Semantic Agreement Maintainance

Andreas Hadiyono, Prasetiyo,
Lintang Y. Banowosari and I Wayan Simri W.
Gunadarma University
andre_cun, prazethe_juga(yahoo.com)
lintang, iwayan(staff.gunadarma.ac.id)

Abstract:
The differences data sources can make problem in accessing information in different sources, especially when implemented in network model, for example P2P(Peer to Peer). This paper focused to solve problem for accessing information in different P2P sources especially in bridging query from user to peer with similar property or object. To solve that problem we use method of semantic agreement maintenance in implemented with semantic web. The goal of this paper is to reduce problem in accessing information with our offered method and implemented method to the web services.

Keywords : local scheme, ontology, P2P, query, semantics

1 Introduction

Many people has become accustomed to the Internet's rapid growth. One function of internet is for searching and sharing information. All existed information of internet kept by various data sources. Many source sometimes present information at different model databases, including highest level until lower level. Every level needs different kind, attribute and properties which can be saved in database. The differences data sources can make problem in accessing information in different sources, especially when implemented in network model, for example P2P(Peer to Peer). Recently, the computer science community has become accustomed to the Internet's continuing rapid growth, but even to such jaded observers the explosive increase in Peer-to-Peer (P2P) network usage has been astounding [5].

In this paper we focus to solve problem for accessing information in different P2P sources especially in bridging query from user to peer with similar property or object. To solve that problem we use method of semantic agreement maintenance in implemented with semantic web. The goal of this paper is to reduce problem in accessing information with our offered method and implemented method to the web services.

Here some motivations of our approach. We divide it for user and system. For user, source of information contains various models to represent their content. Problem occurs when they want to get information for different databases. We propose a new approach to solve that problem, so for future retrieval user can get more relevant information. For system, we hope we can bridges the differences between databases by maintenance semantic agreement. So we can minimize manual monitoring, which high failure and high cost. And finally we can deliver an automatic monitoring and improvement agreement idea.

Problems in developing semantic agreement maintenance method are : (1)To detect changes of data sources. Some modification can be founded in data source during data operations. Here, we compare each version of data source modification to (2)count how big changes of it by giving value to each operation (see General Overview for details). After we get the total value of its operations (border value), so we can (3)choose algorithm that we'll used for maintain data source.

The purpose of paper is to present a semantic agreement maintenance for solve problem accessing information in different data sources and give information about semantic web which can be solve the problem for accessing information and developing one modification approach of distance semantic theory to know change of local scheme or ontology so that can be used to conduct conservancy of agreement between a common ontology and of provider existing peer.

1.1 State of The Art

Semantic integration is an active research in several diciplines, such as databases, information-integration, and ontologies and to representation mapping ontology we can be used several tools, the one tools is PROMPT, the tools are extensions to the Protege ontology-development environmnet[4]. Semantic similarity relates to computing the between concept which are not lexicographically similar.Some of the most popular semantic similarity methods are implemented and evaluated using WordNet as the underlying reference ontology [2].

2 Approach

Semantics is the study of language meaning. In the computer science, semantics have meaning of program or function. Semantics is growing up become Semantic Web which development of World Wide Web through implant with semantic metadata [2]. Semantic conflict arise when two system do not use same interpretation to information. Simplest form of disagreement in interpreting information is homonym (using word which is equal to different meaning), and synonym (using word differ from is same meaning). In this case, semantics of information have to be considered by for the agenda of deciding how different information item correlate one with the other. Yaser [1]have divided schematic variety to in a few groups: Differing in

class like synonym, homonym, differing in class attribute, integrity constrain and method. Differing in attribute, like domain, unit, assess data type and default Differ in hierarchy, like class, attribute, generalizing storey;level and of aggregation.

One approach with semantic agreement maintenance is used four steps to detect and determine maintenance or not(see figure 1).

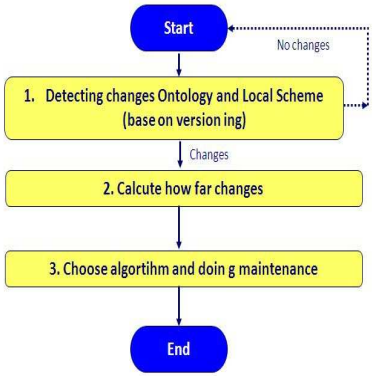

Figure 1: Steps to detect and determine maintenance or not

1. Detecting change of ontology and Local Scheme

   We can knew the changes of one ontology and local scheme with ontology versioning [6]. Ontology versioning is build and maintenance different version from ontology and providing access for those version. Nowadays, versioning sophisticated mechanism doesn't exist. Actually, ontology changes and the old version has been lost forever, because only the old version can be accessed. Sometimes, new version and old version from ontology has been archived. So if we changes property or another we must change the version manually.

2. Calculating how far the changes The changes of ontology can be found with one algorithm. The algorithm is PROMPTDIFF. The algorithm has been represented Noy and Klein [3]. PROMPTDIFF become plug-in in PROTEGE 2000 ontology-editing. With this algorithm we can compare ontology different version. And with PROMPTDIFF have average recall value 96% and average precision value 93

3. Choose algorithm for maintenance and Doing Maintenance

   We can choose algorithm maintenance if we have border value of ontology. The border value of ontology can be found if we do several experiment, so we can found average value of ontology. With this border value we can choose two algorithm for maintenance. The algorithm is simple and complex. Complex algorithm is algorithm which include label matching with Jiang&Cornath, internal matching. And the algorithm simple is a

part of complex algorithm, but simple algorithm only have one step i.e. label mathcing using Jiang & Cornath. Jiang & Cornath we choose for tools of label matching because many paper in semantic web suggest to use it.

3 Evaluation

Here we deliver some scenario to test the semantic agreement model by doing the modification to one or more scheme. First of all, we modify (add, delete, and change) the local scheme of peer for its class and property. The modify of local scheme gives a border value. By using a border value, we can take the PromptDiff approach (from PROMPT TAB tool) to find out algorithm that we'll use. To test the algorithm that we'll use, we can make an agreement from two algorithms. And finally, we count the Recall Value (The proportion of relevant document, beyond all exist relevant document), Precision Value (The proportion of retrieved and relevant document for all retrieved document), and F-Measure (harmonic average weight from precision and recall value.

From those scenario, we got the following result :

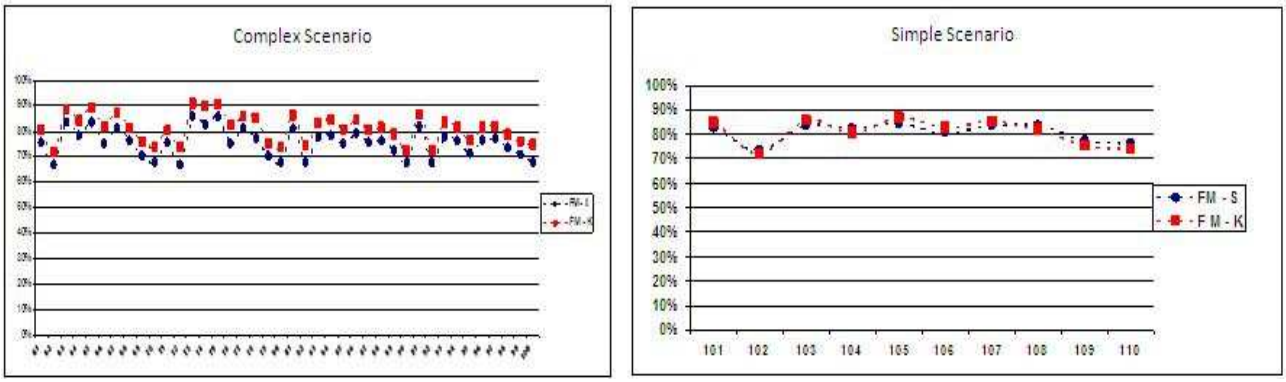

Figure 2: The Result for Complex and Simple Algorithm

1. If complex algorithm used, the F-measure always above the simple algorithm.

2. If simple algorithm used, the F-measure relatively equals with complex algorithm. Indeed, the F-measure of simple algorithm better than complex algorithm (2 percent)

4 Conclusion

- More modification of local scheme and ontology makes F-Measure for semantic agreement become worst.

- Choosing a proper algorithm is the most important thing for agreement maintenance respect to cost computation and F-Measure. The using of complex algorithm for case

that should be use a simple algorithm makes higher complexity with relatively small F-measure distinction. Otherwise, the F-measure for complex algorithm higher than simple algorithm with higher complexity.

4.1 Future Work

- It is possible to expand the evaluation domain. Ideally, the semantic agreement maintenance approach is generic for broader domain. So, the further evaluation for other domains is needed.

- There are also possibilities to divide local scheme and common ontology to some portion and implement the different algorithm for them.

References

[1] Y Bishr. Semantic aspects of interoperable gis. phd thesis, wageningen agricultural university, netherland. 1997.

[2] Paraskevi Raftopoulou Euripides G.M Petrakis Evangelos E. Milos Gianis Varelas, Epimendisi Voutsakis. Semantic similarity methods in wordnet and their applocation to information retrieval on the web.

[3] M. Klein and Noy. A component based framework for ontology evolution. in in proc. of ijcai 2003 workshop on ontologies and distributed systems. 2003.

[4] Nataly F. Noy. Semantic integration : A survey of ontology based approaches.

[5] Kai li Scott Shenker Qin Liv, Pei Cao Edith Cohen. Search and replication in unstructured peer-to-peer networks. 2001.

[6] B. Yildiz. Ontology evolution and versionng the state fo the art. vienna university of technology, institute of softwrae techol- ogy & interactive systems (isis), http://publik.tuwien.ac.at/_les/pub inf_4603.pdf. 2006.