VISUALIZATION – A KEY CONCEPT FOR MULTIDIMENSIONAL PERFORMANCE MODELING IN SOFTWARE ENGINEERING MANAGEMENT

Vasile Stroian
École de Technologie Supérieure
vstroian@yahoo.com

Pierre Bourque
École de Technologie Supérieure
pierre.bourque@etsmtl.ca

Alain Abran
École de Technologie Supérieure
alain.abran@etsmtl.ca

***Abstract* – A major difficulty with current organizational performance models in software engineering management is to represent many possible viewpoints quantitatively and in a consolidated manner, while at the same time keeping track of the values of the individual dimensions of performance. The models currently proposed do not meet the analytical requirements of software engineering management when various viewpoints must be taken into account concurrently. This difficulty is compounded by the fact that the underlying quantitative data is of high dimensionality and therefore the usual two and three dimensional approaches to visualization are generally not sufficient for representing such models.**

This paper describes the proposed concepts for a tool for multidimensional performance modeling in software engineering management. Due to the continuously increasing amount and the high dimensionality of the data underlying these models, a particular focus is given in this paper on potential visualization concepts and techniques that could be incorporated into the proposed tool.

***Keywords:* Performance modeling, Visualization, Multidimensional management models, Software engineering management, International Software Benchmarking Standards Group, ISBSG, QEST**

I. INTRODUCTION

A major difficulty with current performance models in software engineering management is to represent many possible viewpoints quantitatively and in a consolidated manner, while at the same time keeping track of the values of the individual dimensions of performance [1]. This difficulty is compounded by the fact that the underlying quantitative data is of high dimensionality and therefore the usual two and three dimensional approaches to visualization are not sufficient for representing such models. "Performance" in this paper means the performance of a software engineering organization, rather than that of software or hardware.

There already exist a significant number of one-dimensional models of performance which integrate individual measurements into a single performance index. These models can be found in software engineering management [2], as well as in other disciplines. However, these models do not meet the analytical requirements of software engineering management when various viewpoints must be taken into account concurrently.

Section II of this paper presents a selection of multidimensional models of performance in software engineering and in management. A prototype tool supporting one of these models is also reviewed.

The data underlying such models is of high dimensionality since the performance models discussed in this paper are organizational rather than project or departmental and because the data can be historical and at various levels of granularity. Section III therefore presents a review of visualization approaches and techniques for multidimensional data.

Based on an analysis, presented in Section IV, of the selection of multidimensional models and the review of visualization techniques for multidimensional data, Section V describes the proposed concepts for a new tool for multidimensional performance modeling in software engineering management. A brief summary concludes the paper in Section VI.

II. A SELECTION OF MULTIDIMENSIONAL PERFORMANCE MODELS

This section presents a selection of multidimensional performance models found in the literature. The first subsection discusses multidimensional models of software quality and of performance found in the software engineering literature in particular; a review of a prototype tool is also presented. The next subsection presents management-oriented multidimensional performance models. The subset of multidimensional performance models selected for this section was chosen because its models were deemed to be representative of the wider set of models found in the literature and because of the possibility of applying elements of the selected models in the tool proposed in this paper.

A. MULTIDIMENSIONAL MODELS OF SOFTWARE PERFORMANCE AND SOFTWARE QUALITY

The ISO 9126 [3] model of software quality is generic and can be applied to any software product by tailoring it to a specific purpose. The high-level internal and external quality characteristics, are assumed to be independent of one another.

The QEST (Quality factor + Economic, Social and Technical dimensions) model is capable of handling independent sets of dimensions without predefined ratios and weights. The three dimensions taken into consideration, as shown in Figure 1, are combined through the use of a regular geometrical representation of a pyramid (tetrahedron), the sides of which represent the normalized values of each of the software engineering project dimensions being measured. The apex of the tetrahedron represents the performance target. With this 3D representation, it is possible to determine and represent performance considering the usual and distinct geometrical concepts of distance, area and volume. Several papers cover different aspects of the QEST Model : the theoretical aspects [4], the geometrical and statistical foundations of the model [5] and the implementation of the model [6]. An extension of the QEST model to *n* possible dimensions [7], called QEST *n*D, targets complex software projects when a greater number of dimensions must be taken into account. The overall project performance (*p*), as shown in Figure 1, is determined using classic geometrical formulae, such as the volume of a truncated tetrahedron defined by the individual perspective values (*Qe*, *Qs* and *Qt*) divided by the total volume of the tetrahedron. The LIME (LIfe cycle MEasurement) model [8] extends the QEST model concepts to a dynamic context, and can be applicable to each step of any topology of software life cycle model. The QEST model also produces a unique performance indicator.

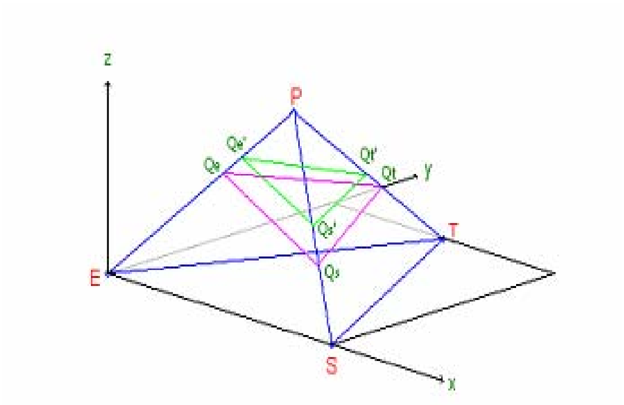

Figure 1 QEST model

A prototype tool based on QEST, described in [9], for software product quality measurement enables the user to enter three values (E,S,T) for every node of the ISO9126 quality_tree, one for each dimension as shown in Figure 2. Definitions are provided in the prototype tool as well for every ISO 9126 tree element as shown in Figure 3. For any project, its value of each dimension is given by the weighted sum of a list of *n* normalized measures having been selected as representative of each of the three viewpoints. The sides of the tetrahedron must be equal. The three dimensions are: *Economic* -represented by the managers' perspective, *Social* - represented by the users' perspective and *Technical* - represented by the developers' perspective.

However the prototype tool requires that each data entry screen be filled out in sequence, even if the individual entering the data is responsible for only a subset of it in a decentralized environment. The prototype tool is also limited to the basic model of QEST, not QEST *n*d or LIME, and to the series of measures defined in ISO912. There is also no database included with the prototype. One must therefore reenter the required data at every usage and cannot compare current results with historical data or future predicted results. No visualization techniques are available in the prototype tool for analyzing data and results other than the pyramidal representation defined in QEST.

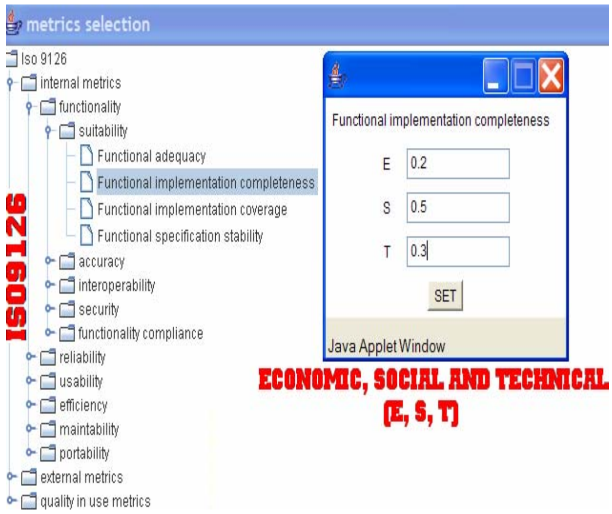

Figure 2 QEST prototype with ISO9126

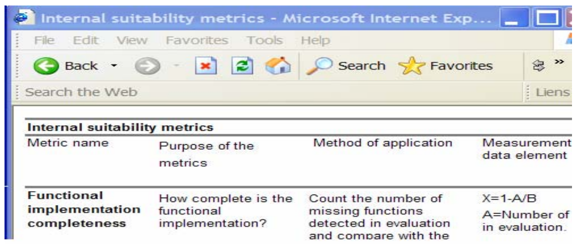

Figure 3 ISO 9126 documentation included in the QEST prototype

B. Generic Performance Management Models

Sink and Tuttle argue that the performance of an organization is a complex interrelationship between seven criteria [10, 11], as shown in Figure 4 :

1. Effectiveness is expressed as the ratio of actual output to expected output;
2. Efficiency is defined as the ratio of resources expected to be consumed to resources actually consumed;
3. Quality represents the quality criterion at the position in the systems model where it must be operationally defined, measured and managed; quality being a critical criterion at all stages of the life cycle of an organizational system;

4. Productivity is defined as the ratio of output to input, productivity being viewed has having the strongest impact on performance, as well as giving insight into effectiveness, efficiency and quality.
5. Quality of work life is an essential contribution to a system which performs well and moderates the equation between productivity and profitability. Poor results in this area usually spell failure for an organization in the long term.
6. Innovation is the creative process of successfully changing whatever it takes to survive and grow; it also moderates the equation between productivity and profitability. Poor results in this area spell failure for an organization in the long term.
7. Profitability represents the ultimate goal for any organization.

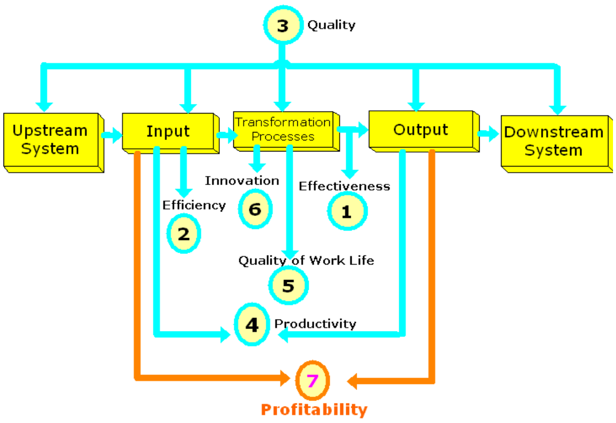

Figure 4 Sink and Tuttle Performance Criteria [10, 11]

In response to the shortcomings of traditional accounting data for performance evaluation, a new multidimensional performance model (Balanced Scorecard or BSC) was proposed by Kaplan and Norton in 1992 [12-14]. The initial BSC model evaluates corporate performance from four different perspectives: the financial perspective, the internal business process perspective, the customer perspective, and the innovation and learning perspective.

The Balanced IT Scorecard (BITS) [15, 16] proposed by the European Software Institute (ESI) is a specific version of the four original perspectives for the information technology industry[1] of the Balanced Scorecard (BSC) presented by Kaplan and Norton: financial, customer, internal process, infrastructure and innovation, and adds a fifth, the people perspective.

The Baldrige award [17] provides a full framework that any organization can use to improve overall performance. Its framework is composed of seven criteria: leadership, strategic planning, customer and market focus, measurement, analysis and knowledge management, human resource, process management and business results.

The Skandia Navigator [18] is based on the identification of critical indicators in five perspectives, all of them linked to the value creation process: a financial focus, a customer focus, a process focus, a renewal and development focus, and a human focus. A second-generation model is also available which attempts to consolidate all the different individual indicators into a single index and to correlate the changes in intellectual capital with changes in market value.

The Performance Prism measurement framework [19] has five facets representing each of five perspectives on performance in the form of a prism.

III. VISUALIZATION

The term visualization means a graphical representation of data or concepts [20]. Because having the right information at the right moment is crucial to making the right decisions, there is growing interest in data visualization in all disciplines including engineering and management.

As shown in Figure 5, the type of data to be visualized may be: one-dimensional data (eg.temporal data, textual document), two-dimensional data (eg. geographical maps), multidimensional data (eg. relational tables, statistical databases), text and hypertext (eg. Web text), hierarchies and graphs, and algorithms and software. Data visualization is used both in the initial exploration before statistical analysis and in the final display of results and model building.

Exploring and analyzing vast volumes of multidimensional data is becoming increasingly difficult without proper visualization and analysis tools. This is particularly true when little is known about the data itself and when the goals are indistinct or unclear. One of the greatest benefits of data visualization is therefore the sheer quantity of information of high dimensionality that can be rapidly interpreted if it is presented well.

In analysing data we are typically interested in detecting trends, anomalies, discontinuities, and correlations. A static presentation in itself is therefore often inadequate and interactive capabilities are required to allow effective exploration of the visualisation. Interactive visualization is a process made up of a number of interlocking feedback loops [20]: data manipulation (objects are selected and moved using the basic skills of eye-hand coordination), exploration and navigation, and a problem-solving (forms hypotheses about the data).

A number of visualization techniques exist, as shown in Figure 5, each with their strengths and weaknesses such as:

- standard 2D/3D displays: bar charts, box plots, pie charts, and scatterplots;
- geometrical transformed displays : Parallel Coordinates, Prosection Views, and Hyperslices;
- icon based displays: mapping the attribute values of each data record to the features of icons;

[1] See next subsection.

- dense pixel display: map each dimension value to a colored pixel and grouping the pixels belonging to each dimension into adjacent areas;
- stacked displays: stack the data partitioned in a hierarchical fashion.

Traditional visualization techniques such as parallel coordinates, and scatterplot diagrams, do not scale up well to data sets of high dimensionality. A common approach to solving this problem is dimensionality reduction.

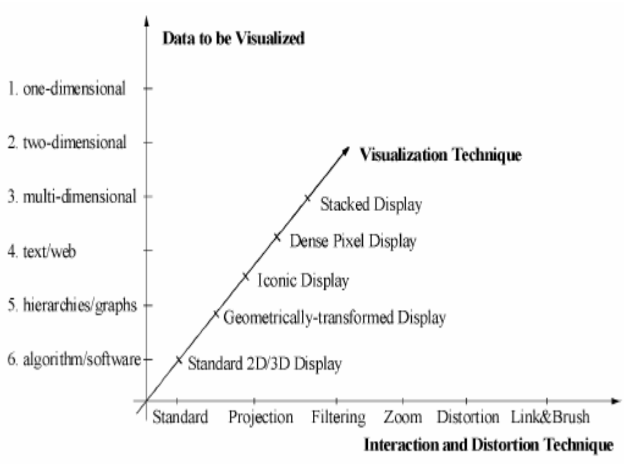

Figure 5 Classification of information visualisation techniques. Reproduced from [21]

There are three major approaches to dimensionality reduction [21] :

- Principal Component Analysis (PCA): reduction to a few dimensions that account for most of the variance within the data.
- Multidimensional Scaling (MDS): an iterative non-linear optimization algorithm for projecting multidimensional data down to a reduced number of dimensions.
- Kohonen's Self Organizing Maps (SOM): an unsupervised learning method for reducing multidimensional data to 2D feature maps.

Visual Hierarchical Dimension Reduction for Exploration of High Dimensional Datasets (VHDR) [21] generates lower dimensional spaces and allows user interactions in the dimension reduction process. DOSFA [22] (Dimension Ordering, Spacing, and Filtering Approach) is based on dimension hierarchies derived from similarities among dimensions.

Many criteria can be used to evaluate the effectiveness of a visualisation technique [23] [24]:

- Maximum number of dimensions: when the number of dimensions or variable exceeds a certain limit there is a degradation of the usefulness of the visualization;
- Maximum data set size: there is a certain amount of screen space available for effective visualization;
- Level of occlusion: different data points will map the same location on the screen- the viewer should obtain a view which avoids a given overlap;
- Support of user interaction: each projection technique has a logical set of interactive capabilities for viewing, modification and enhancement;
- Included interpretation guides: reference points to help interpret the data and its context;
- Use of color to highlight one or more variables;
- Use of 3-Dimensional cues: shading, translucency, and motion.

Any of these visualization techniques may be used in conjunction with different interaction techniques that can categorized into interactive projection, interactive filtering, interactive zooming, interactive linking and brushing. These visualization techniques can help the user complete the following primary tasks [25]:

- Classification – discovery of a predictive learning function that classifies a data item into one of several predefined classes.
- Regression – discovery of a predictive learning function, which maps a data item to a prediction variable.
- Clustering – a common descriptive task in which one seeks to identify a finite set of categories or clusters to describe the data.
- Summarization – calculates descriptive statistics for a set of data.
- Dependency Modeling – determines a model that describes dependencies between variables or between the values of a feature in a data set.

IV. Discussion

The previous sections on multidimensional performance models and on contemporary visualization techniques shows that:

- Performance management is inherently multidimensional and thus a very complex activity. This may be even more the case for a relatively immature field such as software engineering, the end-product of which is, by definition, intangible.
- Performance management is viewed quite differently from one model to another. In fact, the selected models are quite different in terms of the adopted terminology, the number of perspectives included in the model, the chosen perspectives themselves, and the indicators or measurements within each chosen perspective.
- Performance management models have been studied and applied quite extensively in management. Performance models used by software engineering managers must adopt terminology and a framework that is recognized by managers and executives outside their own software engineering organizations.

- Performance management models in software engineering must also, however, support concepts and terminology which are specific to software engineering.
- Performance management models that are specific to software engineering have been studied relatively little.
- Even though many models support multiple levels of goals, objectives and measurements or their equivalents, almost none of the models presented include a mathematical framework for handling these concurrently in an integrated manner. In addition, almost none of the models include mathematical formulae for consolidating the various performance perspectives and indicators or measurements into a single index.
- Contemporary visualization techniques are increasingly being investigated and applied in many disciplines and are seen as key to multidimensional performance modeling in software engineering.

V. PROPOSED CONCEPTS FOR A TOOL FOR MULTIDIMENSIONAL PERFORMANCE MODELING IN SOFTWARE ENGINEERING MANAGEMENT

This section describes the high-level characteristics, as shown in Figure 6, of a proposed tool for multidimensional performance modeling for software engineering managers. The proposed tool would:

- adopt the Sink and Tuttle organizational framework of performance;
- build upon the open, generic and geometrical QEST approach to performance modeling;
- enable the selection, by the user, of different visualization techniques to analyze data and show results;
- enable analysis of the impact of future potential scenarios on performance;
- use the International Software Benchmarking Standards Group (ISBSG) database as the initial test bed of data.

Modeling from various viewpoints and at various levels of abstraction is common in software engineering, especially in software design but is not very prevalent in software engineering management. The first predefined set of viewpoints adopted by the proposed tool would be the performance criteria of Sink and Tuttle: effectiveness, efficiency, quality, productivity, quality of work life, innovation and profitability. This is a comprehensive framework for understanding organizational performance that is easy to understand and which adopts terminology that is familiar to all types of managers. Open indicators will also be available to complete the Sink and Tuttle indicators and to adapt the model to any particular situation. "Open" means that an indicator can be filled and renamed according to organizational needs.

The QEST model will provide the mathematical and geometrical basis for the tool, because it produces a unit performance indicator and can handle both quantitative and qualitative measurements. The QEST *n*D model can handle the seven initial viewpoints proposed by Sink and Tuttle. Additional viewpoints can also be defined according to the objectives of the manager.

To enable the software engineering manager to analyze the multiple dimensions of performance, the tool will offer a set of visualization techniques. Moreover, the tool will allow users to interactively select and combine different visualization and standard data analysis techniques. There will notably be an interactive approach to dimension ordering, spacing and filtering for high dimensional datasets based on dimension hierarchies. The specific visualization techniques that will be available in the proposed tool remain, however, to be defined.

Since obtaining access to existing confidential organizational performance data is very difficult and collecting this type of data may require years of effort and duration for an organization embarking on a performance measurement program, the initial test bed for the proposed tool will be the software engineering project data made available by the International Software Benchmarking Standards Group (ISBSG)[2].

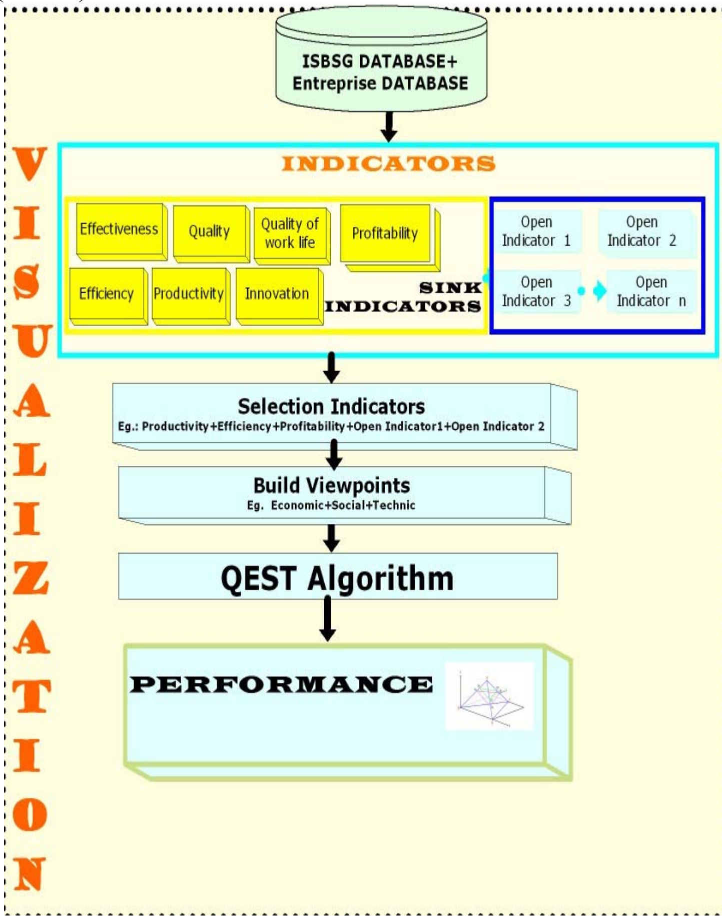

Figure 6 Proposed concepts for a tool

[2] See www.isbsg.org

The ISBSG is a not-for-profit organization whose goal is the development and management of a multiorganizational repository of software project data. This repository is available to organizations, for a nominal fee, and any organization can use it for estimation and benchmarking purposes. The 2005 version of the repository holds data on over 3,000 projects worldwide

VI. Summary

Organizational performance modeling in software engineering management is inherently multidimensional. Although many organizational performance models are proposed in the literature, current performance models in software engineering management do not enable the user to represent many possible viewpoints quantitatively and in a consolidated manner, while at the same time keeping track of the values of the individual dimensions of performance. These models therefore do not meet the analytical requirements of software engineering management when various viewpoints must be taken into account concurrently. Visualization approaches and techniques for multidimensional data are being applied in many and varied disciplines and are seen as a key enabler to organizational performance modeling in software engineering management.

This paper presented a selection of multidimensional models of performance found in software engineering and in management and presented a review of visualization techniques. A prototype tool based on the QEST model was also briefly presented and reviewed. The paper then described a set of proposed concepts for a new proposed tool for multidimensional performance modeling for software engineering managers.

References

[1] P. Bourque, V. Stroian, and A. Abran, "Proposed concepts for a tool for multidimensional performance modeling in software engineering management," presented at (Accepted for Publication in IEEE-IES International Symposium on Industrial Electronics (ISIE)), Montreal, 2006.

[2] S. E. B. Donaldson and S. G. B. Siegel, "Successful Software Development," U. S. River, Ed., 2001, pp. 745.

[3] ISO/IEC, "9126-1: Information Technology - Software Product Quality -Part 1: Quality Model." Geneva, Swizterland, 2001, pp. 36.

[4] L. Buglione and A. Abran, "A Multidimensionality in software performance measurement: the QEST/LIME models.," presented at Proceedings of SSGRR (Second International Conference on Advances in Infrastructure for Electronic Business), L'Aquila, 2001.

[5] L. Buglione and A. Abran, "Geometrical and statistical foundations of a three-dimensional model of software performance," *Advances in Engineering Software*, vol. 30, pp. 913-919, 1999.

[6] A. Abran and L. Buglione, "A multidimensional performance model for consolidating balanced scorecards," *Advances in Engineering Software*, vol. 34, pp. 339-49, 2003.

[7] L. Buglione and A. Abran, "QEST nD: n-dimensional extension and generalisation of a software performance measurement model," *Advances in Engineering Software*, vol. 33, pp. 1-7, 2002.

[8] L. Buglione and A. Abran, " A Multidimensionality in software performance measurement: the QEST/LIME models.," presented at Proceedings of SSGRR (Second International Conference on Advances in Infrastructure for Electronic Business), L'Aquila, 2001.

[9] A. Abran, M. Kunz, R. R., Dumke, and L. Buglione, "A prototype Web-based implementation of the QEST model," presented at International Workshop on Software Measurement, Montreal, 2003.

[10] D. S. Sink and T. C. Tuttle, *Planning and measurement in your organization of the future*: Norcross, Ga : Industrial Engineering and Management Press, 1989.

[11] D. S. Sink, *Productivity management : planning, measurement and evaluation, control, and improvement*: USA, 1985.

[12] R. S. Kaplan and D. P. Norton, *Strategy maps : converting intangible assets into tangible outcomes*. Boston: Harvard Business School Press, 2004.

[13] R. Kaplan and D. Norton, *The Balanced Scorecard: Translating Strategy into action*: Harvard Business School Press, 1996.

[14] R. Kaplan and D. Norton, *Putting the Balanced Scorecard to Work*, vol. 71:5: Harvard Business Review, 1993.

[15] Reo D, Quintano N., and L. Buglione, "ESI Balanced IT Scorecard Infrastructure&Innovation Perspective Technical Report," *European Software Institute*, vol. ESI-1999-TR-043, 1999.

[16] Reo D, "Applying the Balanced Scorecard for process improvement-a case study by the European software industry. Presentet ad Measuring The Business Value Of IT With The Balanced Scorecard," Phoenix, February 23 and 24, 2000.

[17] Baldrige, "Baldrige Framework For Excellence," Accessed at: http://www.quality.nist.gov/PDF_files/2006_Business_Criteria.pdf, 2006.

[18] SKANDIA, "Skandia Insurance Company," Accessed at : http://www.skandia.com/en/index/, 2005.

[19] A. Neely and C. Adams, "Perspectives on Performance: The Performance Prism," vol. 2005: http://www.exinfm.com/pdffiles/prismarticle.pdf, 2005.

[20] C. Ware, *Information Visualization: Perception for Design*: Morgan Kaufmann Publishers, 2000.

[21] J. Yang, M. O. Ward, E. A. Rundensteiner, and S. Huang, "Visual Hierarchical Dimension Reduction for Exploration of High Dimensional Datasets," *Eurographics/IEEE TCVG Symposium on Visualization*, pp. 9, 2003.

[22] W. P. Jing Yang, Matthew 0. Ward and Elke A. Rundensteiner, "Interactive Hierarchical Dimension Ordering, Spacing and Filtering for Exploration of High Dimensional Datasets," *IEEE Symposium on Information Visualization*, pp. 7, 2003.

[23] M. O. Ward, "XmdvTool : Integrating Multiple Methods for Visualizing Multivariate Data," *IEEE Visualization*, pp. 8, 1994.

[24] M. Tory and T. Moller, "Evaluating visualizations: do expert reviews work?," *IEEE Computer Graphics and Applications*, vol. 25, pp. 8-11, 2005.

[25] M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*: John Wiley & Sons, 2003.