A BUSINESS INTELLIGENCE SOLUTION FOR PUBLIC TRANSPORTATION SECTOR

Paulo Sampaio[1]
Guilherme Pereira[1]
Maria Sameiro Carvalho[1]
José Telhada[1]
António Paisana[1]
Paulo Paixão[2]
Albertino Fonseca[1]
[1]University of Minho, Systems and Production Department, Campus Gualtar, 4710-057 Braga, Portugal
E-mail: {paulosampaio, gui, sameiro, telhada, apaisana}@dps.uminho.pt; tino.fonseca@gmail.com
[2]EFACEC Sistemas de Electrónica, Rua Eng. Frederico Ulrich – Guardeiras, Apartado 3081, 4471-907 Moreira da Maia
E-mail: paulo.paixao@efacec.com

KEYWORDS
Business intelligence, performance measurement, intelligent transportation system, decision support system.

ABSTRACT

Smart companies in the 21st century use business intelligence solutions to gain a clearer picture of their internal operations, customers, supply chain and financial performance. In short, they are using business intelligence solutions to become intelligent about the way they do business. Our paper reflects the result of a project that we conducted in Portugal related to the development of a business intelligence solution for the intelligent transportation system (ITS). This project was developed on a close partnership between the University of Minho and Efacec, that is one of the most important and large Portuguese companies. More specifically, our aim was to: (1) perform a strict identification and selection of the information requirements; (2) develop technical specifications; (3) develop a business intelligence solution that will be used to support the companies top management decisions. Our approach consists on integrating in the computer-aided support system all the relevant information by mixing up different visualization forms (tables, graphs and maps), customized according to each performance indicator, thus enhancing reading efficiency and correct interpretation of every depicted situation being analyzed.

INTRODUCTION

This paper is focused on providing a systematization of the performance measures that are relevant to the transport system of EFACEC (public transport for passengers) – section 2, on proposing a data base model to support the information management system – section 3, and on proposing an adequate BI solution, creating some forms of visualization to enhance the effective and efficient use of the data and information available – section 4.

The company providing BI solutions for the transportation sector

With a history of already 100 years, the Efacec Group had its origin in "A Moderna", a company established in 1905.

Formed in 1948, Efacec, the largest Portuguese Group in the field of electricity, employs around 4500 people and has a volume of orders that already exceeded 1000 MEuros, being present in more than 65 countries.

The portfolio of Efacec business activities is composed by:

- Energy Solutions:
  - Transformers.
  - High and Medium Voltage Switchgear.
  - Service.
- Engineering solutions and Services:
  - Engineering.
  - Automation.
  - Maintenance.
  - Environment.
  - Renewable.
- Transport and Logistics:
  - Transport.
  - Logistics.

The Efacec portfolio is based on a Systems Integrating Contractor philosophy that covers the present needs of the market and turns the different Group capabilities profitable.

Efacec's strategy, focusing on International market, as well as important investments made in Innovation and New Technologies Development, backed by the resources of Efacec traditional technology led the Group to a position at the forefront of the Portuguese Industry and overseas markets.

The business intelligence concept

Business Intelligence (BI) solutions create learning organizations by enabling companies to follow a virtuous cycle of collecting and analyzing information, devising and acting on plans, and reviewing and refining the results. To support this cycle and gain the insights that BI delivers, organizations need to implement a BI system comprised of data warehousing and analytical environments.

According to Eckerson (2003), smart companies recognize that the systems that support BI solutions are very different from other systems in the company. Well-designed BI systems are adaptive by nature and they continually change to answer new and different business questions. The best way to adapt effectively is to start small and grow organically. Each new increment refines and extends the solution, adjusting to user feedback and new requirements. Additionally, the BI solutions combine data with analytical tools in order to provide information that will support top management strategic decisions (Cody and Kreulen 2002).

The main goal of a BI solution is to collect data, store the data collected and analyze the data gathered in order to produce knowledge (Santos and Ramos 2006).

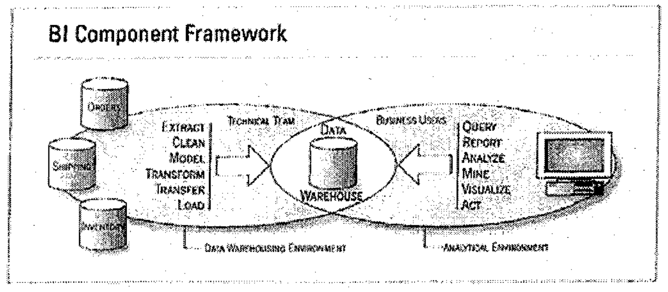

Figure 1: BI environment (Eckerson, 2003)

It is important that organizations understand the key indicators of success so they can surmount the challenges associated with every BI project. The successful BI solutions should have the following characteristics (Eckerson 2003):

1. Business sponsors highly committed and actively involved in the project.
2. Business users and the BI technical team working together closely.
3. The BI system viewed as an enterprise resource and given adequate funding and guidance to ensure long-term growth and viability.
4. Firms should provide users both static and interactive online views of data.
5. The BI team should have prior experience with BI.
6. All the company should be committed and involved with the BI solution.

Organizations that have deployed BI solutions cite many tangible and intangible benefits. Although it's difficult to associate a concrete Return on Investment (ROI) resulting from these benefits, most enlightened executives place huge value on having a "single version of the truth", better information for strategic and tactical decision making and more efficient processes (Eckerson 2003). The most common benefits cited in the literature are the followings:

1. Time savings.
2. "Single version of truth".
3. Better strategic and plans.
4. Better tactics and decisions.
5. More efficient processes.
6. Cost savings.
7. Greater customer and supplier satisfaction.
8. Greater employee satisfaction.
9. Return on Investment.
10. New revenues.
11. Total cost of ownership.
12. Shareholder value.

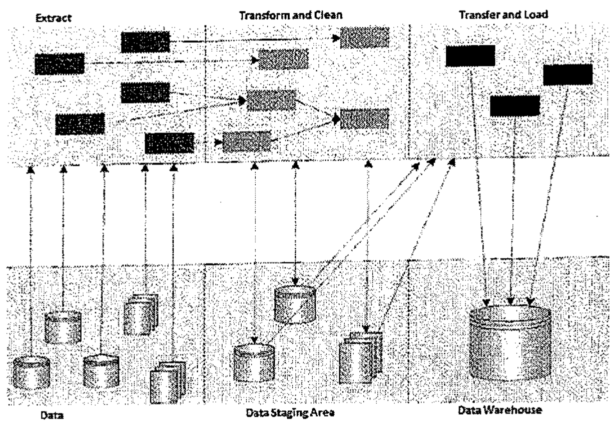

Figure 2: Data Extraction, Transformation and Cleaning, and Transfer and Load (Santos and Ramos, 2006)

PERFORMANCE MEASUREMENT FOR TRANSPORTATION SECTOR

Performance measurement is a way of monitoring progress towards a result or goal (Sousa *et al.*, 2005, Juran and Godfrey, 1998). It is also a process of gathering information to make well-informed decisions. Performance measures are valuable and provide several useful benefits (NCHRP 2006):

1. Greater accountability to policy-makers, organization customers and other stakeholders.
2. Improved communication of information about the transportation system to customers, political leaders, the public and other stakeholders.
3. Increased organizational efficiency in keeping agency staff focused on priorities and enabling managers to make decisions and adjustments in programs with greater confidence that their actions will have the desired effect.
4. Greater effectiveness in achieving meaningful objectives that have been identified through long-range planning and policy formulation.
5. A better understanding of the impacts of alternative courses of action that performance measures can provide.
6. Ongoing improvement of business processes and associated information through feedback.

Performance measures traditionally have been largely technical in nature. However, today transportation executives and managers must address an increasingly complicated and wide-ranging set of issues regarding the best solutions on balance to transportation problems, the cost-effectiveness of proposed projects and anticipated impacts of those projects.

Based on the literature related to performance measurement systems for transportation, there is a large number of measures within the following categories:

1. Preservation of assets.
2. Mobility and accessibility.
3. Operations and maintenance.
4. Safety.

The Transportation Research Board – TRB (NCHRP 2006) had developed a framework that transportation organizations should use to:

1. Identify performance measures that are most useful for asset management: assessing existing performance measures that are in place, identifying gaps, and considering new measures to fill gaps.
2. Integrate these performance measures effectively within the organization: engaging stakeholders to ensure buy-in, designing families of measures that can be used at different organizational levels and for different types of decisions, ensuring consistency across measures, identifying needed improvements to data collection and analytic tools, designing communication devices, and documenting measurement and reporting procedures.
3. Set performance targets: establishing both long-term goals and short-to medium-term targets for performance measures.

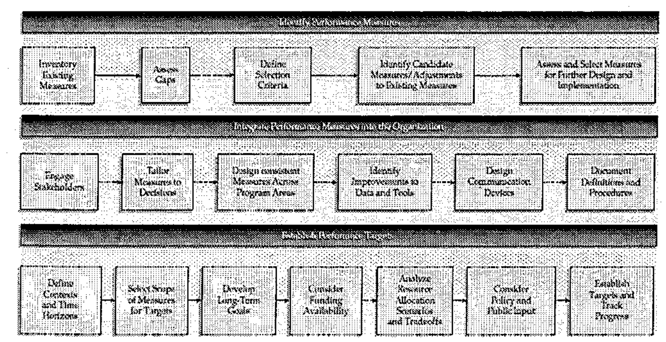
Figure 3: Guidance for Performance Measures and Targets (NCHRP 2006)

Rather than a single, prescriptive set of measures that might provide a good textbook example but surely would not be suitable for all companies, the framework is in the form of guidance on the preferred types and forms of measures and the process by which companies should identify and implement them. The framework was developed with the recognition that each company will have a different set of circumstances and needs to consider in selecting and implementing performance measures (NCHRP 2006).

One could not define the best set of performance measures for transportation companies, because each company has their own characteristics. However, the Transportation Research Board suggests a set of "good measures" that if implemented properly, will produce good results. The performance measures categories that have been defined by TRB are the following:

- Preservation – measures the condition of the transportation system and actions to keep the system in a state of good repair.
- Mobility and Accessibility – measures the ease of movement of people and goods.
- Operations and Maintenance – measures the effectiveness of the transportation system in terms of throughput and travel costs and revenues from a system perspective and maintenance level of service measures focused on the customer experience of the system.
- Safety – measures the quality of transportation service in terms of crashes or incidents that are harmful to people and damaging for freight, vehicles and transportation infrastructure.
- Economic development – measures direct and indirect impacts of transportation on the economy.
- Environmental impacts – measures effects on environment.
- Social impacts – measures effects on broader society or on population groups.
- Security – measures protection of travelers, freight, vehicles and system infrastructure from criminal and terrorist actions.
- Service Level – measures the delivery of transportation projects and services to customer.

The absence of measurement limits organizations ability to evaluate the changes and therefore precludes systematic improvement. Thus, good performance measures are a necessity for any progressive organization to recognize successful strategies and discard the unsuccessful. The performance measures categories defined for the model developed were supported in the TRB framework.

DATA BASE CONCEPTUAL MODEL

As it was already stated, this project aim was to develop a business intelligence solution for passenger transportation operators companies from the public sector.

Based on the literature review carried out and on the indicators identified, our proposal was to create a model based in the set of performance indicators that have been developed by the Transportation Research Board, integrated with the performance indicators defined and used by Efacec.

The model developed is supported in Data Base Model philosophy with the aim of organizing and managing the company information.

The most common model frameworks developed for business intelligence solutions could be based on a pure star schema design (one fact table connected with multiple dimension tables) that could evolve to a star schema design with multiple fact tables (one for each group of performance indicators) connected with multiple dimension tables, that, additionally, could evolve to a star scheme design with multiple fact tables sharing different dimension tables, usually known as constellation scheme design (Moody and Kortink, 2003).

The adopted scheme will have a significant impact in the rationalization of data storage space, concerning the information access and in terms of hardware necessities.

The scheme that best fit to the project goals was the constellation scheme design.

One important limitation of the constellation model is related to its dimension. It is important to point out that, regardless being the best solution, this öption will require a higher computation performance that will be reflected in terms of higher software and hardware costs.

The constellation model proposal is based on one Data Warehouse that is composed by three Data Marts – Courses, Events and Incidents, with multiple dimension tables, some of them shared between more than one fact tables (Date and Time). These Data Marts intend to register historical information on facts occurring during delivering of transportation services (Figure 4).

Courses Data Mart is composed by 10 dimensions tables:

- Date
- Time
- Route
- Alternative Route
- Path
- Link
- Stop
- Schedule
- Vehicle
- Driver

Concerning Events Data Mart, there are 5 dimensions tables defined:

- Date
- Time
- System
- Equipment
- Failure type

The Incidents Data Mart is composed by 7 dimensions tables:

- Date
- Time
- Place
- Operator
- Tender
- Road
- Incident type

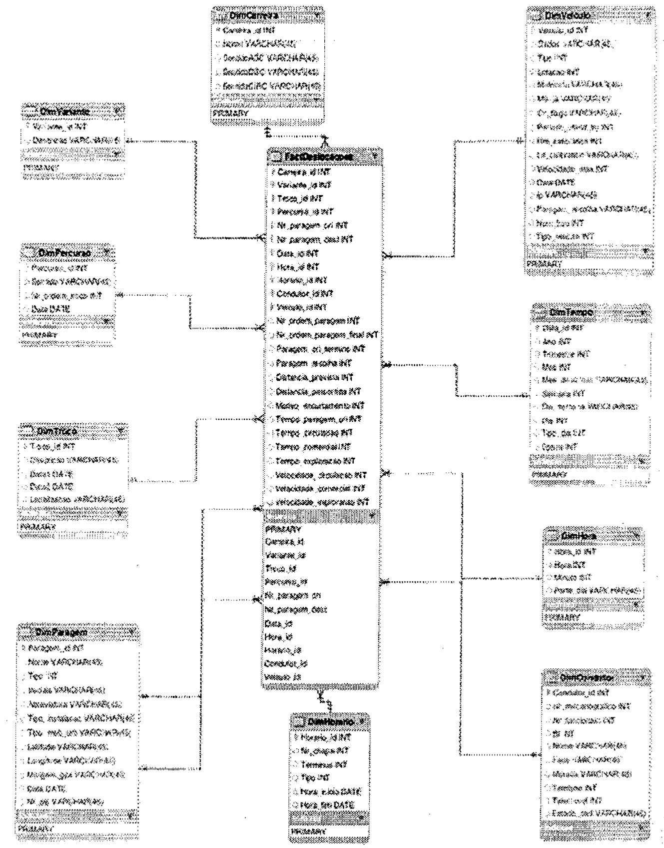

Figure 4: Conceptual Model developed for Courses Data Mart

BUSSINESS INTELLIGENCE SOLUTION

Intelligent transportation Systems (ITS) integrate a broad range of IS/IT elements for gathering on-site data and transfer it to the central database(s) of the enterprises or organizations. ITS also integrate the required hardware and software to analyze and manage the information extracted from data.

In general, the amount of data available (e.g. detailed traffic operations events and conditions) is huge, and only a relatively small fraction of that is adequately translated into information that is used by managers. In most cases, this is restricted to the information that is easily obtained by applying traditional query database operations and graphical displays, including standard statistical elements (e.g. historical means and deviations of traffic volumes) in general-purposed tables, graphs and maps. But, in fact, much more information can be potentially gathered from data: (1) further performance measures can still be obtained by traditional forms; (2) deeper information somehow hidden in the databases that can be highlighted by applying more advanced analytical methods, such as data mining, optimization and simulation, as well as advanced forecasting (e.g. auto-correlative, multi-regressive) and statistical techniques; and, very important, (3) most of all this information can be better highlighted by using alternative, more adequate forms of visualization such as well-designed graphs and maps; also, some information (e.g. spatial-temporal trends) can be exclusively evidenced by specific forms of visualization. It is widely recognized that it is (often) easier to detect a pattern from a "picture" than from a textual or numeric explanation.

From the above, it is obvious that one of the most important problems of ITS is the analysis and management of information. This problem is becoming further relevant due to the permanent increase on the availability of data and their related inexpensive storage and processing power as a result of the widely implementation of modern IS/IT systems.

The BI solution developed to Efacec is based on commercial available software, and most of the performance measures implemented were easily translated into the form of traditional tables, graphs and reports. This includes making use of simple dashboards. Most of these elements can be easily customized by end users, according to their specific needs and options.

Data visualization can be seen as the process of translating information into a visual form (usually graphs and maps), enabling users (scientists, engineers and managers) to perceive the features of the data available, and, from that point, to help them to make more effective decisions at different levels: strategic, tactical or operational. For example, deciding on the design of service operations strategies or, simply deciding on the realization of a functional regulatory (e.g. corrective) measure to ensure the normal functioning of the system or to minimize the impact of an anomalous event (incident) that has occurred.

In the following paragraphs three different examples show the importance of adequate visualization patterns according to specific nature of data recorded and information needed.

In order to detect peak periods and understand correlation to service delays may be worth using simple line charts of average traffic volumes (widely available from public data sources) and actual delays observed in the system, as a function of day-time (Figure 5). This should be done by taking into account the average figures for relevant zones or intersections that affects each PT service. Such information (patterns) can be also incorporated into adequate predictive models of time arrivals at stops (e.g. for public panel information).

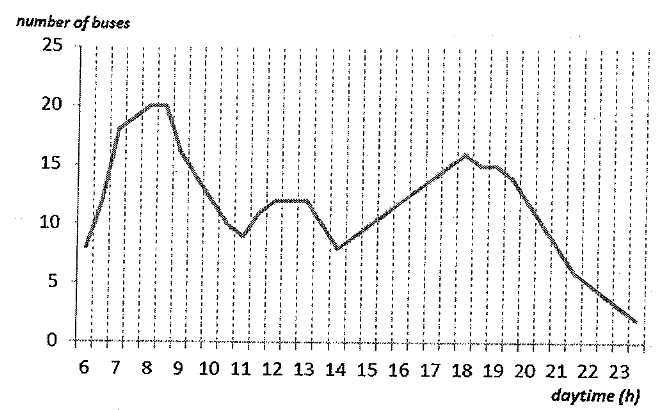

Figure: 5 Line chart of the average volume at a given location (e.g. PT stop or servive zone).

In order to perceive and monitor service levels in different directions from the city center (e.g. incoming PT services bringing people to work at the morning peak), may be worth using a rose-like diagram (Figure 6). This can be done per service or group of services in each direction (e.g., north, west, south and east bands).

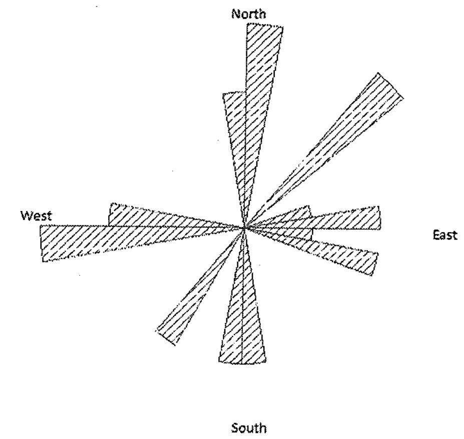

Figure 6: Rose-like graph of frequencies for 12 outbound PT services at the evening peak: each route is represented by a triangle illustrating its direction and relative frequency.

In order to infer the patterns of transit volumes in each direction per time of day (in average), may be worth using clustering analysis, e.g. putting in evidence similar directional service patterns among different time intervals. In these cases, the use of colored "graph-image" may be useful (Figure 7).

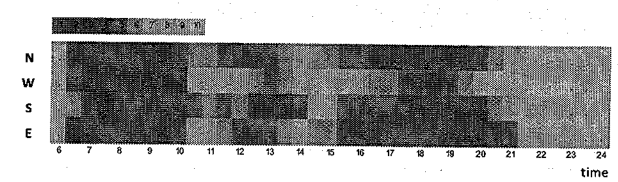

Figure 7: Average inbound and outbound public transit volume by direction and time of day (30 min clusters).

CONCLUSIONS

The Efacec BI solution developed produce performance reports for the transportation operators companies. The reports have two different levels. There are strategic reports that are only used by the companies' top management and there are operational reports to be used by the medium level management.

In this paper we do present the literature review carried out to this project and the model developed to support the Efacec business intelligence solution. The model adopted is supported in the framework developed by the Transportation Research Board for transportation companies, which define nine performance indicators categories.

One of the most important problems of ITS is the analysis and management of information. The BI solution presented, based on commercial available software, implemented relevant performance measures, translating into the form of traditional tables, graphs and reports, and allowing easy customization by end users. Thus, this work promoted the use of general proposed BI tools that, along with the integration of advanced analytical techniques, support the inference of both strategic and operational performance measures helping decision makers in the transportation sector – specific emphasis was devoted to visualization techniques that allow the detection of spatial-temporal patterns of behavior.

It is obvious that PT services must be monitored by means of dynamic (digital and vectorial) maps inter-operating with real-time data gathering devices (e.g. in-vehicle equipment including GPS). This sub-system of the ITS is commonly recognized as being a geographic information system (GIS). Most spatial-temporal patterns will be only detected by displaying the related information on the map, so this sub-system constitutes a must-be in a BI for transportation systems. Further work consists on establishing the integration of the BI application with a GIS model and software in order to enhance the usefulness of the overall decision support tool.

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by the European Commission and the Portuguese Government through the research grant QREN, in compliance with SmartCenter.

REFERENCES

Cody, W F, Kreulen, J T, 2002. The integration of business intellingence and knowledge management. *IBM Systems Journal* 41:697-713.

Eckerson, W, 2003. *Smart Companies in the 21st Century: The Secrets of Creating Successful Business Intelligence Solutions*. USA. The Data Warehousing Institute.

Juran, J and Godfrey, A B, 1998. *Juran's Quality Handbook*. USA. McGraw-Hill.

Moody, D L and Kortink M A R, 2003. From ER Models to Dimensional Models: Bridging the Gap between OLTP and OLAP Design, Part I. *Business Intelligence Journal* 8: 7-24.

NCHRP, 2006. *Performance Measures and Targets for Transportation Asset Management*. USA. Transportation Research Board.

Santos, M Y, Ramos, I, 2006. *Business Intelligence: Tecnologias da Informação na Gestão do Conhecimento*. Portugal. FCA-Editora de Informática.

Sousa, S, Aspinwall, E, Sampaio, P, Rodrigues, A G, 2005. Performance Measures and Quality Tools in Portuguese Small Medium Enterprises: Survey Results. *Total Quality Management and Business Excellence* 16: 277–307.

AUTHORS BIOGRAPHY

PAULO SAMPAIO is graduated in Industrial Engineering and Management (University of Minho – Portugal, 2002) and holds a PhD degree in Systems and Production Engineering (University of Minho – Portugal, 2008). His main research interests are Quality Management, Operational Research and Applied Statistics.

GUILHERME PEREIRA was born in 1961 in Porto, Portugal. He graduated in Industrial Engineering and Management in the University of Minho, Portugal. He holds an MSc degree in Operational Research and a PhD degree in Manufacturing and Mechanical Engineering from the University of Birmingham, UK. His main research interests are Operational Research and Simulation.

MARIA SAMEIRO CARVALHO is graduated in Computer and Systems Engineering (University of Minho – Portugal, 1986) and holds a MSC degree in Transportation Planning and Engineering (University of Leeds - UK, 1990) and a PhD degree in Transport Planning and Engineering (University of Leeds - UK, 1996). Her main research interests are in Transportation, Supply Chain Management and Operational Research.

JOSÉ TELHADA holds a PhD degree in Transport Planning and Engineering (ITS-University of Leeds – UK, 2001). His main research interests are Operational Research and Logistics applied to transportation systems.

ANTÓNIO PAISANA is graduated in Economics (University of Birmingham, 1978) and holds a PhD degree in Production Engineering from the Loughborough University, UK (1992). His main research interest is Strategic Management.

PAULO PAIXÃO is graduated in Electrical Engineer and Computers at Oporto University. He has been responsible for several National and European R&D projects. At present, Paulo is R&D Director of EFACEC Transports - Telecommunications and Signalling Division and Project Manager of SmartCenter activities.

ALBERTINO FONSECA is graduated in Informatics Engineering (University do Minho, 2010). At present he works at EFACEC.