International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-5, Issue-2)

Research Article | **February 2016**

Study of Conceptual Models from the Perspective of Quality Metrics

Naveen Dahiya
Depaertment of CSE, MSIT,
Delhi, India

Abstract—

The paper discuss the role of quality metrics to predict the quality of data warehouse conceptual models in terms of understandability. Also a new quality metric is proposed and its significance discussed in predicting quality of conceptual models.

Keywords-- Conceptual Model, Quality metrics, Data warehouse, Understandability, Quality evaluation.

I. INTRODUCTION

According to W.H. Inmon "A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process" [1]. Quick retrieval of reliable information is the requirement of time to make correct decisions and to sustain growing market competition. An efficient data warehouse enables organizations to achieve the goal. There are three design phases of data warehouse: conceptual phase, logical phase and physical phase. A good quality conceptual model will give better design at logical and physical phases as well. Various quality metrics had been proposed by researchers to evaluate the quality at conceptual level of data warehouse design. A brief literature review on the role of various quality metrics for the conceptual models is presented in the paper. A new quality metric is proposed namely: the number of multilevel hierarchies in the conceptual models.

II. LITERATURE REVIEW

This section provides a review on the metrics proposed by researchers towards quality evaluation of conceptual data warehouse models.

- Gray et al. (1991) [2] proposed various objective and open ended metrics to evaluate the quality of an entity relationship diagram. These metrics helps to identify the problems of the database design and optimize it.
- Moody (1998) [3] introduced twenty nine metrics which measure the quality of a conceptual data model in various aspects.
- Serrano et al (2007) [4] proposed a set of metrics defined as follows:
  i. NDC(S) Number of dimension classes.
  ii. NBC(S) Number of base classes.
  iii. NC(S) Total number of classes.
  iv. RBC(S) Ratio of base classes. Number of base classes per dimension class.
  v. NAFC(S) Number of FA attributes of the fact class.
  vi. NADC(S) Number of D and DA attributes of the dimension classes.
  vii. NABC(S) Number of D and DA attributes of the base classes.
  viii. NA(S) Total number of FA, D and DA attributes.
  ix. NH(S) Number of hierarchy relationships.
  x. DHP(S) Maximum depth of the hierarchy relationships.
  xi. RSA(S) Ratio of attributes.
- Serrano et al (2008) [5]described various that helps to choose the best schema among the alternative schemas, defined as follows:
  i. NFT(Sc) Number of fact tables in the schema
  ii. NDT(Sc) Number of dimension tables in the schema
  iii. NFK(Sc) Number of foreign keys in all the fact tables of the schema
  iv. NMFT(Sc). Number of facts in the fact tables
- Dahiya et al (2015a) [6] proposed a metric NRFD i.e. number of relation between fact class and dimension classes with in a schema and gave its theoretical validation.

III. PROPOSED METRIC AND PRELIMINARIES

A data warehouse schema has fact and dimension classes. Fact class represents the main subject that is to be analysed and is numeric. The fact is analysed against various dimensions.

This section provides a demonstration to calculate values of metrics, already proposed by Serrano et al (2007) [4]and Dahiya et al (2015a) [6], for a conceptual data warehouse model for analysing manufactured part of cars along multiple dimensions namely plant, supplier and package shown in Fig. 1. Part is fact class which contains specific measures called fact attributes namely number, unit and weight of a part to be analysed. Plant, supplier and package are dimension classes.

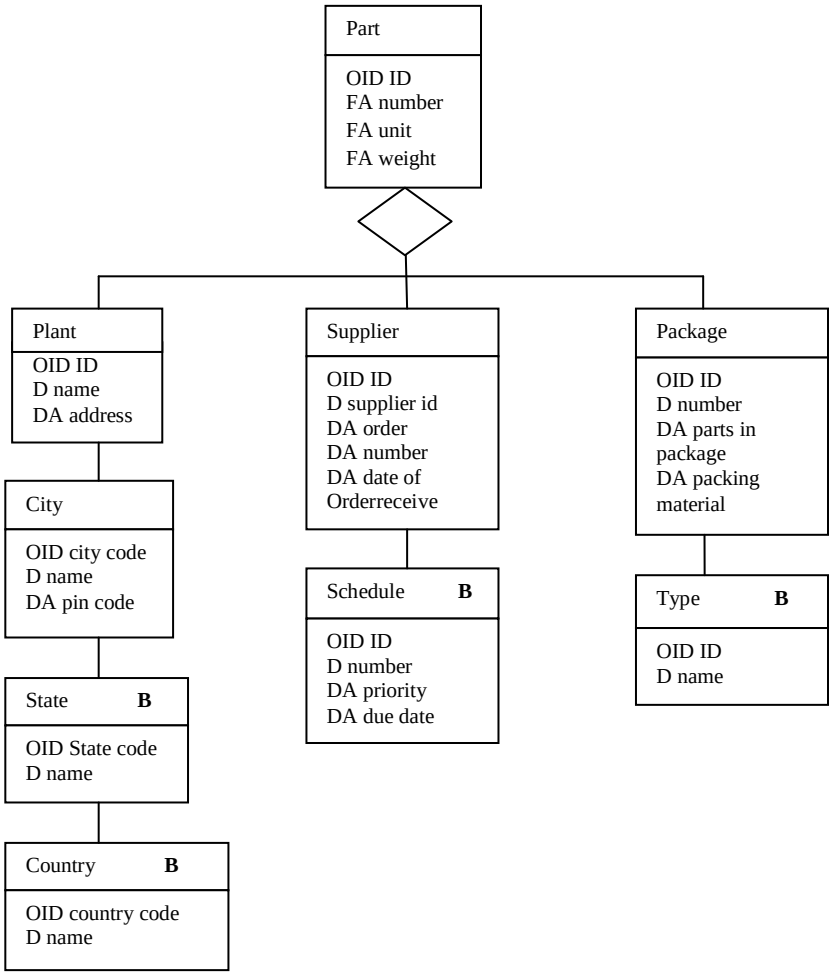

Fig 1. UML Class Diagram for manufacturing parts (Dahiya et al 2015b) [7]

The values of quality metrics for UML class diagram of manufacturing parts is shown in Table I.

Table I 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 study of quality metrics gave required motivation to propose a new quality metric. A situation might arise to find best possible configuration from existing conceptual models for same domain, where all of the proposed metrics (defined above) may have the same values. This lead to the investigation of other quality factors that might affect the understandability (time taken by users to understand a conceptual model and answer questions based on the model) of models other than already proposed. The authors find out that a new quality metric namely number of multi-level hierarchies (NMH). Number of multilevel hierarchy (NMH) in a conceptual model is the number of hierarchies where one dimension is related simultaneously to more than one base class. The proposed metric may have a significant effect on the structural complexity of conceptual models along with their understandability.

IV. IMPORTANCE OF PROPOSED METRIC

As we know that conceptual model consists of fact classes and dimension classes. The fact class contains important subjects that are to be analyzed by business executives, managers. Dimensions provide a way to measure facts. Multi-level hierarchies may exist at the dimension level. The increase in the number of multi-level hierarchies, increase the structural complexity of data warehouse models. With the increase in structural complexity, it is predicted that understanding time of the model also increases. Understanding time is directly proportional to structural complexity of data warehouse conceptual model. It is predicted that proposed metric will play an important role in predicting the understandability of conceptual models.

V. CONCLUSION

The author presented the role of quality metrics in predicting the understandability of conceptual models. Author proposed a new metric based on size and structural properties of data warehouse conceptual model that may play a significant role in predicting the quality of data warehouse conceptual model. The future research can be carried out to prove the significance of proposed quality metrics by carrying out theoretical and empirical validation of proposed metric.

REFERENCES

[1] W.H. Inmon, "Building the data warehouse," *Johnwiley& sons*, 2005.

[2] R. Gray, Carey, N. McGlynn, and A. Pengelly, "Design metrics for database systems," *Journal of BT Technology*, vol. 9(4), pp. 69-79,1991.

[3] D. L. Moody, "Metrics for evaluating the quality of entity relationship models," in p*roceedings of the Seventeenth International Conference on Conceptual Modeling–ER'98,* Springer Berlin Heidelberg, pp. 211-225, 1998.

[4] M. Serrano, J. Trujillo, C. Calero, and M. Piattini, "Metrics for data warehouse conceptual models understandability," *Information and Software Technology*, vol. *49*(8), pp. 851-870, 2007.

[5] M. A. Serrano, C. Calero, H. A.Sahraoui, and M. Piattini, "Empirical studies to assess the understandability of data warehouse schemas using structural metrics," *Software Quality Journal*, vol. *16*(1), pp. 79-106, 2008.

[6] N. Dahiya, V. Bhatnagar, and M. Singh, "Enhancing consistency of conceptual data warehouse design," *International Journal of Computational Systems Engineering*, vol. *2*(1), pp. 11-24, 2015a.

[7] N. Dahiya, V. Bhatnagar, and M. Singh, "An empirical experimentation towards predicting understandability of conceptual schemas using quality metric," *International Journal of Big Data Intelligence*, vol. *2*(1), pp. 9-22, 2015b.

