João Mendes Moreira and Jorge Freire de Sousa
INEGI – Instituto de Engenharia Mecânica e Gestão Industrial
Rua dos Bragas - 4099 Porto Codex - Portugal
Email: jmoreira@fe.up.pt and jfsousa@fe.up.pt, Tel.: +351-2-2073650, Fax: +351-2-2073659

PLANNING AND CONTROL INDICATORS FOR MASS TRANSIT COMPANIES

The Planning Indicators Board (PIB) is a top level oriented module for mass transit companies. The proposal is to collect and to show the planning and control information to support the decision making process. It uses Data Warehousing technology.

1. The GIST system

The GIST system [2,4] is a computer application for supporting operational planning in public transport companies. It was developed as a decision support system which aims to help transport companies to improve the operation of critical resources, such as vehicles, drivers and planning staff. The system is also an important software tool to support tactical and strategic management studies regarding companies operations. A consortium of 5 leading Portuguese transport companies (CARRIS, STCP, Horários do Funchal, Empresa Barraqueiro and Vimeca) and 2 R&D institutes (INEGI-FEUP and ICAT-FCUL) is responsible for the GIST system. The companies involved in this consortium operate daily about 6000 vehicles, corresponding broadly to half of the road public transport market in Portugal, including Madeira and Azores.

The GIST system was successfully installed in those companies in 1996. Nowadays, the upgrades of this application are being developed under the name of GIST98/EUROBUS. This evolution represents both improvements to the GIST present functions and extensions of its functionality.

2. The GIST98/EUROBUS system

The GIST98/EUROBUS system contains the following modules:

- Network Module, allowing the definition of the transportation network;
- Gist-Line Module, the route information module;
- Trip and Vehicle Scheduling Module, allowing the trip timetable definition and the vehicle scheduling information management and optimisation;
- Crew Scheduling Module, the crew scheduling information management and optimisation module;
- Crew Management and Rostering Module, in which is defined who will do each duty each day and in which various optimisation algorithms are applied to the rostering rules. This module is, in fact, divided in four modules;
- User Information Module, an user-oriented module which provides information to the users of public transports;
- Performance Indicators Board Module, a top level oriented module which gives performance indicators to support the decision making process.

The first 3 modules are upgrades of the present GIST modules. The Crew Scheduling Module, the Crew Management and Rostering Module, the User Information Module and the Performance Indicators Board Module belong to the EUROBUS project that has been financially supported by a public institution named ‘Agência de Inovação’. All these modules form the GIST98/EUROBUS system. Figure 1 shows the GIST98/EUROBUS system architecture model.

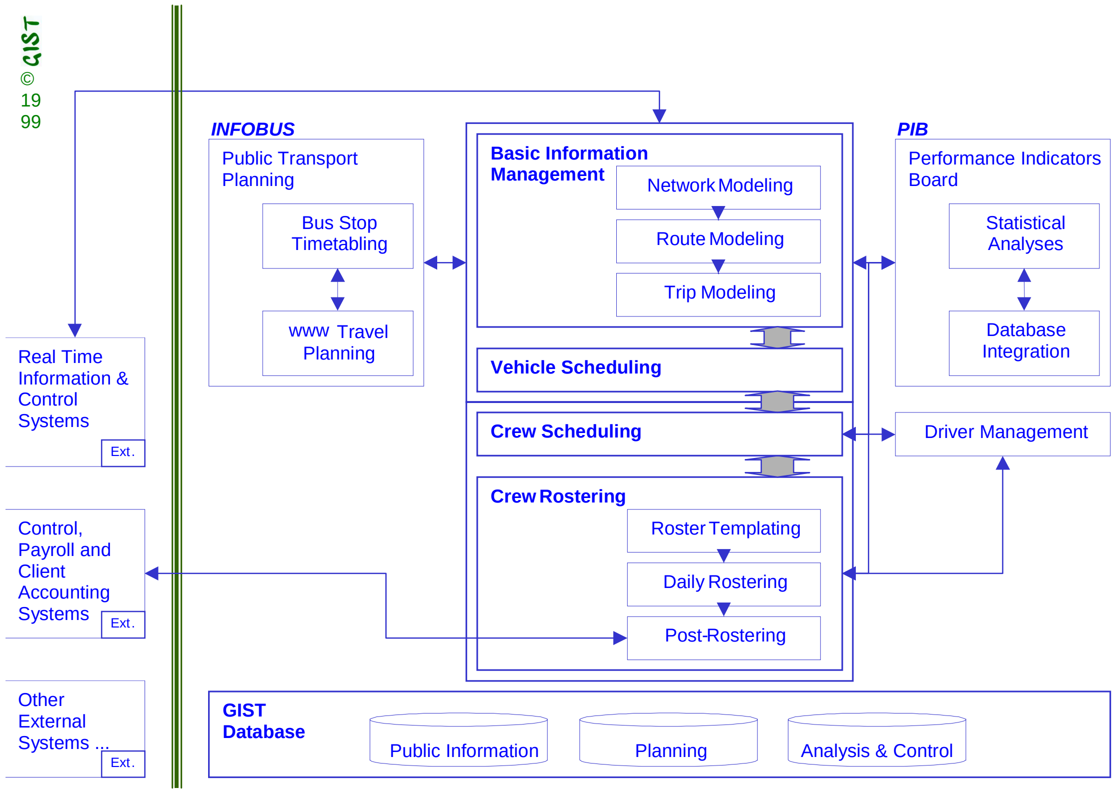

Figure 1: The GIST98/EUROBUS system architecture model

3. The Performance Indicators Board Module (PIB)

The PIB Module filters the information derived from the other modules so that the top-level managers can access, in an easy way, the relevant information for the decision making process. The main problem in structuring the PIB Module is to define those indicators that are significant to top-level managers and those that are important to the operational planners but not necessarily to the managers. The PIB Module just contains the information useful to top-level managers. All the other information is included in the corresponding module. As the GIST98/EUROBUS is a decision support system to the operational planning, the information we can get from it is essentially the planning information.

The first questions we tried to answer were: "Which are the big groups of indicators that top-level managers need?" and "Which indicators must be included in each group?".

The results presented in this text are derived from the study accomplished at the STCP Company (the Oporto Public Transport Authority).

3.1. "Which are the big groups of indicators that operational managers need?"

Taking attention to the information that GIST98 contains and to the big areas that top-level managers work with, we have defined two main groups of performance indicators: service indicators and crew indicators. The first group gives information about the level of service that the company is offering, which is especially important to the marketing department, while the crew indicators give information relevant to the human resource department.

3.2. "Which indicators must be included in each group?"

The answer to this question is not easy. The complexity involved in the definition of these indicators creates problems in a systematic approach. The way we solved this question will be explained in the next chapter.

4. Service indicators definition process

The first thing to do was to collect information from the companies. It was a long, interactive process between the analysts and the companies' staff. The service reports produced monthly by the companies were another major source of information.

Based on the information collected it was possible to notice that each potential indicator has no more than 4 parameters. This means that it is possible to identify a specific indicator by answering to 4 questions: "Which is the indicator denomination?", "Which entity does the indicator refers to?", "Which is the aggregation level used?" and "Which aggregation function does it represent (in other words, is it an average, a maximum or any other kind of function)?". As an example: if we want to know the average distance of trips per route, the indicator is the distance, the entity is the trip, the aggregation level is the route, and the aggregation function is the average. Using this methodology it was possible to identify the indicators group, the entities group and the aggregation function group. The aggregation level was more complex to define. In fact, using the example above, the trips can be aggregated by route, by line, by day type, by route and by day type at the same time, etc., i.e., they can be aggregated by one entity or by a combination of entities. The first thing we observed was that when an indicator can be aggregated by one entity it can also be aggregated by more generic entities. In the example, if routes can aggregate trips, lines can also aggregate trips (notice that a line is a set of routes). The step forward was to define the different dimensions to be used. Each dimension refers to a sequence of entities. The order by which they appear on table 1 is defined by their degree of detail. This means that the first entity of each dimension is the most generic one and the last one is the most specific. In other words, the entity referred at the first column of each dimension is a set of the entities referred at the next one, and so on.

Tables 1 and 2 present part of the results obtained by this methodological approach (the aggregation functions were ignored just for the sake of simplicity).

Table 1: Dimensions

Companies	Company			
Net Types	Net Type			
Network	Gist-Line	Line	Route	
Network Dates	Network Date			
Timetable Dates	Timetable Date			
Year Seasons	Year Season			
Day Types	Day Type			
Trip Types	Trip Type			
Vehicle	Net Type	Depot	Logic Vehicle	Stretch

Table 2: Service Indicators

	Dimensions								
Service Indicators	Companies	Net Types	Network	Network Dates	Timetable Dates	Year Seasons	Day Types	Travel Types	Vehicle
Route Extension	x	x	x	x					
Number of Trips	x		x		x	x	x	x	x
Trip Distance	x		x		x	x	x	x	x
Trip Start Hour	x		x		x	x	x	x	x
Trip End Hour	x		x		x	x	x	x	x
Trip Duration	x		x		x	x	x	x	x
Trip Amplitude	x		x		x	x	x	x	x
Trip Speed	x		x		x	x	x	x	x

With these tables we can easily see two groups of indicators: the extension indicator in the first group and all the others in the second one. This division is related with the entity the indicator refers to and the set of dimensions used. Each group will be implemented using the 'cube' concept as discussed next.

5. The implementation

To implement the module the choice fell on the SQL Server 7 with its OLAP services [1], because it has a very competitive ratio price/quality. In the implementation of a Data Warehouse there are three main steps to consider [5]:

- The Data Acquisition Processes [7]: It includes the extraction, transformation, transport and load processes. Previously it is important to analyse the source data and clean it. This is a very important issue because without reliable data it is not possible to guarantee the results. The source of information can be multiple: operational databases, other Data Warehouses/Marts or external data. In PIB's case, the source data can be one or more GIST's operational databases. This step takes typically around 80% of all the development effort.
- The Data Store Definition: This is a very important step. Notice that the information stored is directly related with the indicators and dimensions pretended, so, the database shall keep all the strategic information for business even if some of it will not be used in a near future. The SQL Server 7 stores that information in a relational database.
- Cubes definition [6]: A cube is a structure that can be seen as a huge table (fact table) where the primary key is the set of all dimensions' identifiers and the attributes are the indicators. Taking attention to tables 1 and 2, it is possible to define the cubes easily. The first thing to do is to define the dimensions as they are defined in table 1. The second step is to implement the cubes defining the indicators and selecting the related dimensions.

6. Using the analysis tool

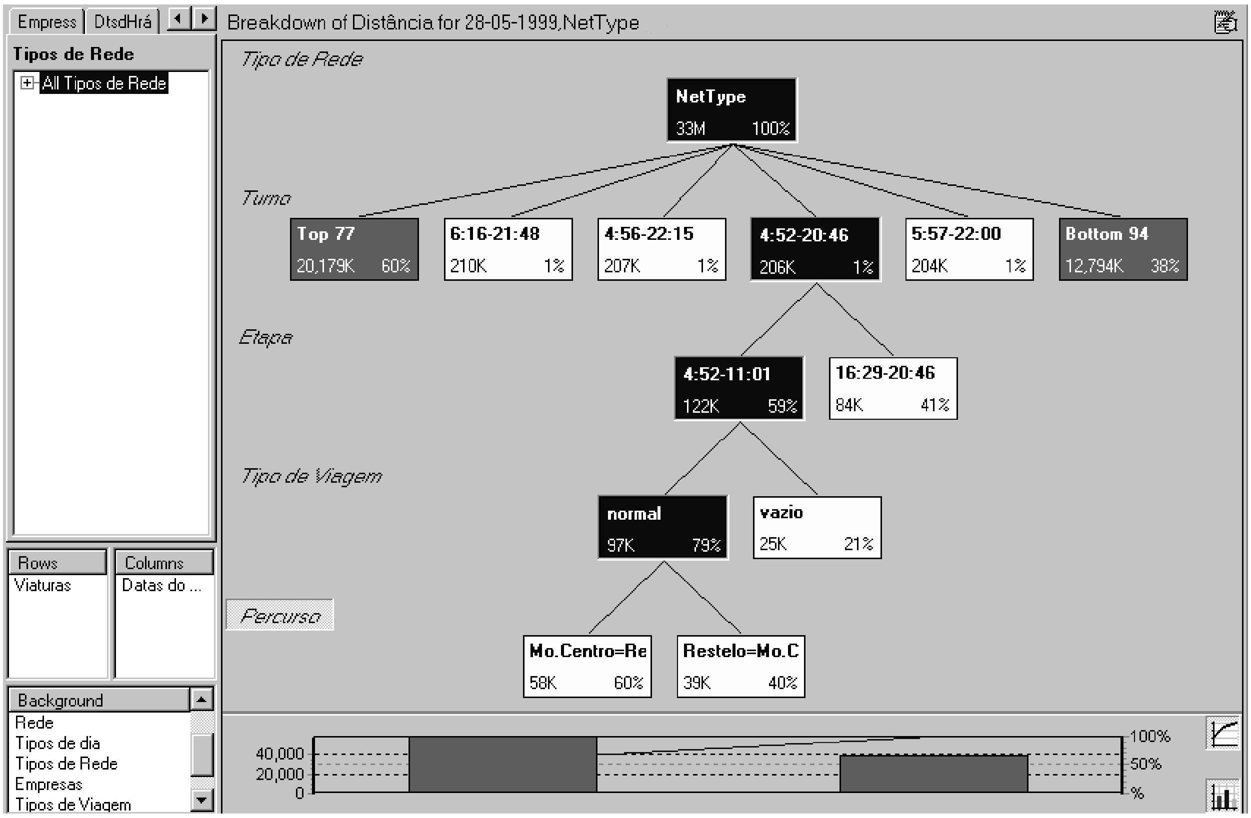

Figure2: Decomposition Tree

Finally, how can the information be useful to the decision making process? The answer is easy. There are several non-expensive analyses and reporting tools that allow the user to do all kind of analysis just executing basic commands with the mouse. In figures 2 and 3 some examples using the Knosys – ProClarity 1.0 are presented. They show two specific types of analysis. Figure 2 shows the decomposition tree with the information about the distance planned to the day 28-05-1999 with different levels of detail: by net type; by logic vehicle; by stretch; by travel's type and by route. Figure 3 shows the relationship between the planned distance and the number of trips by logic vehicle for the day 28-05-1999.

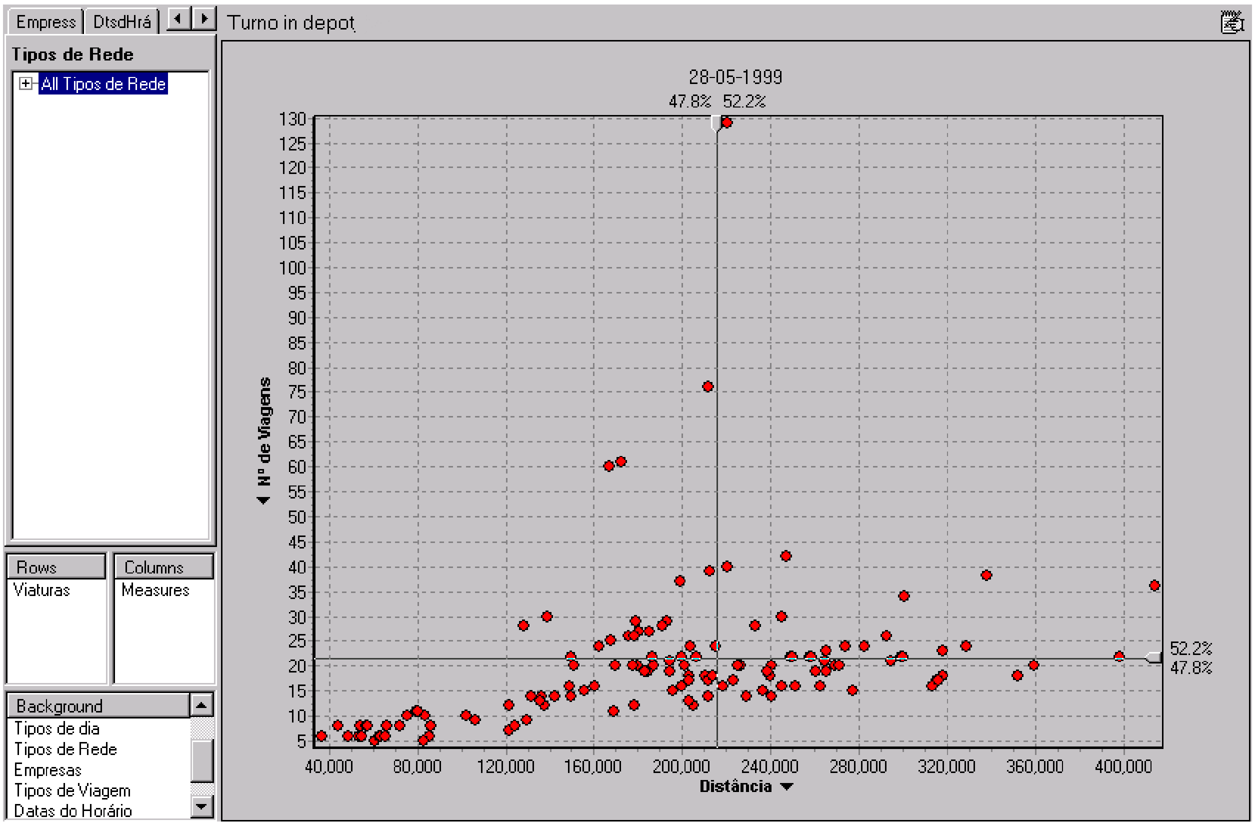

Figure 3: Perspective chart

Using this tool it is possible to do much more types of analysis such as, to compare indicators for the same month of consecutive years, or for consecutive months, to see line, column, pie, bar or other chart types, to create other indicators using formulae, etc. All this may be done using the aggregation level chosen by the user. To aggregate or detail data, it is enough to drill up or down.

7. But, why Data Warehousing technology?

There are three main reasons to build a Data Warehouse [3]:

- There is the perception that the information exists but it is not so useful as it could be
- Each department has its own language and communication is difficult.
- The reports production is expensive and inneficient

The Data Warehouse is a large or even huge database (it can reach Tera Bytes) that stores all the data strategically important to the company, along time; so the data once stored is not updated. The Data Warehouse guarantees data consistency along time and it stores the information in such a way that the data selection operations become more efficient. Once all the information is centralised, everybody is using the same indicators with the same definitions, which makes communication easier. But the main advantage of the Data Warehouse is to turn data into a competitive advantage because it can be used to analyse relationships between variables, to analyse trends, in a word, to put the information at the service of the decision making process.

References

[1]. Corey M., Abbey M., Abramson I., Barnes L., Taub B., Venkitachalam R., *SQL Server 7 Data Warehousing*, Osborne/McGraw-Hill, 1999

[2]. Cunha J.F., Sousa J.P., Galvão T., Borges, J.L., *A Decision Support System for the Operational Planning at Mass Transport Companies*, presented at the 6th International Workshop on Computer-Aided Scheduling of Public Transport, July 1993, Lisbon, Portugal.

[3]. Onder J., Nash T., *Building a business-driven Data Warehouse*, Data Management Review, October 1998

[4]. Sousa, J.P., Sousa, J.F. and Guimarães, R.C., *Un système informatique d'aide à la génération d'horaires de bus et de chauffeurs, in: Gestion de l'économie et de l'entreprise - l'approche quantitative*, Editions CORE, Série Balises, De Boeck, Brussels, 1988, pp. 477-492 (in French).

[5]. The Data Warehousing Institute, *1998 Data Warehousing Buyer's Guide*, http://www.dw-institute/buyersguide/, 1998

[6]. Thomsen E., *OLAP Solutions*, Wiley, 1997, pp. 229-264

[7]. Watterson K., *Attention, Data-Mart shoppers*, BYTE Magazine - July 1997