SURVEY BASED ON AN APPROACH TO RETRIEVE LOGICAL SCHEMA FROM SHOPPING CART DATABASE

Rinkalkumar Patel[1], Shilpa Serasiya[2]

[1] *Student-ME, Computer Engineering Department, Kalol Institute of Technology & Research Center-Kalol, Gujarat, India*
[2] *Professor, Computer Engineering Department, Kalol Institute of Technology & Research Center-Kalol, Gujarat, India*

ABSTRACT

In this paper, we present an approach to retrieve logical schema from shopping cart database systems. In general, an e-commerce system is built by following one of two approaches. The first approach is the customization approach using a suite of tools such as IBM's Web Sphere Commerce Suite. For example, the Commerce Suite provides tools for creating the infrastructure of a virtual shopping mall, including catalog templates, registration, shopping cart, order and payment processing, and a generalized database. The second approach is the bottom-up development of a system in-house by experts of an individual company. In this case, the developer is manually building a virtual shopping mall with mix-and-match tools. In addition, a database supporting the business model of the e-commerce system must be manually developed. Whether a developer is using the customization or the bottom-up approach, understanding the structure of e-commerce database systems will help the database designers effectively develop and maintain the system.

Keyword: - *Denormalization, RDF, XML, Ontology vocabulary, Unstructured to structured data*

1. INTRODUCTION

Data is not stored on a single computer, because current era is the era of information technology & social media, that user provided data can be stored in many more computers on the internet, so it is difficult for them to access quickly and easily. The data is to be in the format as RDF/XML, N-Triples and OWL or with the same specifications [1]. [2]Research in the field of data mining in semantic web analysis and data applied to various algorithms of data mining, such as data classification, association rule mining etc. From the above it can be seen that the present data are not stored in a single computer always. This research is proposed for methods to mine the data in E-commerce systems and improve the efficiency and scalability in relation database systems.

2. DENORMALIZATION

Database Denormalization is a well-known way of achieving performance improvements. Denormalization is the process to optimize the performance of a database by structuring data from an unstructured data or by grouping data. In some cases, Denormalization can actually increase the performance or scalability in relational database software. In the suggested system, the database administrator can define the Entity Relationship Model of the schema, and use

the queries that are built and mapped using Process Action Diagram language. Then the administrator can select the tables to join, and the system can automatically transforms the queries to match the new schema model. The system keeps a record of the mappings between the denormalized fields and the base fields from which they are derived and if the base fields were to be selected or updated, the new fields are returned or modified. The described system hides the denormalization process from the database users by converting the internal queries into structured or grouping. In our work, we need a similar method which can denormalize the database schema and rebuild the queries for a new schema.
Denormalize process has used to generate schema fee database.

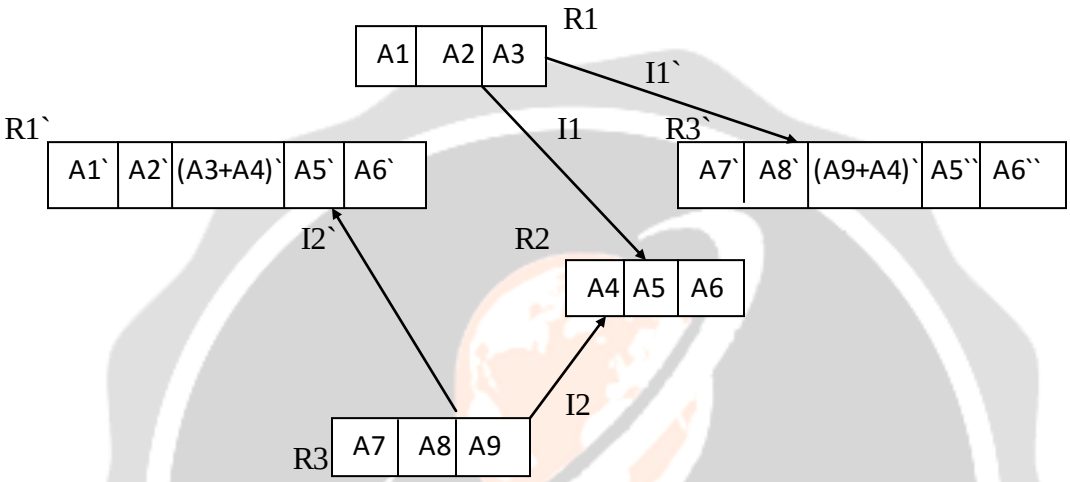


Fig -1: Sample denormalized schema without lost schemas.[1]

Although the above denormalization is valid, it is not enough for our purpose of generating all the possible data schemas. We should not delete the original two relations from the schema because with this deletion we are losing possibly important schema options. Consider that there is an additional relation R3 (A7; A8; A9), from which attribute A9 references A4. In this case, if the two relations are merged, and the original relations R1 and R2 are deleted , we cannot reach the schema containing relations R01 and R03 illustrated in Fig-1, which is gained by not removing the original relations. Note that R1` ≠ R3` .Thus, in the algorithm when generating all the viable schemas, the denormalizationstep should not delete the source relations.

3. ERM TO RDF

Below figure shows how an ERM model is transformed into a RDF model. Two resources, one for a customer instance and another for a product instance is required for the description of the below mentioned model. Relationship between customer and its related product is shown through the customer instance. An RDF model is about linking the different instances, whereas an ERM system is about linking entities and its relationships.

.

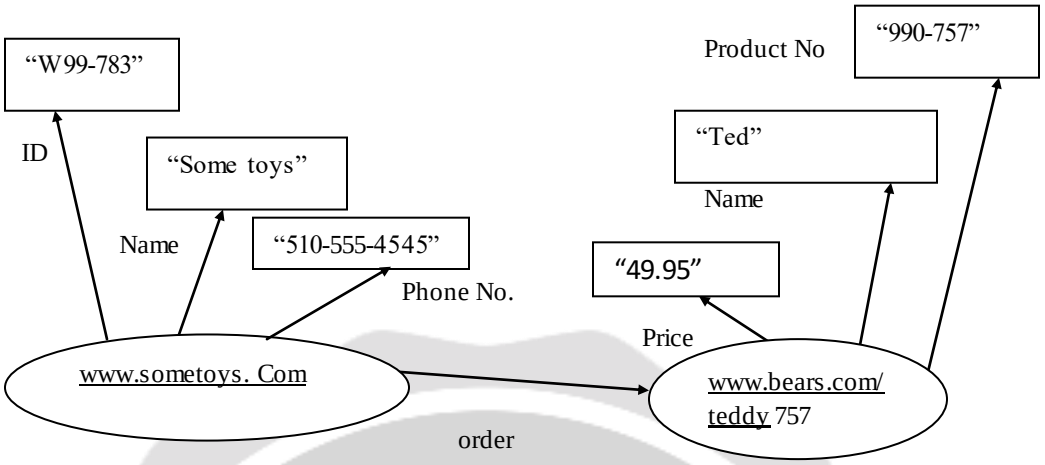


Fig -2: Convert ERM to RDF

Following code snippet shows RDF serialization of an Entity Relation as in Figure 2. [4]

```
<rdf:RDF>
<rdf:Description about="http://www.some-toys.com">
<sales:ID> W99-783 </sales:ID>
<sales:Name> Some Toys </sales:Name>
<sales:PhoneNo> 510-555-4545</sales:PhoneNo>
<sales:orders
rdf:resource="http://www.bears.com/teddy757"/>
</rdf:Description>
<rdf:Description about="http://www.bears.com/teddy757">
<bears:ProductNo> 900-757 </bears:ID>
<bears:Name> Ted </bears:Name>
<bears:Price> 49.95 </bears:Price>
</rdf:Description>
</rdf:RDF>
```

4. ONTOLOGY VOCABULARY

Ontology vocabulary is the main layer of semantic web architecture.Which consist of hierarchical distribution of important concept in a domain.,along with descriptions of the properties of each concept. Ontologies play a pivotal role in the semantic web by providing a source of shared and precisely defined terms that can be used in metadata. The recognition of the key role in ontologies are likely to play in the future of the web that has led to extension of web mark up languages like XML Schema, RDF and RDF Schema. The recognition of the limitations in mark up languages led to the development of new web ontology languages such as OWL.OWL is used when the information contained in documents needs to be processed by applications.It also used to explicitly represent the meaning of terms in vocabularies and the relationships between those terms. This representation of terms and their interrelationships is called ontology. OWL has more facilities for expressing meaning and semantics than XML, RDF, and RDF-S, and thus OWL goes beyond these languages in its ability to represent machine interpretable content on the Web.It has been designed to meet the requirements of RDF, RDFS, XMLSchema [4].

5.UNSTRUCTURED TO STRUCTURED DATA

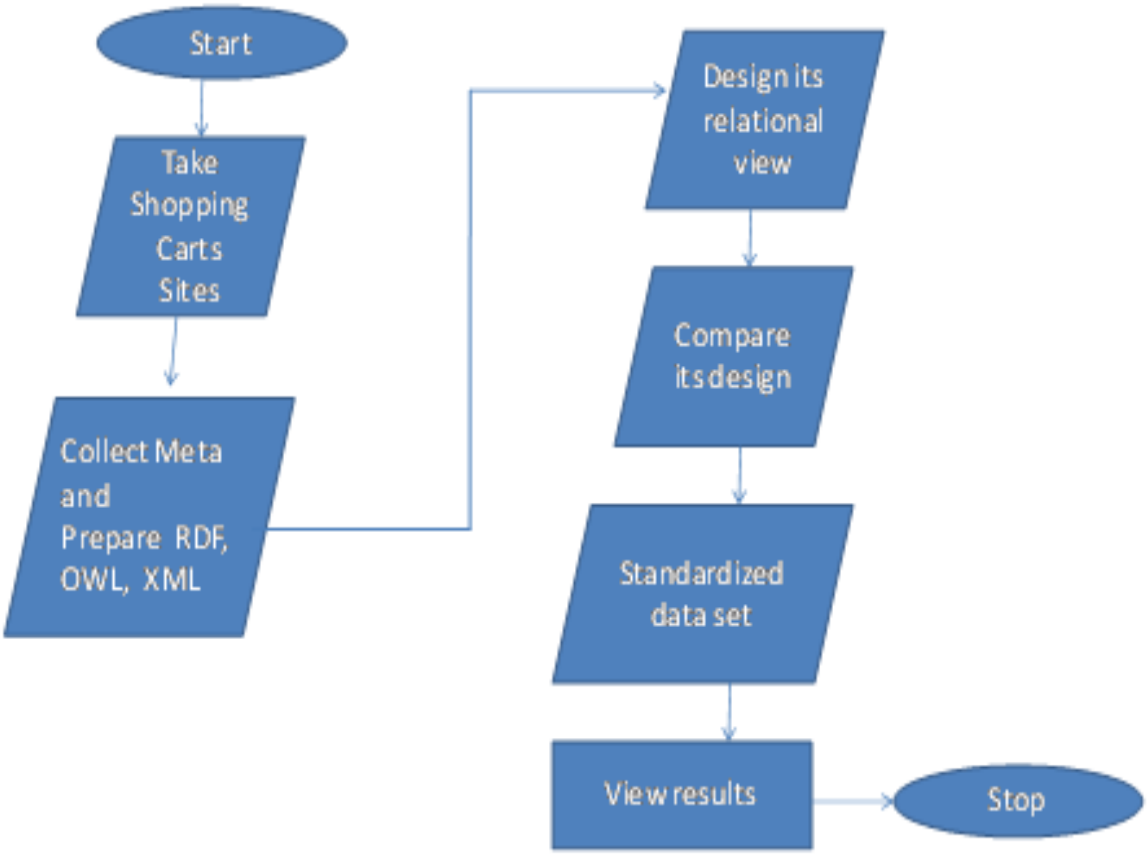
Unstructured data represent the largest and fastest growing basis of information accessible to businesses and governments. There are several methods available to find the structured data from unstructured data environment. An unstructured data communities such as semantic web, AI, IR, KDD, and Web to industrial users such as Microsoft, Google, and Yahoo. In Distributed data mining the unstructured data like voice mail, e-mail messages, still images, complicated reports, video and presentations. Most of the existing methods of unstructured to structured data contain some bias. So we are focusing on the structured data environment from unstructured data.

Table-1: Difference between unstructured and structured data[6]

	Unstructured data	**Structured data**
Technology	character and binary data	relational database tables
Transaction management	no transaction management and no concurrency	matured transaction management, various concurrency techniques
Version management	versioned as a whole	versioning over tuples, rows, tables, etc
Flexibility	very flexible, absence of schema	schema-dependent, rigorous schema
Scalability	very scalable	scaling database schema is different
Robustness	-	very robust, enhancements since 30 years

6. FLOW OF PROPOSED WORK

Here we describe the flow of proposed work as below:
Step: 1 First we will get 4-5 e-commerce sites.
Step: 2 we will take its metadata and other <head> tags.
Step: 3 we will convert it into RDF.
Step: 4 Based on the RDF, those values will be converted to relational designs (ERD).
Step: 5 Comparison of all sites ERD, we will study the standardization process.
Step: 6 Standardization process will be performed with De-normalization.
We will display the results based on the comparison of above process in table format.



7. CONCLUSIONS

This research aims to study about the retrieve logical schema from shopping cart database systems using RDF/XML format. We have introduced an algorithm that can automatically denormalize the initial data structure and also translate the predefined queries along with the schemas. From the database designing, an interesting research issue is what database structures are needed to support customization and personalization most effectively. [10] A lot of work is being done in studying different methods of database designs for E-commerce systems but all that work does not provide specific standard format. This paper gives an idea to retrieve logical schema from shopping cart database.

5. ACKNOWLEDGEMENT

I would like thank my Institution, all my faculty members in Department of Computer Science and my colleagues without whom this project would have been a distant reality. Last, but not the least, no words are enough to acknowledge constant support and sacrifices of my parents because of whom I am able to complete my work successfully.

6. REFERENCES

[1] Pascal Hitzler, Markus Krötzsch, Sebastian Rudolph, "Foundations of Semantic Web Technologies," Textbooks in Computing, Chapman and Hall/CRC Press, 2009.
[2] Victoria Nebot, Rafael Berlanga "Finding association rules in semantic web data," Knowledge-Based Systems, vol. 25, no. 1, pp. 55-62, 2012.
[3] Tamás Vajk, Péter Fehér, Krisztián Fekete and Hassan Charaf, "Denormalizing Data into Schema-free Databases", 4th IEEE International Conference on Cognitive Infocommunications December 2–5, 2013 , Budapest, Hungary.
[4] Karim Heidari, Serajodin Katebi, and Ali Reza Mahdavi Far, "New Methods for E- Commerce

Databases Designing in Semantic Web Systems (Modern Systems)", World Academy of Science, Engineering and Technology Vol:3 2014-03-27

[5] Nikitas N. Karanikolas and Michael Gr. Vassilakopoulos,"Database Design with Real- World Structures", Department of Informatics Technological Educational Institute of Athens, Greece

[6] Padmapriya.G, M.Hemalatha, "A Recent Survey on Unstructured Data to Structured Data in Distributed Data Mining", Int.J.Computer Technology & Applications,Vol 5 (2),338-344, March-April 2014

[7] K. Chomboon, N. Kaoungku, K. Kerdprasop, and N. Kerdprasop, "Data Mining in Semantic Web Data", International Journal of Computer Theory and Engineering, Vol. 6, No. 6, December 2014