national roads. Since the motorways and highways often end outside harbour cities, the omission of municipal roads can result in large detours in a transportation model. Even some motorways may be owned and maintained by counties, municipalities or private companies.

As such, there are many benefits in integrating data from different sources and at different quality levels. However, figure 1 illustrates the possible problems doing this. This includes:

- Transportation models will often be inhomogeneous in their conceptual structures, which makes coordination difficult. Furthermore the data models are in some cases not transparent; e.g. software packages are for competitive reasons not fully documented, or they have not been documented properly due to time- and budget constraints.
- Software packages have inhomogeneous formats (even if they build on similar conceptual models).
- Some metadata are implicitly given by the software package, and some by the data model; e.g. that an organisation always uses the same unit definition and data collection method. As such the unit definition, quality, year etc. are not stated explicitly in the data itself.
- Translators are not always sufficient; data may have been aggregated during export, some topological relationships may have been lost during translations, metadata is not exchanged, etc.

- When data from different sources are combined into one conceptual or data model, there are a number of consistency problems as well as problems stemming from different levels of aggregation.

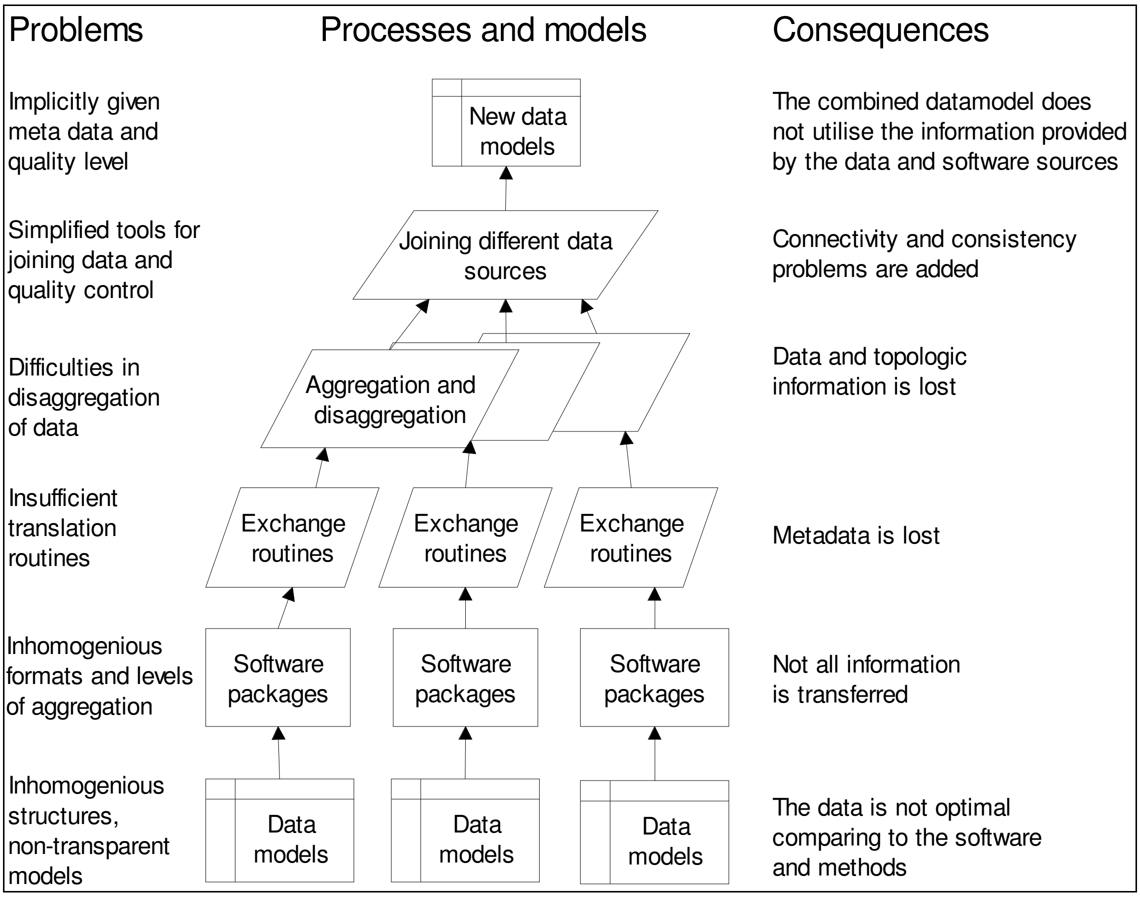

***Figure 1**. Possible problems integrating different models and data sources.*

The problems also apply to the databases of the results from transportation models (not only on the input databases), and hereby to the comparison of results from different transportation models for different projects, or for the sake of quality control.

2.2 Problems due to data of transportation models

Even if the above technical problems are solved, problems may still prevail. Transportation models are in general very demanding concerning the amount and quality of input and calibration data. The main problems with current data and databases are:

- Data required by the transportation model, e.g. for estimation, is not available. For example, a pan-European passenger transport model requires homogeneous input data from all countries.
- The composition of the available data required by the transportation model does not match and re-composition is not possible. For example, the data acquired for a transportation model has different levels of aggregation or use

different segmentation, that cannot be matched to the one needed by the transportation model.

- The data itself does not match, e.g. that units have been defined differently without an easy way of reformatting this. An example is traffic counts as weekday traffic defined as September to June average, versus traffic counts as Annual Daily Traffic (ADT).

3 VISIONS BEHIND THE GENERALISED TRANSPORTATION-DATA FORMAT (GTF)

Because of the problems mentioned in section 2, the value of transportation models' databases can be significantly increased by homogenising them and by defining an openly available specification of the homogenised conceptual model. The first (and main) advantage would be to have databases, which can be exchanged, enriched, corrected and used in a transparent manner since all would be based on the same conceptual model. Secondly, it can be ensured that the required information is actually contained in the data and that the information can be exchanged. The structure of a "Generalised Transportation-data Format" accordingly accomplishes the following:

- Instead of having disparate and manifold software applications and databases, GTF contains all necessary elements and provides one single and homogenous data specification and format.
- Instead of having incompatible proprietary formats and informational contents, GTF should be used throughout any computer system, by providing translators to / from the proprietary formats to GTF.

⇒ an implemented GTF format serves as a mediator format between disparate other formats

To achieve this, GTF consists of:

- A conceptual model (GTF-CM, called GTF-Conceptual Model). This defines the framework for a given transportation model, while it does not contain the data within the transportation model and the implementation of the transportation model (i.e. it does not constrain any implementation for example as relational tables or as software in anyway).
- A standard exchange format (GTF-XML), including meta-data as well as the data itself (i.e. "tags" encapsulating raw data giving it meaning).
- Generic commands to run transportation models and retrieve results (GTF-TIP, "Transportation-data Interchange Protocol").

3.1 Basic concepts

Basically, GTF is a framework, which can be used to define the information that is contained in data. It wraps data into information classes describing the basic data and the necessary supplementary information (meta–data) to give a meaning to the raw data.

A potential problem is, that most transportation models, standard software, and exchange formats define data with implicit information, where only the developer or in the best cases the practitioner with good knowledge of a well-written documentation know the exact definition of a data element, e.g. speed. This needs further definitions, *metadata*, to be defined precisely, e.g.:

- What *type* of speed; free flow, at congestion, in average, measured, modelled, signed?
- At what *level of aggregation*; for all lanes, for passenger cars, rush hour, weekday average, all week average, yearly average?
- *Quality*; measured at each link, judged from road category, guessed on intuition, and method of establishing the data?
- *Origin*; what is the year of data measurement and updates?
- *Organisation*; who established the data?

3.2 XML

In GTF, XML (se e.g. Marchal, 1998 or Booch *et al*, 1999) is used as a framework to ease the definition and exchange of data. The ideas behind XML are a bit similar to those of object-oriented programming (see Brown, 1997, Budd, 1997 or Rumbaugh *et al* for an introduction to object oriented concepts, or to the brief introduction in Nielsen & Frederiksen 2001b). Accordingly, the GTF-XML file includes class instances and definitions on the relationships between them. The main advantage of this is the minimal amount of different abstract concepts used to cover a wide range of concrete things.

The GTF specification defines its exchange format as an application of XML.

3.3 Main classes in GTF

Very generally speaking, most transportation models use the following information items for their computations (although some transportation models have more advanced input requirements):

- <u>Zonal data:</u> any kind of zonal description, e.g. socio–economic data, ecological data, zonal boundaries, transport data, indicators, transport matrices etc.
- <u>Network data:</u> data describing the relations between the elements, e.g. link characteristics, a link has a starting node and an ending node (i.e. topological characteristics), link/network clusters etc.
- <u>Transport data:</u> data describing services in the public transport, pre-defined routes, etc.
- <u>GIS data:</u> the necessary information for visualisation purposes, e.g. the underlying projection of the node and its co–ordinates.

The *basic classes (or business layer classes)* that were used to create the conceptual model are a total of only 11, namely *Node, Factor, Link, Mode, Vessel, Chain, DynamicSegmentation, Alternative, Unit, Group* and *Meta*, which are called "topmost classes" or "top–levels".

The top–levels and their children can be combined using the defined relationships.

3.4 GTF Data Pool

With GTF, the structure of the numerous software applications and databases would be accessible in a homogeneous and compatible manner. A set of GTF Translators would provide a single access point to all transportation models and data, see figure 2.

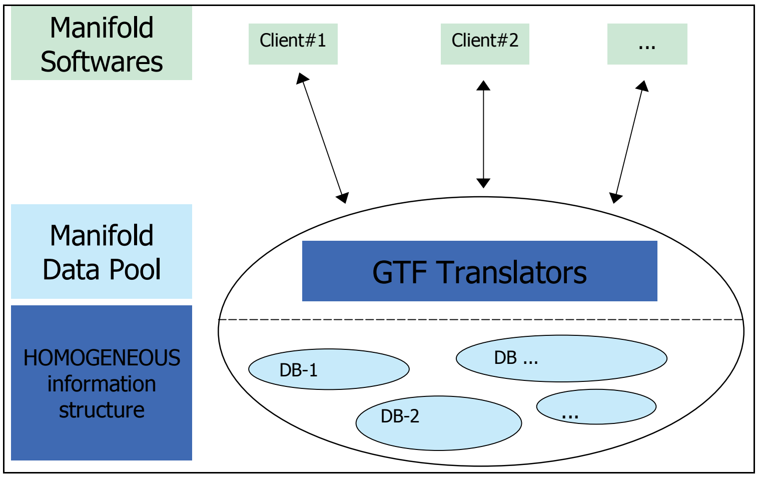

Figure 2. *GTF data pool.*

The numerous databases can either be restructured according to GTF's conceptual model. Or a specific GTF Translator for each database could be developed providing a homogeneous and single access possibility.

3.5 Implications / ramifications of GTF

The impact of GTF has many ensuing commercial and practical benefits:

- Synergy effects emerge from the possibility of transferring knowledge between systems.
- It will be possible to compare different transportation models' results (and their quality) as the transportation models can be used on the same data(base).
- Transportation model users may avoid to (re)create their own databases over and over again like in the past, but will have access to standard data(bases).
- Data(bases) will gain in quality as time passes, because the data providers will have an incentive to update their databases regularly and properly, since only the "good" databases will be used.
- Users will request new transportation models or combination of transportation models, which previously could have been denied by the consultants, because of lack of transparency on the business.
- The clients / users will have the possibility of choosing and combining transportation models.

- People dealing with problems appearing in different working areas can exchange information, e.g. to analyse side effects when changing from a higher to a lower aggregation level.

All these effects will have a vigorous impact on research in the transportation modelling and other fields.

3.6 Comparison with other transportation models and formats

GTF is designed to be a general conceptual model (for data) and a format, mainly addressing the demand for strategic - and hereby often multi-modal – transport models. GTF can describe domain specific objects, e.g. by using subclasses. But this is not predefined as detailed as in some specialised formats, e.g. to describe details in rail switches or turn lane geometry at road intersections.

The next figure depicts (rudimentarily) the data models and other specifications gathered and evaluated in the process of defining the initial GTF data model.

This specification will be submitted to all participants of the GTF/eConference and GTF/Workshop. After these two events a GTF Version 1.0 and eventually a Memorandum of Understanding signed by all participants will be available.

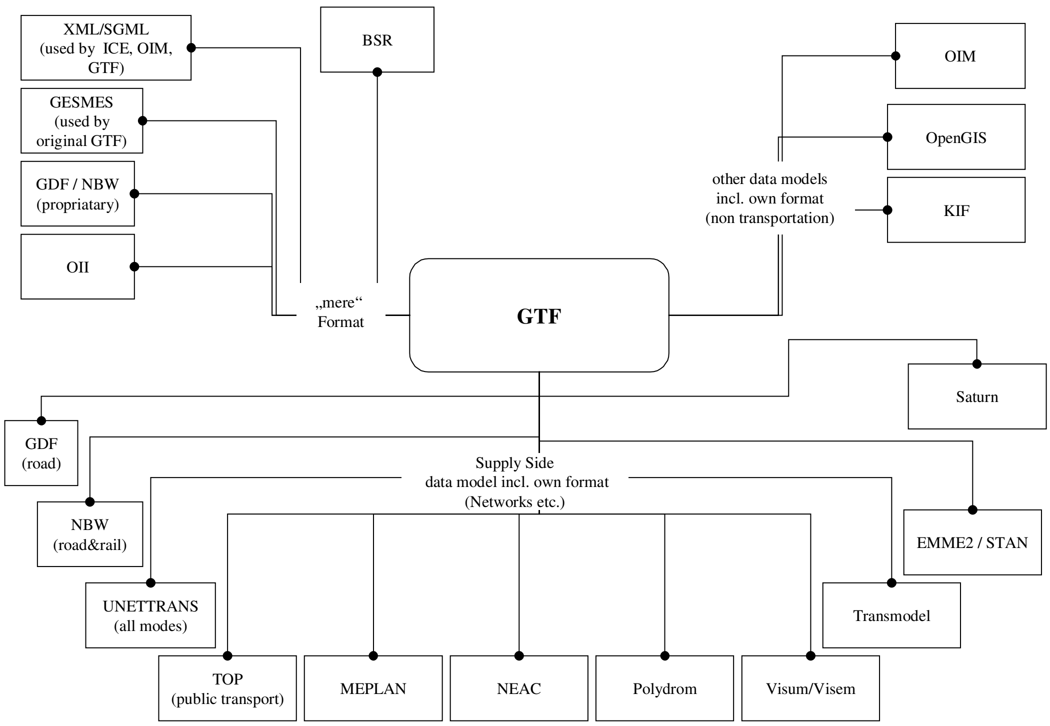

Figure 3. GTF and the evaluated DMs and formats

For comparison, figure 3 suggests the domain of application of GTF and other conceptual/data models and/or formats.

The Transport Object Platform (TOP) is a conceptual model and its implementation for ArcGIS 8 systems. It has been inspired by the work in the

BRIDGES project. TOP is in its predefined version less general than GTF. But its object model is completely open; users can add objects – even parent objects – that inherit, connect, or relates to other objects in the conceptual model. TOP is mainly developed as a general model for multi-modal transport. Furthermore it includes domain specific objects, e.g. turns at road networks, stops and terminals in transit networks, and complex demand for freight transport. Nevertheless it is less detailed than GTF in the way, that it only considers topological objects. It is up to the user to define attributes.

The UNETRANS transport data model (http://www.ncgia.ucsb.edu/vital/unetrans/) was also developed in relation to the ArcInfo GIS (and with some coordination with TOP). UNETRANS is, to a higher extent than TOP, a pre-defined model. Even many attributes are predefined. Its focus is mainly infrastructure (rather than transportation), with details especially concerning road networks. On the other hand the description of topological relationships in public transport is less comprehensive than in TOP.

The European TRANSMODEL for public transport (CEN-norm prENV 00278021) is a detailed – as well as large-scale – oriented model for public transport – especially bus-networks. It is more comprehensive within this domain than e.g. TOP version 1.0.

The European GDF-format mainly for road traffic (http://www.ertico.com/links/gdf/gdf.htm) has a higher degree of pre-definitions than UNETRANS. Although GDF is a format, and UNETRANS a model, their underlying conceptual models have similarities, as GDF was reviewed before defining UNETRANS

The RAIL model being developed as a counterpart to TOP is a detailed object-oriented model for rail infrastructure building on ArcInfo. It includes a number of domain-specific objects such as switches, signals, control sections, blocks, etc. RAIL is being co-ordinated with TOP as some TOP-objects can get aggregated information from RAIL, and RAIL can disaggregate information from TOP (e.g. the delay along a path from a rail-simulation can be aggregated into the Time-Table within TOP).

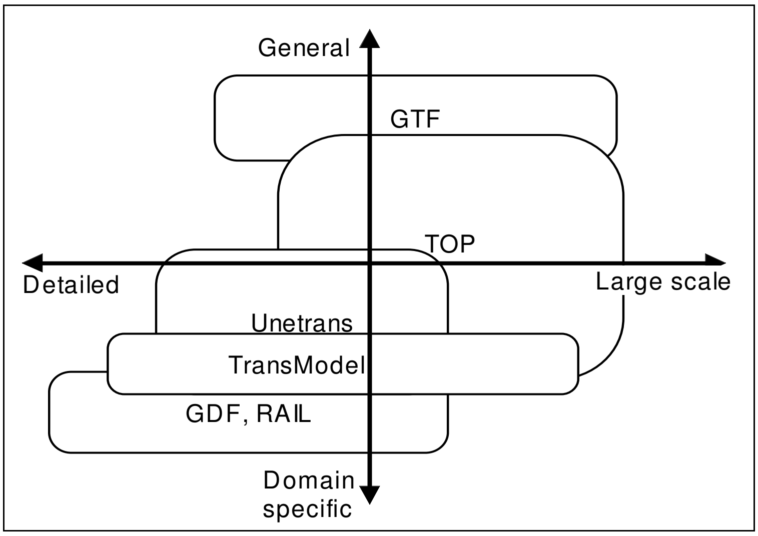

Figure 4. *Domain of different models and formats.*

Finally, commercial transportation modelling packages could be classified within the framework in figure 4. The main difference is, that their formats are less open and that the conceptual/data models are predefined to a large extent. Detailed domain-specific software are e.g. rail and road simulations packages. Most transport modelling software is fairly general, with different specialisations from detailed modelling to large-scale. Some are even comprehensive covering multiple scales. GIS-packages are typical general packages, with few predefined domain-specific objects for the transport sector (Neither for detailed nor for aggregated purposes). This was the background for the development of TOP, RAIL and UNETRANS as extensions to ArcInfo. TransCAD (developed by the US firm Caliper) is an exception, since it both has GIS and transportation modelling capabilities tailored for the transport sector (http://www.caliper.com/tcovu.htm). In addition, “GIS-like” features are emerging in some commercial transport modelling packages.

4 CLASSES IN GTF

This section introduces the fundamental classes that are the foundation of the GTF conceptual model. The transport data that is covered is primarily data used in strategic transport models. Thus it covers interurban, regional or international travel on all transport modes for both passengers and freight. More specifically, the meaning of “Transportation Class” in this paper is:

- Transportation = “The act of moving passengers or freight in space.”
- Transportation Class = “All that produces (generates or attracts), enables or hinders movement of passengers or freight.”
- Transportation Relationship = “The connection between two transportation classes.”
- Transportation Attribute = “A quality or feature of a transportation class that is a central part of its nature distinguishing its instances.”

The definition of e.g. Factor in GTF contains not only the raw data, but also the meta-data, e.g. “statistical source = EUROSTAT, type = statistics”.

The definitions above cannot be used for direct implementation. The goal of these definitions is to be able to define a conceptual model of transportation and not to implement this as a data model. The implementation is left to eventual providers who have to adopt GTF as one of the exchange formats of their software/model.

Figure 5 depicts the main conceptual classes used in transportation modelling information (without going into details).

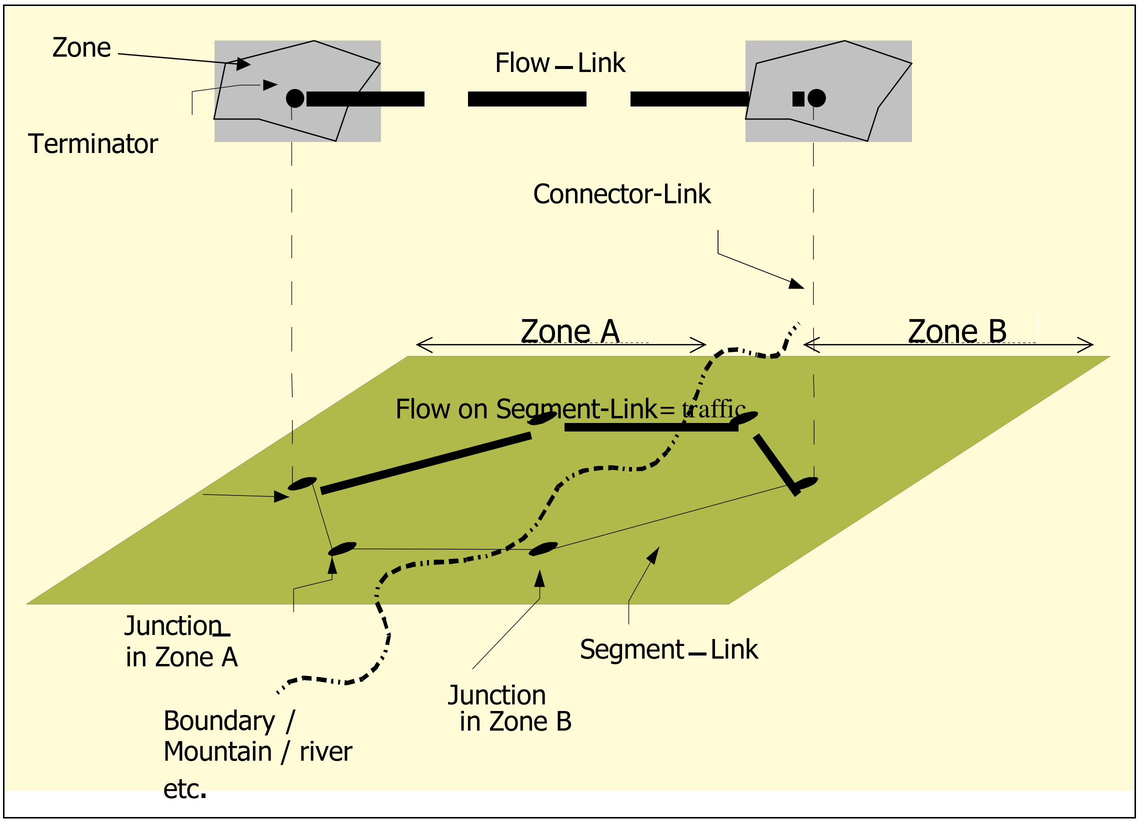

Figure 5. *Problem domain overview.*

A Terminator is a virtual point for input & output (source & sink) of movement in networks. It is associated with Zone, which contains the Factors of an area. In many transport models, the concept of a centroid is used to describe the same as Terminator in GTF. However, since centroids in some GIS implicitly are the geometrical centre of the zone (rather than the activity based centre like in transport model), the more general word Terminator is used. A Terminator is connected to an infrastructure network through a ConnectorLink.
The ConnectorLink is the virtual description of the impedance(s) that is needed in average to enter / leave a Zone and thus creating inter-zonal transportation / movement called LinkFlow.

A LinkFlow is the result of Factors generating and attracting movement across the limits of Zones. It can therefore be described as a connection (relationship) between two Zones. This is “flow” e.g. in the sense of demand for transportation. Thus, a LinkFlow is a connection between Terminators with information about the amount and types of flows (vehicles etc.) between the two Terminators. “flow” in the sense of observed movement is an attribute in the GTF Conceptual Model attached to a Link (or Segment).

A Node performs two functions in transportation modelling. The first function is to relate (connect) a Zone to some point in the network as access and egress points for mobility (the Node being a Terminator). The other function is that of being a Junction in the transport network. For generalisation purposes, Zone is a derived class from Node, too, as Zone’s can be starting / ending points for Links not only Terminators. (Terminators –centroids- were introduced in the transportation modelling field as a necessary entity for software/computational purposes.)

A Link is a topological relation between two Nodes. The Nodes in turn usually are associated to specific geographical co-ordinates in real world space. But this is mostly needed for visualisation and presentation purposes.

Following this kind of logical decomposition and analysis, 11 top level classes were defined in the GTF: Node, Factor, Link, Mode, Vessel, Chain, DynamicSegmentation, Alternative, Unit, Group and Meta. Using these high-level definitions further child-classes are defined in the conceptual model. The next table gives definitions for each top level (Mandel & Ruffert, 2000, elaborates further on the table, while Mandel & Ruffert (2001) describes all sub-classes.

Class Name	Description
Node	The generalisation of the concept “start or ending point of Links” and thus a generalisation (class) of the Terminator, Junction and Zone classes. The generic class Link is determined by exactly two Node classes. This secures a more homogenous view on the problem domain.
Factor	These determine the generated or attracted movement of a Zone, which together induce the demand for movement and transport. Factor objects contain for example the GDP, age distribution, level of income etc. for one Zone, a group of Zones or an aggregation of Zones. Additionally, since Factor is a subclass of Meta, any Factor object can be associated to any other GTF object in a data set, not only to Zone but e.g. to Node or Link etc. as many transportation models attach generation or production factors onto

Class Name	Description
	Nodes (Junctions etc.) or Links.
Link	The Link class is not only an abstraction for all types of infrastructure network links, but it incorporates the connections to Zones (through their Terminators). Terminators, Junctions and Zones in different combinations act as Nodes to define three possible types of Links: 1. the Segment (LinkInfrastructure) connects two Junctions in the transport network 2. the LinkConnector between a Junction and a Terminator describes the disutility to reach (any) point in the Zone from the main transport network 3. the LinkFlow between two Terminators or Zones is a Link that holds the flow information that results when two Zones to describe the movement between two areas in space. For technical reasons this class is actually named "MatrixElement".
Mode	A Mode is the type of immobile infrastructure used by Vessels for the transportation of Units from Zone to Zone or between Junctions etc. i.e. on Links.
Vessel	Vessel is the abstraction of everything that move on Links. In transportation models typical Vessels are cars, trains, aeroplanes, trucks etc.
Chain	This class represents the abstract concept of sequence of Links or activities. For example, a child class is Service that provides a traveller with the means to travel with relevant choices already made in advance by the service operator. The Service class defines the type of service, the used carrier Vessel(s), the level of security attributed to this type of service, and the timetable for the service.
DynamicSeg mentation	Contains information of milestones, e.g. their position (distance from starting Node and distance form ending Node) and other data that is attached to a specific point of a Link Use this class to associate of dynamical segmentation of e.g. attribute values along a Link.
Alternative	Transportation models use choice alternatives (e.g. usage of road or rail or air mode for transportation etc.) to describe the situation the behavioural units face in certain situations. From a transportation modelling point-of-view the networks (groupings of Nodes, Links etc. which form a logical whole) need often to be distinguished according to different "main modes". To broaden the definition, the more precise term Alternative defines "choice alternatives".
Unit	Units define the type of element being moved or transported (The purpose of the movement or the date / time schedule of a movement are stored in Meta.)
Group	This class can be used to group any class, class instance in order to define "result sets". This class is not like the others in the Toplevel. It is simply for grouping purposes. The other Toplevels contain business logic, e.g. topologic information etc. To add a level of semantics for the grouping one of the children classes should be used.
Meta	Metas are objects to define meta-information describing dimensions of measurements etc. The Metas can be used to associate dimension information with all/any other class instance.

Figure 6 depicts the top-level objects and their relationships in an UML diagram.

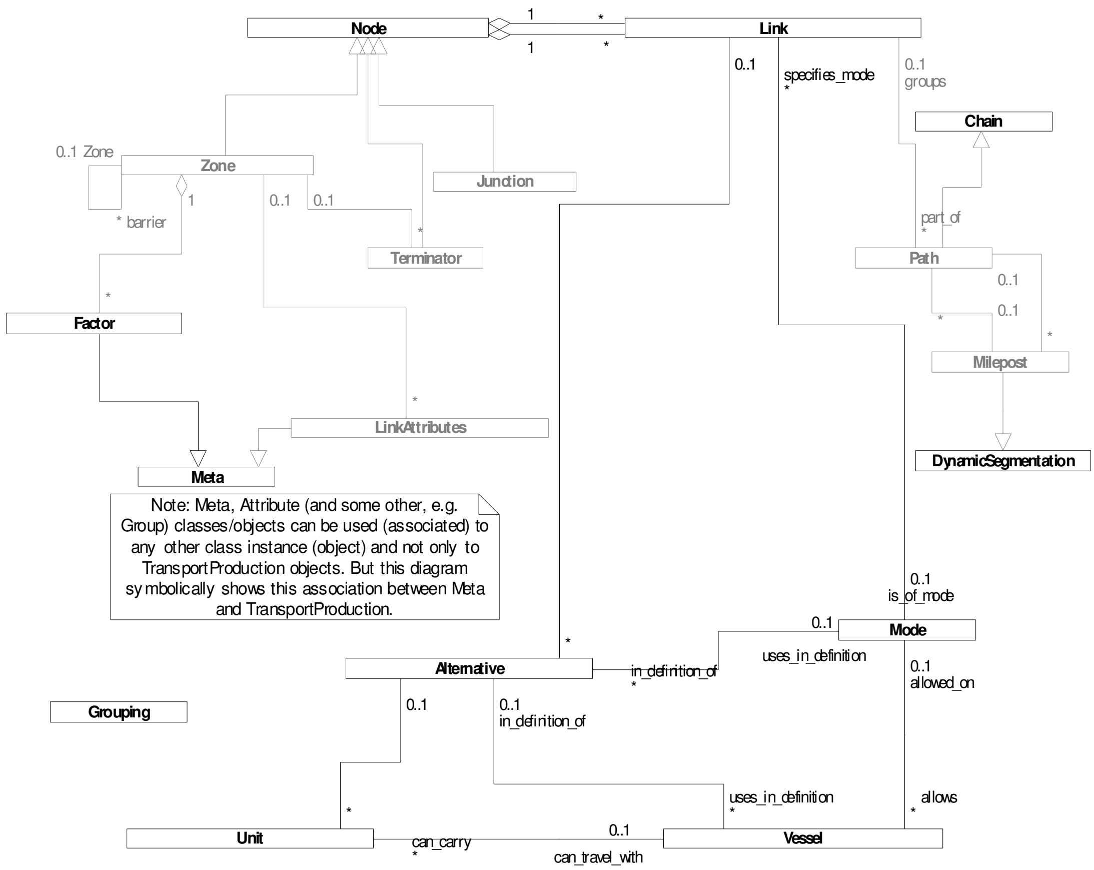

Figure 6. *Overview of the GTF model class and relationship structure.*

Note that all classes described in the model are instances of the class GTFObject that has a GIS pointer and a KIF conceptual pointer. The GIS pointer should be used to point to (an external) GIS object, e.g. a contour of polylines object. The KIF (Knowledge Interchange Format) can be used to contain a piece of text in KIF syntax. This can be used to describe some knowledge information according to KIF, e.g. "f(origin, destination)=time + cost + weather" or knowledge of an abstract nature like "for travelling business people time is much more important than cost". This kind of information can be described formally in KIF thus it is understandable / can be processed by computers.

5 USING THE GTF-FORMAT

From the description of the requirements of systems supporting GTF follows that transportation modelling-data needs to be transferred across different platforms, mainly Windows and UNIX platforms. The structural system requirements are depicted in figure 7 (Mandel & Ruffert, 1999).

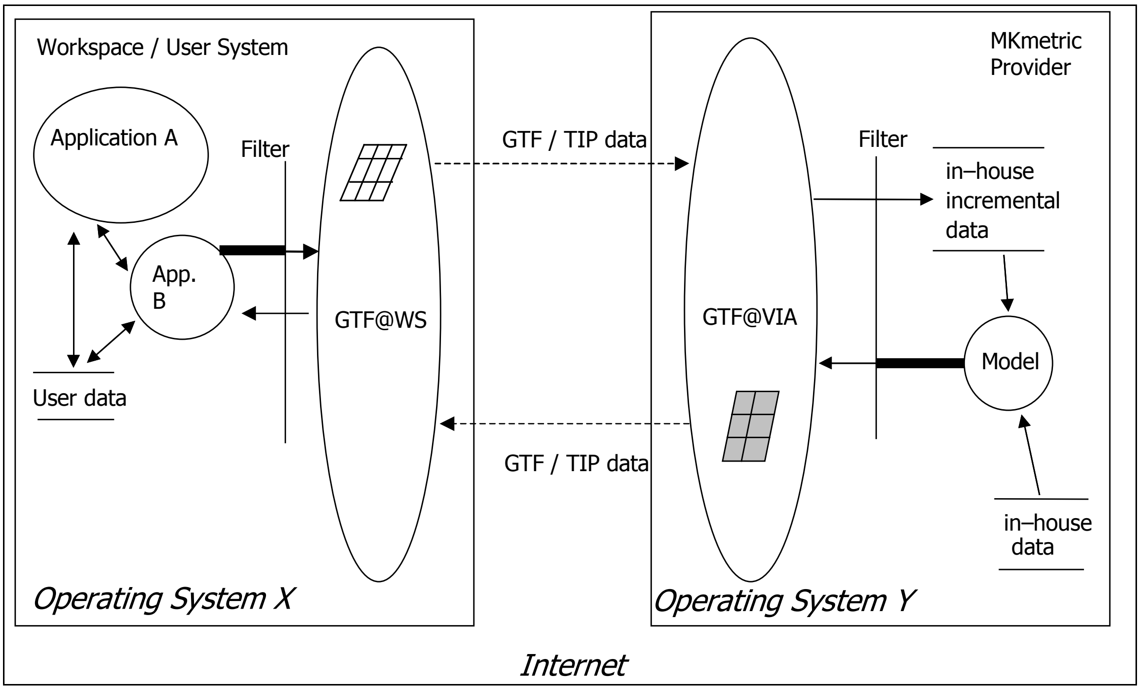

Figure 7. *Typical Exchange Structure.*

5.1 Workflow when exchanging information

As an example of a typical workflow, a user modifies his local "User data" through his system. The user then formulates a request for a transportation model, and the data to be used. A filter is used to make sure that only relevant data (-data not unknown to the transportation model provider) gets translated by the GTF@VIA Translator (e.g. VIA = MKmetric's transportation model package). The resulting GTF file is transferred to the user's account at the transportation model provider's server. There the data from the GTF file is extracted and incorporated as incremental data into the data already available at the provider's (in–house data) site. The complete data is then fed into the chosen transportation model according to the TIP information in the GTF file and the requested computations are done. The requested results are extracted by the filter and translated into GTF by the GTF@VIA Translator. The user's system gets notified that the requested results are ready for download at the provider's site. The user downloads the data. The user can then view the results with his favourite applications.

The consequences for the actual structure of a GTF file are:

- Cross-platform / Human-readability: A non–binary code must be used. The choice has fallen to the ASCII code, because this format has the least problems when being exchanged between heterogeneous platforms. ASCII also has the additional effect that a GTF file in ASCII can also be read and understood by a human, e.g. in case of problems.
- Segmented & Self-describing: As the data and control information to a

transportation model needs to be put together by the user's system the exchange format must be very flexible and powerful.

5.2 Transportation–data Interchange Protocol (TIP)

The GTF – TIP specification document includes a number of commands that can be issued to a transportation model provider's GTF enabled system. These are part of the GTF file and will enable a transportation model provider to process the GTF data file so that the requested answers are computed. This is necessary, because the GTF conceptual model alone does not contain any information on what shall be done with the data. TIP is a generalisation of "usual" commands (queries) to a transportation model to compute results. The development of TIP is based on the classic four-step transportation model. The commands can be independent of the actual transportation model or the transportation model's philosophy and can be issued to the transportation model in order to produce intermediate data or final transportation model results. These results can then be passed through a filter defined in a TIP command file that is part of a GTF-XML file. The filter extracts data from transportation model results corresponding to the user's query, and notifies the user's system that the requested results are available for download from the transportation model provider.

6 SUMMARY, DISCUSSION AND CONCLUSIONS

GTF is an acronym for "Generalised Transportation-data Format"; with the goal of standardising the information used by transportation modelling software for the purpose of electronic data interchange (EDI). The GTF specification uses already defined standards wherever possible in order to maximise acceptance and to minimise redundant work. To accomplish this the GTF specification comprises the following parts:

1. A standardised definition of transport information, but without constraining the possible information to any specific domain. This is called the *GTF conceptual model (GTF-CM)*.
2. A standardised set of commands to run transportation models and to retrieve relevant data. This is called *Transportation-data Interchange Protocol (TIP)*.
3. A standard format for arranging data in a file used for *Electronic Data Exchange (EDI)* and a standard protocol for exchanging the data file. For this XML is used. (GTF-XML)

This paper addresses primarily the main components of the GTF conceptual model (section 4). As the technical descriptions of the other two components, and details on the conceptual model are comprehensive, they are only briefly described in the paper with reference to Mandel & Ruffert (2000 & 2001).

During the discussions within the EU-project BRIDGES, followed by the thematic network SPOTLIGHT, it was realised, that formulating a fixed data model was virtually impossible at the European level, due to the large differences in conceptual models, data definitions and software solutions found in different countries, within different domains (e.g. transport sectors), and at different levels of aggregation. Realising this, it was decided to implement a flexible format, which can be read and interpreted from any software platform (given GTF-translators have been implemented).

6.1 GTF specification

The GTF specification was developed to enable transportation model providers to offer their transportation models' results in a standard fashion. Subsequently, this enables computer systems to present the results in the form a user wishes. A complete system furthermore should assist the user with the tasks of finding appropriate data and appropriate transportation model providers to answer a user's transportation query.

The specification does not cover everything in detail, but tests showed that transportation models of urban transport, freight and passenger models, special models for shipping, road specific information on load or damages, schedules as well as indicators or indexes can be handled by GTF (Mandel & Ruffert, 1999).

6.2 Technical development of GTF

During the work with GTF – and discussions with transportation model providers, transportation model users and transportation modellers – it has been realised, that the balance between flexibility and predefinitions in a format is difficult.

Without offering the possibility to add sub-classes or new parent classes, one risks that GTF cannot contain the richness of a certain transportation model, whereby it becomes useless for certain data sets.

However, if many users add their own extensions to GTF it become less general with the risk of being a set of tailor-made formats for which all other transportation modellers need to develop specific versions of their GTF-translators. The ultimate problem with this may be different GTF definitions of transportation models that in fact are conceptually equal. Hereby, GTF would de facto degenerate into several – related – exchange formats.

The solution is not easy. However, the best approach seems to:

- Extend GTF with new core-objects if several transportation models need these.
- Extend GTF with parent and sub-classes "labelled GTF-versions", when more than one transportation model need additions that describe the same conceptual phenomena.
- Extend GTF with tailor-made additions only for phenomena that is contained in one transportation model only. These additions should only be sub-classes, since other transportation models that do not use this richness can interpretation an exported data-model using the implicitly given parent classes.

A procedure of submitting "change requests" (e.g. like the "Request for

Comments” RFC commonly used in the Internet on all kinds of topics – the HTML specification was an RFC) to *http://gtf.mkm.de* or *spotlights@mkm.de* should be installed and formalised.

6.3 Future use of GTF

In our point of view, there is a widespread waste of resources within the transportation modelling community due to inconsistent data and lack of reuse of existing data. However, transportation modelling is a complicated field. And the present version of GTF became very complicated in order to capture the complexity of transport models. Even as such, it covers mainly more well known transportation model types.

On the other hand, resources for transportation modelling are often low. Furthermore, some software products trap their customers by using closed proprietary data models, and/or insufficient exchange routines. As such, the vendors have neither economic nor businesslike reasons for implementing a unifying exchange format (in this context the OpenGIS consortia by the leading GIS-vendors is a revolutionary step within the GIS-community).

There are also business-like and political reasons that hinder the exchange of transportation modelling data, e.g. between competitive rail operators, or certain regions that do not want other organisations to question transportation modelling results they use to advocate for certain subsidy from the government or EU.

Besides technical issues within GTF, these organisational and political issues have to be solved, before the visions within GTF can migrate into practice.

7 REFERENCES

Booch, G., Christerson, M., Fuchs, M. and Koistinen, J. (1999) *UML for XML Schema Mapping Specification.* http://www.rational.com/media/uml/resources/media/uml_xmlschema33.doc

Brown, David (1997) *An Introduction to Object-Oriented Analysis: Objects in Plane Language*, John Wiley & Sons

Budd, Timothy (1997), *An Introduction to Object-Oriented Programming*, Addison-Wesley

EUROSTAT (1996) *GESMES 93 – Exchange of Multidimensional Statistical Arrays and Time–series Data*, Volume 1: Guidance to Users, Volume 2: Reference Guide, EUROSTAT, Luxembourg.

Mandel B. & Ruffert E. (1999) *GTF Final Report.* MKmetric GmbH. EU-project BRIDGES.

Mandel B., Ruffert E. (2000) Generalised Transportation-data Format (GTF): Data, Model and Machine Interaction; paper presented at the *1st ITEM Workshop; Montreal/Canada; 10/2000.*

Mandel B. & Ruffert E. (2001) *GTF Specification.* MKmetric GmbH. EU-project SPOTLIGHT.

Marchal, Benoit (1998) *XML by Example*, Que; ISBN: 0789722429

Nielsen, O. A., Israelsen, T. & Nielsen, E. R. (1998) *BRIDGES TO GIS – Methodology,* Deliverable D5 & D6. BRIDGES Contract No PL96-1138. EU, DG7, 4th Framework Programme.

Nielsen, O. A., Ruffert, E., Mandel, B. *Generalised Transportation-data Format (GTF) – data, model and machine interaction,* presented at the European Transport Conference (ETC), 10-12th September 2001.

Rumbaugh J, Blaha M, Premerlani W, Eddy F & Lorensen W (1991) *Object–Oriented Modelling and Design*, Prentice Hall, New Jersey.

UML, resource (documentation of UML, XML as well as example of the use of XML) http://www.rational.com/uml/index.jtmpl

7.1 Projects

BRIDGES, "Building Bridges between Digital Transport Databases, GIS Applications and Transport Models to Develop ETIS Software Structure" (contract no. ST-96-AM-1138), on behalf of the Commission of the European Community – DG VII, 1997-1999

Spotlights(TN); "Scientific forum for making advanced transport models fully transparent to end-users, open and more integrated into policy-making"; on behalf of the Commission of the European Communities– DG-Energy and Transport; Actual Cost Contract No.: 1999-TN.10941; 2000-2003.

7.2 Acronyms and definitions

Conceptual model	The description of objects and their relationships in a model, i.e. the structure of a model – not its implementation
Data Base	The data in a specific model stored electronically
Data Format	Specific format for exchanging data
Data Model	A conceptual model, with precise definition of all objects, their data definitions as well as each data-element
EDI	Electronic Data Interchange
EC	European Commission
EU	European Union
EUROSTAT	The statistical bureau of EU
GTF	Generalised Transportation-data Format
GTF-CM	GTF-Conceptual Model
GTF-XML	GTF's EXM-based exchange format
ITS	Intelligent Transport Systems
KIF	Knowledge Interchange Format
TIP	Transportation-data Interchange Protocol
Model	The implementation of a data model in a specific software system including all needed data (and implicitly build in methods as well)
TOP	Transportation Object Platform
UML	Unified Modelling Language
XML	eXtensible Mark-up Language (Metagrammar for interorganizational communication around the Internet)