International Journal of Multidisciplinary Research and Development
Online ISSN: 2349-4182 Print ISSN: 2349-5979
www.allsubjectjournal.com
Volume 2; Issue 11; November 2015; Page No. 686-688

Analysis of conceptual models and metrics

Naveen Dahiya

Department of CSE, MSIT, New Delhi, India

Abstract
The paper presents study of data warehouse conceptual models in context of quality metrics and the role played by quality metrics in quality evaluation of models. A new quality metric is proposed and the importance of metric in quality evaluation of conceptual models is discussed.

Keywords: Conceptual Model, Quality metrics, Data warehouse, Understand ability, Quality evaluation

1. Introduction

An efficient data warehouse system is capable enough to provide strategic information to the business analyststhat include knowledge of variable trends about customer needs and preferences, emerging technologies, sales and marketing techniques, quality of products and services over a period of time. Efficient data warehouses are need of the time as they help to gain competitive advantage to business organizations (Inmon 2005) [3].

Data warehouse system development follows incremental approach starting with conceptual, than to logical and finally physical design phase. The significance of conceptual design phase in data warehouse development towards building of an efficient data warehouse can be understood from the fact that a strong foundation gives way to the erection of a long lasting building. The conceptual phase is the foundation of data warehouse design process. The output of conceptual phase is a conceptual model. The quality of conceptual model can be evaluated using various quality metrics. A brief study of the quality metrics is presented in the successive section.

2. Literature Review

The section throws light on the quality metrics proposed by researchers for predicting quality of conceptual data warehouse models.

- Basili *et al* 1996 [5] discussed the importance and need of experimentation in empirical validation. The main focus was on defining goal and hypothesis, experimentation and validity threats to experiments.
- Calero *et al* 2001 proposed various metrics for different configurations of data warehouse models.
- Table metrics: NA, NFK
- Star metrics: NDT, NT, NADT, NAFT, NA, NFK, RSA, RFK
- Model metrics: NFT, NDT, NSDT, NT, NAFT, NADT, NASDT, NA, NFK, RSDT, RT, RFK, RSDTA
- Moody (1998) [6] introduced twenty nine metrics which measure the quality of a conceptual data model in various aspects.
- Serrano *et al* (2007) [8] proposed a set of metrics defined as follows:

1. NDC(S) Number of dimension classes.
2. NBC(S) Number of base classes.
3. NC(S) Total number of classes.
4. RBC(S) Ratio of base classes. Number of base classes per dimension class.
5. NAFC(S) Number of FA attributes of the fact class.
6. NADC(S) Number of D and DA attributes of the dimension classes.
7. NABC(S) Number of D and DA attributes of the base classes.
8. NA(S) Total number of FA, D and DA attributes.
9. NH(S) Number of hierarchy relationships.
10. DHP(S) Maximum depth of the hierarchy relationships.
11. RSA(S) Ratio of attributes.

- Serrano *et al* (2008) [7] described various that helps to choose the best schema among the alternative schemas, defined as follows:

1. NFT(Sc) Number of fact tables in the schema
2. NDT(Sc) Number of dimension tables in the schema
3. NFK(Sc) Number of foreign keys in all the fact tables of the schema
4. NMFT (Sc). Number of facts in the fact tables

- Dahiya *et al* (2015a) [1] proposed a metric NRFD i.e. number of relation between fact class and dimension classes.

3. Introduction to Proposed Metric

A data warehouse model consist of fact and dimension classes. The fact is analyzed against various dimensions. Facts are measurable and are numeric while dimensions are non-numeric in nature.

Let us understand this with the help of an example for the model shown by Fig. 1. The model analyzes manufactured part of cars along multiple dimensions namely plant, supplier and package. Part is contains fact attributes namely number, unit and weight.

The values of quality metrics for model of Fig. 1 is shown in Table I. The metrics taken for reference in Table 1 are those proposed by Serrano *et al* (2007) [8] and Dahiya *et al* (2015a) [1].

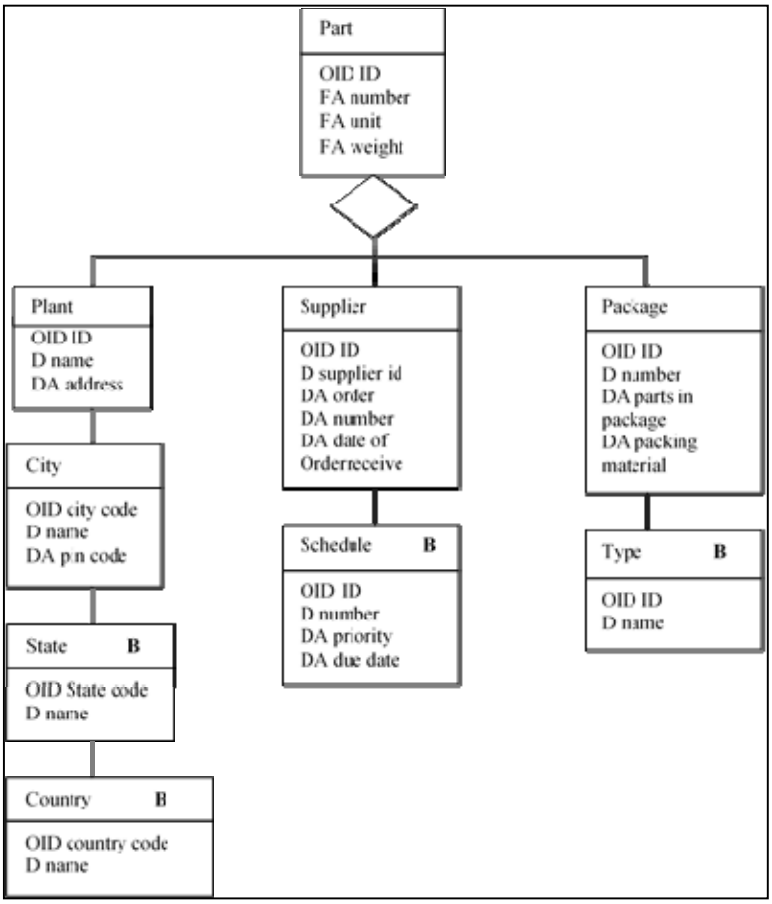

Fig 1: Conceptual model for manufacturing parts (Dahiya *et al* 2015b) [2]

Table 1: Metrics

Metrics	NFC	NDC	NBC	NC	RBC	NAFC	NADC	NABC	NA	NH	DHP	RSA	NRFD
Values	1	3	5	9	1.6	3	9	8	20	3	3	.18	3

The authors searched for the new quality factors that might affect the quality (understand ability) of conceptual data warehouse models, other than the already proposed metrics. The authors came out with a new quality metric namely number of relations (NR). NR is the sum total of all the relations between facts and dimensions, relations between dimension classes and base classes, relation between different base classes. The proposed metric may have a significant effect on the structural complexity of conceptual models along with their understand ability.

A. Importance of proposed metric

In a conceptual model there exist fact classes and dimension classes. The fact are analyzed along dimensions. There exist relations between facts and dimensions, relations between dimension classes and base classes, relation between different base classes. As the number of relations increase in a conceptual model, its size and structural complexity also increase in proportion. As the size and structural complexity of a conceptual model increase, its understanding time also
Increase, as the users will take more time to understand a complex conceptual model. So it is predicted that the proposed Metric will play a significant part in quality evaluation of conceptual models.

4. Conclusion

In the paper, the study of conceptual data warehouse models was presented in relation to the role played by quality metrics in quality evaluation of models. Authors proposed a new quality metric that is based on size and structural properties of conceptual data warehouse models and the importance of proposed metric was discussed. The current work can be extended to carry out theoretical and empirical validation of the proposed metric to judge its significance in quality evaluation of conceptual data warehouse models.

5. References

1. Dahiya N, Bhatnagar V, Singh M. Enhancing consistency of conceptual data warehouse design, International Journal of Computational Systems Engineering. 2015; 2(1):11-24 a.
2. Dahiya N, Bhatnagar V, Singh M. An empirical experimentation towards predicting understandability of conceptual schemas using quality metric," International Journal of Big Data Intelligence. 2015; 2(1):9-22 b.
3. Inmon WH. "Building the data warehouse," *Johnwiley& sons*, 2005.
4. Calero C, Piattini M, Pascual C, Serrano M. Towards Data warehouse Quality Metrics, International

Workshop on Design and Management of Data Warehouses (DMDW'01), 2001.
5. Basili V, Briand L, Melo W. A validation of object-oriented design metrics as quality indicators, IEEE transactions software engineering, 1996; 22(10):751-61.
6. Moody DL, Metrics for evaluating the quality of entity relationship models, in Proceedings of the Seventeenth International Conference on Conceptual Modeling–ER'98 (pp. 211-225). Springer Berlin Heidelberg, 1998; 211-225.
7. Serrano MA, Calero C, Sahraoui HA, Piattini M, Empirical studies to assess the understandability of data warehouse schemas using structural metrics, Software Quality Journal. 2008; 16(1):79-106.
8. Serrano M, Trujillo J, Calero C, Piattini M, Metrics for data warehouse conceptual models understandability, Information and Software Technology, 2007; 49(8):851-870,