VISUAL DATA MINING TOOLS: QUALITY METRICS DEFINITION AND APPLICATION

Edwige Fangseu Badjio, François Poulet
ESIEA Recherche,
Parc Universitaire Laval – Changé
38, rue des Docteurs Calmette et Guérin
53000 Laval France

Keywords: HCI, visual data mining, software quality, utility, usability, acceptability, evaluation, GQM model

Abstract: The main purpose of this work is to integrate HCI (Human Computer Interaction) requirements in visual data mining tools engineering. We present the definition of metrics/measurements in order to improve the quality of those tools at all the steps or after the development process. On the basis of these metrics/measurements, we have derived a questionnaire for the evaluation of the utility, the usability and the acceptability of visual data mining environments. A case study enables us to concretely materialize the contribution of the measurements and also to detect and explain (design and usage) errors.

1 INTRODUCTION

According to (Fayyad et al. 1996) data mining is the non-trivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data. Visual data mining consists in the use of visualization as a communication channel for data mining. For (Wong, 1999), visual data mining lies in tightly coupling the visualizations and analytical process into one data mining tool that takes advantage of the strengths of all worlds.

The standards ISO 9241, ISO/IEC 9126, ISO 13407, (Nielsen, 1993), (Bastien et Scapin, 1999) criteria propose attributes which characterize software quality in terms of usability. These attributes include: the facility to learn the software, the user's satisfaction, comprehensibility, efficacy, operability, attractiveness etc... How can we concretely incorporate these various attributes in a visual data mining environment? We know that finding and fixing a software problem after delivery is often 100 times more expensive than finding and fixing it during the requirements and design phase (Boehm et Basili, 2001).

From the user's point of view, the interface is the most important element of the software because it is the user's mediator with the system for the achievement of his task (Costabile, 2000). We try to bring replies to this question. Indeed, some research works (Kolski et al., 2001) insists on the fact that when the software is highly interactive, an adapted methodological step is essential, the traditional cycles of the software engineering being insufficient.

We have defined a set of metrics/measurements (recommendations) for the improvement of the quality of visual data mining tools. The application of these recommendations allows the correction of errors likely to occur with less expense than at the end of the design process. Based on the formal model GQM (Basili et al., 1994), the proposed approach also allows an evaluation after the tool design and development in order to explain the minor and major errors detected.

The overview of this paper is the following: first, we explain the theoretical foundations of our work before the measurement definition and the questionnaire presentation, section 4 presents a case study followed by the conclusion and future works.

2 THEORETICAL FOUNDATIONS

The utility of a visual data mining tool relates to the adequacy existing between the functions provided by the system and those necessary to the user in order to achieve the visual mining tasks assigned to him.

The usability is the quality of hardware or software which is easy and pleasant to use and to understand, even by somebody who has little

Fangseu Badjio E. and Poulet F. (2005).
VISUAL DATA MINING TOOLS: QUALITY METRICS DEFINITION AND APPLICATION.
In *Proceedings of the Seventh International Conference on Enterprise Information Systems*, pages 98-103
DOI: 10.5220/0002547700980103

knowledge in data processing. Usability is a critical quality of the success of a software product.

The acceptability relates to the adequacy between the decision to use the visual data mining environment and the compatibility of the tool for the user and the organization to which he belongs.

A measurement is a support allowing answering a variety of questions related to the software development process. It also allows evaluating and measuring the quality of the end products.

A metric is a mathematical number that shows a relationship between two variables.

The theoretical base of this research work is software quality studies. Quality is the aptitude of a product or a service to satisfy the user's needs. One of the earlier works of the software quality field is McCall's (McCall et al, 1977) model which counts around fifty criteria allowing the expression of software quality in general. Another set of software quality factors was carried out by (Boëhm, 1978). These factors are related to the software functionalities and performances from the software engineering point of view and the usability. From the human machine interaction and software engineering point of view, the awaited aptitudes of the software analysis and different level of evaluation: the satisfaction of the user's needs, reliability, interoperability, conformity with the standards and a good ratio cost/performance. More recently, case studies such as (Nielsen, 1993), (Bastien et Scapin, 1999) stress the ergonomic evaluation and improvement of the user interface much more. Cognitive psychology work related to data visualization (Healey, 1996) proposes also interesting primitives in such cases.

The principal advantage of ergonomic inspection is its level of detail. Indeed, this method guarantees an exhaustive analysis of the overall software.

For the needs of data mining algorithm's technical aspect evaluation, data sets repository creation efforts was the object of projects such as UCI (Blake et al, 1998). In visualization for mining purpose, Grinstein's research work (Grinstein et al, 1997) allowed the evaluation of data representation methods. The evaluation criteria were: the memory size of the computers, their execution speed and their graphical capacities. It should be noted that we have taken into account all the various details explained in this section for the development of our evaluation method so that it is as most complete as possible.

For the visual data mining tools measurements specification, our formal framework is GQM "Goal/Question/Metrics" (Basili and al. 1994) model. GQM is a support making it possible to better clarify the objectives to be reached, a set of attributes are identified via questions. Downwardly, measurements are defined. For the evaluation and the explanation, one proceeds in an ascending way.

If one is located in the measurement specification model (figure 1), our objective will be to find the means of specifying each usability criterion described by the various ISO standards in order to apply it to visual data mining specificities. For example in order to develop the user's guidance (goal), the question will be how to advise, direct our end-users? One of the answers is to provide decision support or recommendation for the selection of the data mining algorithm to be performed.

Figure 2 describes the measurement definition process. First, we have done an analysis of end users followed by a task analysis, the usability goals definition in a visual data mining context. More description of the new evaluation approach is the subject of the following section.

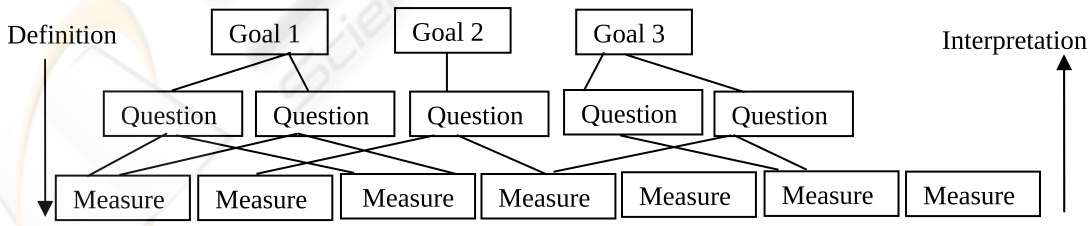

Figure 1: GQM Model (Basili et al 1994)

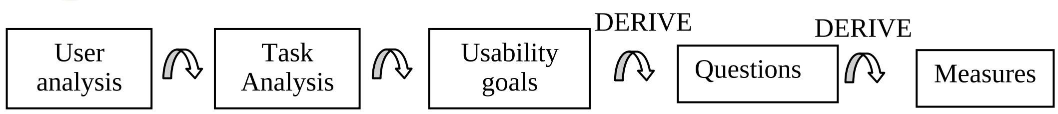

Figure 2: Measure definition process

3 METRIC DEFINITION

Metrics/measurements are important for software engineering activity. Developers use metrics for controlling software quality throughout the project life cycle. By using software measurements, the managers can see measurable attribute of the software quality. Customers look for the measurements in order to determine the quality of the products. Maintainers use metrics as an indicator for reusability or reengineering.

3.1 User analysis

The end-user of visual data mining tools could be the data specialist. He has a basic knowledge in data analysis. This data mining approach has been developed for the need of integrating the user in the knowledge discovery process in order to combine the human potentialities of judgement with the computer calculation capacities.

3.2 Task analysis

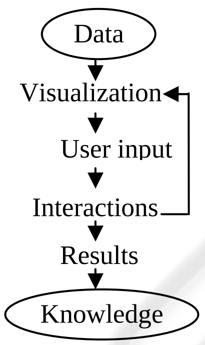

Figure 3: our visual data mining task model

(Ankerst, 2000) describe a task model for visual data mining field. This model includes data visualization and application of mining algorithms steps. Ankerst's model more explicitly reduced the visual data mining to the use of data visualization during the pre-treatment (data exploration phases), the treatment (knowledge discovery step) or the post processing (knowledge representation step). This model does not correspond to our visual data mining model detailed by the figure 3. The user can visualize the data in pre-treatment. The fundamental difference is the fact that the user interacts with a graphical representation (chart) of that data in knowledge discovery in the data. The data model (knowledge) is built in an interactive way.

3.3 Usability goals: G of GQM

The ISO 9241, ISO/IEC 9126, ISO 13407 standards, (Bastien et Scapin, 1999) and (Nielsen, 1993) work propose metrics for software usability improvement. (Bastien et Scapin, 1999) metrics are used like support. These metrics constitute the G (goals) side of the GQM model. These metrics are:

Guidance: means implemented in order to advise, direct, inform and lead users.

Workload: the interface elements must reduce the users' perceptive load, just as in the improvement of the dialogue efficacy.

Explicit control: possibility for the system to take into account the explicit actions of the users and the control they have, relating to their actions treatment.

Adaptability: system capacity to react according to the context, the needs and the user preferences.

Error management: means allowing to reduce errors and to correct them when they occur.

Compatibility: agreement between the users' characteristics, the tasks and their organisation.

In the following sections, we will use the GQM model (figure 1) to formalize these usability needs according to the visual data mining field.

3.4 Questions to answer for improving usability and utility: Q of GQM

After having defined our GQM model goals, the second stage consists of the definition of a set of questions whose responses will allow the achievement of the goals. The task analysis enabled us to define also a set of questions related to utility (technical quality of tools) and to mix them with usability questions. Due to space constraints, the set of questions can not be presented here.

3.5 Usability and utility measures: M of GQM (top level)

We studied the visual data mining process in order to bring responses relating to usability's and utility's questions. These responses include the specificities of visual data mining tools. For example, for the users' workload reduction, it is judicious to give them the possibility of choosing a visualization method among several possible alternatives.

Several analysis methods can be included in a data mining environment. For the usability (guidance) improvement, the developers have to advise users in the choice of the most suitable method. The execution of an analysis method is time consuming. If the method execution does not finish under good conditions it is necessary to envisage the re-use of training data. In order to take into account user preferences, developers must also envisage several visualisation methods.

3.6 Usability and utility measures: M of GQM (lower level)

There are several methods for software evaluation needs. We have chosen a questionnaire for the setting up of our evaluation method because its does not require any user competence. The questionnaire has six topics. It provides a methodology to verify if the desired quality requirements (G of GQM) have been satisfied. It is also used to conduct tests, conduct reviews and audits and to review software program design. Table 1 presents rating levels in order to fill in the questionnaire. The best level possible with our rating level is 5, the worst possible level is 1.

Table 1: Rating levels

Mark	Rating
1	Absent, Strongly disagree, Extremely (difficult, confusing, boring)
2	Poor, Somewhat disagree, Somewhat (difficult, confusing, boring)
3	Fair, Neutral
4	Good, Somewhat agree, Somewhat (easy, clear, fun)
5	Excellent, Strongly agree, Extremely (easy, clear, fun)

The first topic of our questionnaire is the visual data mining tool **technical quality**, (table 2) it relates to technical aspect: data-processing capabilities, operating system, speed, compatibility, etc… The perception of the tool technicality allows the measurement of the tool power given the capacities offered. The evaluation on this level refers to the adaptability to the task, with the implementation precision, the capacity of knowledge prediction. The measurements based on technology also allow the evaluation of the degree with which the system can handle data of variable sizes.

The **Scenario** topic refers to the execution details of the visual data mining software. These details particularly relate to the interaction quality and time necessary for this purpose.

The topic **Interface presentation model (IPM)** refers to the elements of the graphic interface (color, typography …) and allows the consideration of aesthetics and attracting aspects of the tool.

The **Visualization** topic relates to the relevance of the charts and their structure compared to the mining objectives of the user and his profile. It is question here of seeing in which cases the charts used facilitate knowledge perception and comprehension.

The **Usability** topic is based on general recommendations for the ergonomic design of the interface.

The topic **User** allows defining explicitly the user profile and takes it into account for the software improvement. User topic also makes it possible to have an overall perception of the use of the visual data mining tool, of its design features, such as the adaptability and the adequacy of the system, of its communication and control facilities, its robustness and its effectiveness, its simplicity of implementation and comprehension, as well as its convivial and personalized character. This topic makes it possible to understand the total reaction of a user compared to the mode and structure of the interaction, the communication means used, flexibility and assistance.

Table 2: Technical quality

	1	2	3	4	5
Installation					
Assistance					
Portability					
Architecture					
Heterogeneous data access					
Algorithms diversity					
Data models validation					
Results appearance					
Models exportation					
Interoperability					
Efficiency					
Robustness					
Re use of training data					
Treatment of data sets with large dimension					

4 CASE STUDY

In this section, we will apply the questionnaire in order to evaluate the quality of two methods included in a visual data mining tool. These methods are CIAD (Poulet, 2001) and PBC (Ankerst et al., 1999). We listed four visual data mining methods. But two of these methods are not free.

These evaluated methods relate to visual data mining: the use of two dimensional visualizations techniques and the capacity of the system to treat complex interactions between all the possible pairs of attributes. The test is performed by four autonomous users (a researcher and three PhD students in visual data mining field). This evaluation aims to detect and explain design problems (usage step).

4.1 Evaluation tasks

For our evaluation needs, the prescribed task is the interactive construction of decision trees starting from representations of small data sets like Iris (150 records, 4 numerical attributes, 3 classes), glass (214 records, 6 numerical attributes, 9 classes), ionosphere (351 records, 2 numerical attributes, 34 classes) data sets from the UCI (Blake et Merz, 1998). The decision trees allow partitioning a great quantity of data in small segments by application of a series of decision rules. They are very much used in data mining. Their coupling with data visualization leads to tools for interactive construction of decision trees in which the end-users could be the data field specialists. The quality of the models resulting from the visual data mining depends on the quality of the method used. It is thus necessary to develop tools useful, usable and acceptable for data field expert's users. The following paragraph describes the interpretation of the obtained results.

4.2 Evaluation results

Table 3 presents the mean ratings of the case study evaluation questionnaires. The first six lines represent the topics of the evaluation questionnaires. We present also five sub topics (the last five lines of the table 3) which we considered to be more relevant.

The evaluators agree on the fact that the tools are very useful. The error's treatment is done in a relevant manner; CIAD and PBC are thus convenient. Also, for an autonomous user, the tool is very easy to understand and to use. The interfaces presentation models are well developed. The elements disposal on the screen is very good; graphics and colours are well used. CIAD allows the training data set re-use, which makes it possible to reduce the users' workload. Such is not the case for PBC.

The evaluators also agree on the fact that the installation of these tools is not obvious. These tools are designed for experimental purposes. Only one algorithm and only one visualization method are implemented for PBC and CIAD. The users can not assess prefered analysis methods or visualization tools. It is not possible to access various data set formats. For CIAD and PBC, the users are not directed (guided), also, it misses the on line help, the contextual menus, the user manual. The results of this evaluation allow the designers of CIAD and PBC to develop the aspects related to the usability (assistance modules, user manual, several alternatives possible with regard to data analysis methods and data visualization, cognitive aspects of visualization for data mining, user preferences) of these tools and thus to work on their acceptability.

Table 3: Mean rating scores

	PBC	CIAD
Guidance	1	1.5
Workload	1	1.5
Explicit control	1	1.5
Adaptability	1	1
Error management	3.5	3.75
Compatibility	2	3
Usefulness	5	5
Ease of use	3.25	3.75
Assistance	1	1
Learning ability	3.5	3.5
Installation	1	1

5 CONCLUSION

In order to avoid the redevelopment of design which generates wastes of time and high cost of production without however guaranteeing performance, we study the development of useful, usable and acceptable visual data mining software. To this end, we have drawn up a list of criteria having to be taken into account for a development of reliable tools of this type and for their evaluation. Also, in order to cover all the aspects of that type of tool analysis, we listed six topics. Those topics are defined in the form of a tree structure including the principal topic, under topics or meta-criteria and criteria. These various criteria have been used for the evaluation of CIAD and PBC, modules dedicated to the interactive construction of decision trees. The evaluation allowed us to use one of the advantages of our analysis approach. It allows a thorough evaluation of the software (interface, technical quality, ergonomic (usability), visualization and scenario). The design problems are thus discovered on all the tool levels.

The end users of visual data mining tools could be data domain specialists. They don't have knowledge needed for the selection of the best tool available for the treatment of their data. From the method presented above, we work now on the definition of a preference index for the recommendation of visual data tools to these users.

REFERENCES

Ankerst M., 2000. "Visual Data Mining", PhD Thesis, Ludwig Maximilians University of Munich, 2000.

Ankerst, M., Elsen C., Ester, M., and Kriegel, H.-P., 1999. Visual classiffication: An interactive approach to decision tree construction. In Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, p. 392-396.

Basili, V., Caldiera, G. and Rombach, D., 1994. The Goal Question Metric Approach. Encyclopedia of Software Engineering. Wiley 1994.

Bastien, J.M.C., and Scapin, D.L., Leulier C., 1999. The ergonomic criteria and the ISO/DIS 9241-10 dialogue principles: a pilot comparison in an evaluation task. Interacting with Computers, Vol. 11, n° 3, p. 299-322, 1999.

Blake, C., Merz, C., 1998. UCI Repository of machine learning databases, [www.ics.uci.edu/~mlearn/MLRepository.html]. Irvine, University of California, Department of Information and Computer Science.

Boehm, B. and Basili, V., 2001. Software Defect Reduction Top 10 List, IEEE Computer, Vol.. 34, No. 1, January 2001.

Boëhm, B., 1978. Characteristics of software quality, Vol 1 of TRW series on softwaretechnology, North-Holland, Amsterdam, Netherlands, 1978.

Costabile, M.F., 2001. Usability in software life cycle In Handbook of of Software Engineering and Knowledge Engineering, SK Chang Ed., World Scientific, Vol , World Scientific Publ. Company, 2001.

Fangseu Badjio E., Poulet F., 2004, A decision support system for data miners, in proc. of AISTA'04, The International Conference on Advances in Intelligent Systems - Theory and Applications in cooperation with IEEE, Luxembourg-Kirchberg, Luxembourg, November 15 – 18 2004, ISBN 2-9599776-8-8.

Fangseu Badjio E., Poulet F., Usability of Visual Data Mining Tools, in proc. of ICEIS'04, the 6th International Conference on Enterprise Information Systems, Porto, Portugal, April 2004, Vol.5, 254-258, ISBN: 972-8865-00-7.

Fayyad, U. M., Piatetsky-Shapiro, G., and Smyth, P. 1996. editors Advances in Knowledge Discovery and Data Mining. AAAI Press / MIT Press, Menlo Park, CA.

Grinstein ,G. G., Hoffman, P., Laskowski ,S. J. and Pickett, R. M., 1997. Benchmark Development for the Evaluation of Visualization for Data Mining Proc. of the Workshop: Issues in the Integration of Data Mining and Data Visualization, Newport Beach, California, August 17, 1997.

Han, J., Cercone, N., 2001. "Interactive Construction of Decision Trees" in proc. of PAKDD'2001,LNAI 2035, 575-580, 2001.

Healey, C.G., 1996. Choosing Effective Colours for Data Visualization , in Proceedings of the 7th conference on Visualization '96, p 263-270.

ISO (International Organization for Standardization), 1997. ISO 9241: Ergonomics Requirements for Office Work with Visual Display Terminal (VDT) - Parts 1-17.

ISO (International Organization for Standardization), 1998. ISO 13407: Human-Centered Design Process for Interactive Systems, 1998.

ISO/IEC (International Organization for Standardization and International Electrotechnical Commision), 1998. ISO/IEC 9126-1: Information Technology – Software Product Quality, 1998.

Kolski, C., Ezzedine, H., Abed, M., 2001. Développement du logiciel: des cycles classiques aux cycles enrichis sous l'angle des IHM. In Kolski C. (Ed.), Analyse et Conception de l'IHM. Interaction Homme-machine pour les SI, vol. 1, Hermès, Paris, pp. 23-49, ISBN 2-7462-0239-5.

McCall, J.A., Richards, P.K., and Walters, G.F., 1977. Factors in software quality, Vols I-III, Rome Air Development Centre, Italy.

Murine, G. and Carpenter, C., 1984. Measuring software product quality", Quality progress, Vol 7(5), p16-20,.

Nielsen, J., 1993. Usability Engineering, Academic Press, Cambridge, MA.

Nielsen, J., 2000. Why You Only Need to Test With 5 Users http://www.useit.com/alertbox/20000319.html.

Poulet, F., 2001. "CIAD: Interactive Decision Tree Construction, in proc. of VIII Rencontres de la Société Francophone de Classification, Pointe-à-Pitre, 275-282.

Ware, M., Franck, E., Holmes, G., Hall, M., Witten, I., 2001. "Interactive Machine Learning: Letting Users Build Classifiers", in International Journal of Human-Computer Studies (55), 281-292, 2001.

Wong, P.C., Visual Data Mining, in IEEE Computer Graphics and Applications, 19(5), p.20-21, Sept.Oct. 1999.