PERFORMANCE INDICATORS BOARD

João Mendes Moreira[1], Jorge Freire de Sousa[2]

INEGI - FEUP, Rua dos Bragas, 4099 Porto Codex, Portugal
[1]jmoreira@fe.up.pt; [2]jfsousa@fe.up.pt

Abstract:
The Performance Indicators Board (PIB) is a top level oriented module for mass transport companies. The objective of PIB is to collect and to show the planning and control information to support the decision making process. In this paper, the authors describe the indicators identification process and the logic architecture of the module, in particular its query tool.

Keywords: Key Performance Indicators, Transportation, Data Mart, OLAP.

1 The GIST system

The GIST system [4,3] 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 an improvement to the GIST present functions and an extension 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;
- 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;

- 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;
- Hiring Management Module, the module that manages the renting vehicle activity.

The first 3 modules are upgrades of the present GIST modules. The Crew Scheduling Module, the 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 together with the Hiring Management one form the GIST98/EUROBUS system.

3 The Performance Indicators Board Module (PIB)

The PIB Module filters the information derived from the other modules so that the operational managers can access, in an easy way, the relevant information for the decision making process. The main problem in structuring a PIB Module is to define those indicators which are significant to the managers and those that are important to the planners but not necessarily to the managers. The PIB Module just contains the information useful to the operational 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 planning information.

The first questions we tried to answer were: "Which are the big groups of indicators that operational 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) and also from the U.S. Bureau of Transportation Statistics [2].

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 operational 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 specially 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 those indicators creates problems in a systematic approach. The way we solved this question will be explained in the next section.

4 The process of defining the service indicators

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 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 width of trips per route, the indicator is the width, 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 period of the day, by day, by line and by period of the day at the same time, 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 2 is defined by their degree of detail. This means that the first entity of each dimension is the most specific one and the last one is the most generic. In other words, the entity referred at the last column of each dimension is a set of the entities referred at the previous one, and so on.

	Dimensions					
Service Indicators	Space	Day type	Time	Vehicle	Roster	Planned vs. Real
Number of routes	x					
Routes width	x					
Number of lines	x					
Lines width	x	x	x	x	x	x
Number of trips		x	x	x		x
Number of vehicles used	x	x	x	x		x
Number of vehicles used at the same time	x	x	x	x	x	x
Total distance	x	x	x	x	x	x
Commercial distance	x	x	x	x	x	x
Non commercial distance	x	x	x	x	x	x
Commercial time	x	x	x	x	x	x
Non commercial time	x	x	x			x

Table 1: Service Indicators

Space	Segment	Route	Line	Gist-line	Depot	Network
Day type	Day	Day type				
Time	Day	Week	Month	Quarter	Year	
Vehicle	Trip	Vehicle	Depot	Network		
Roster	Duty period	Duty	Duty type	Roster	Depot	Network

Table 2: Dimensions

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).

5 The PIB's logic architecture [5,1]

The data for PIB is stored in an appropriate database. This is an important feature for several reasons. Some companies operate in several towns and they may have a GIST98/EUROBUS system working at each town; if the managers want to know indicators concerning all the company, and not just a part of it, the PIB has to work at the same time with information from several databases in order to provide the right indicators. Moreover, the information that is stored is not necessarily the same that is used to support the operational planning. Sometimes, specific statistics based on that information are enough. So, the information is read from different databases with a given periodicity and only the one that is relevant to the PIB Module is stored.

To extract and load the data to the PIB database, a load manager has been built. The load manager's function is to extract the information from the source databases (just the relational ones considered) and load it to the PIB database. It is also a function of the load manager to perform simple transformations in order to adapt the source data to the PIB database. The extracting and loading functions can be scheduled in order to avoid the overloading of the computer during the working period.

As the companies often just need to keep data during a certain number of years, it is important to have a manner of storing the data older than a given date in a proper device. That data must be available whenever the users need it.

The two main components of the PIB module perceptible to the users are the query tool and the report tool. The query tool allows the definition of the indicators. It is a user-friendly interface allowing the construction of complex queries by non-skilled users. This tool will be described in detail in section 6. The report tool is oriented to the construction of reports. It is possible to include a set of indicators in a report and to format it. There is a set of different layouts for the indicators available to the user. It is possible to define parameters to the reports. The report is a Microsoft Excel file that can be opened and changed using this software.

The main problem of this type of systems is the large amount of information they have to manage. The relational database management systems were traditionally designed to optimise the use of device's space. Roughly we can say that the paradigm was: "The less is the space needed to store the data, the better is the database design". During the last years the information won an increasing strategic importance in the companies and the storage devices became cheaper. This has driven to a change of paradigm. Nowadays the ability to manage huge amounts of data with good performances is a crucial factor for the companies. So, in the design of a database, the performance factor must be studied. How should we design the database to speed up queries? One of the crucial factors to reduce the query's response time is the number of tables used. The ideal to speed up queries is to design the tables according to the queries to be done. Usually, a compromise option between these two paradigms is used. The information is stored in the third normal form and, according to the most frequently queries, temporary tables are created to improve performances. As the most frequently asked questions are not the same along the time, the management of temporary table is dynamic. This is the job of the warehouse manager.

The warehouse manager updates the structure of the database but, to optimise the performance of the queries, we need to read the information from the database in the most efficient way. This is the function of the query manager.

Figure 1 shows how those six components interact with each other.

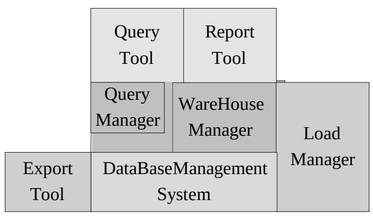

Figure 1

6 The query tool

In section 4 we saw that it is possible to identify a specific indicator by answering to 4 questions. The query tool uses that kind of questions to create a sequence of user-friendly windows.

6.1 Entity

Firstly, the entity is chosen. All the other windows depend on this choice. The entity choice can be seen as the answer to the question: "Which entity does the indicators refers to?" or "From which entity do I want to get information?".

6.2 Indicators

The second step is the selection of the indicators. The indicators available are those that refer to the entity chosen previously. These choices can be seen as the answer to the question: "Which is the indicator denomination?". In this window, the user can also define the aggregation function that he wants to use with each indicator. By doing this, he is answering to the question: "Which aggregation function does the indicator represent?".

6.3 Aggregation levels

Then, the user must choose the aggregation levels he wants, i.e., the dimensions. It is the answer to the question: "Which are the aggregation levels used?". The dimensions available depend on the chosen entity. If there are several identifiers for each dimension, the user must select one of them. When the user finishes this step he has identified the required indicator(s).

6.4 Layout

The layout is defined in a matrix format. The user can set the rows and the columns.

6.5 Constraints

By default, all the records are displayed. To impose constraints the user can write conditions or/and select exhaustively the required data. The window where this is done can be reached by pressing a single button. This is done to hide from the users optional windows that are used only by skilled users.

6.6 SQL statement

Finally, the user can access the SQL statement and change it, if he wants. In the same window, he gives the name to the indicator created so that he can use it later.

With this type of interface it is possible to generate any SQL statement. In order to accomplish this task it has been necessary to add some options to the windows described above such as: to ignore similar rows (DISTINCT clause), to order the data by a given criteria (ORDER BY clause), to operate set operations like union, intersect and minus, or to use sub-queries. Skilled users can edit the SQL statement and write it directly using that language. Any basic SQL statement using the optional clauses WHERE, GROUP BY and HAVING can also be done with this interface. These clauses form a large percentage of the queries most frequently done by the users.

7 Conclusions

In this paper we described a computer application for supporting operational planning in public transport companies. The main feature of the PIB Module is to display the information the system holds in a useful layout to the managers, hence reducing the documentation production costs. PIB is not a competitor of OLAP (On Line Analytical Processing) tools. The decision of developing our own OLAP tool was taken by the consortium based on a cost/benefit analysis.

The methodology used in the Performance Indicators Board Module was user-oriented. During the development process, analists and developers had to interact with the final user in order to adapt the module design to the user needs. As a consequence, changes were successively done to previous plans. This interactive process is essential by several reasons: the involvement of the companies in the development is strong, the redesign effort is minimum, the definition of short-term goals speeds up the development, and companies become more capable to understand technical difficulties. All these features increase the probability of success.

In November 1998 Microsoft launched the SQL Server 7 that includes the OLAP Server, all at a price around $1400 for 5 users. With this announcement our strategy was altered because the cost of this software is clearly lower than the PIB's development cost and because this tool is better to achieve the purpose of our work.

References

[Anahory, Sam and Murray, Dennis, 1997], Data Warehousing in the Real World, Addison-Wesley

[Bureau of Transportation Statistics (BTS) of the U.S. Department of Transportation, 1997], National Transportation Statistics 1997.

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

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

[Thomsen, E., 1997], OLAP Solutions: Building Multidimensional Information Systems, John Wiley & Sons, Inc.