International Journal of Innovative Research in Information Security (IJIRIS)
Volume 1 Issue 6 (December 2014)
ISSN: 2349-7017(O)
ISSN: 2349-7009(P)
www.ijiris.com

Generic Algorithm based Data Retrieval Technique in Data Mining

T.Sumithgowri
Master of Philosophy,Research Scholar
Tamilnadu Open University,Saidapet,Chennai

N. Sivashanmugam,
Assistant Professor,
Tamilnadu Open University,Saidapet,Chennai

Abstract-- This system Hybrid extraction of robust model (GA), a dynamic XAML based mechanism for the adaptive management and reuse of e-learning resources in a distributed environment like the Web. This proposed system argues that to achieve the on-demand semantic-based resource management for Web-based e-learning, one should go beyond using domain ontology's statically. So the propose XAML based matching process involves semantic mapping has done on both the open dataset and closed dataset mechanism to integrate e-learning databases by using ontology semantics. It defines context-specific portions from the whole ontology as optimized data and proposes an XAML based resource reuse approach by using an evolution algorithm. It explains the context aware based evolution algorithm for dynamic e-learning resource reuse in detail. This system is going to conduct a simulation experiment and evaluate the proposed approach with a xaml based e-learning scenario. The proposed approach for matching process in web cluster databases from different database servers can be easily integrated and deliver highly dimensional e-learning resource management and reuse is far from being mature. However, e-learning is also a widely open research area, and there is still much room for improvement on the method. This research mechanism includes 1) improving the proposed evolution approach by making use of and comparing different evolutionary algorithms, 2) applying the proposed approach to support more applications, and 3) extending to the situation with multiple e-learning systems or services.

Keywords -- XAML, Dataset, Data Management

1.INTRODUCTION

In many application domains (e.g., medicine or biology), comprehensive schemas resulting from collaborative initiatives are made available. This proposed system argues that to achieve the on-demand semantic-based resource management for Web-based e-learning, one should go beyond using domain ontology's statically. So the propose XAML based matching process involves semantic mapping has done on both the open dataset and closed dataset mechanism to integrate e-learning databases by using ontology semantics.

It defines context-specific portions from the whole ontology as optimized data and an XAML based resource reuse approach by using an evolution algorithm. Such well established schemas are often associated with reliable data that have been carefully collected, cleansed, and verified, thus providing reference ontology-based data management systems (DMSs) in different application domains. A good practice is therefore to build on the efforts made to design reference DMSs whenever we have to develop our own DMS with specific needs.

A way to do this is to extract from the reference DMS the piece of schema relevant to our application needs, possibly to personalize it with extra constraints w.r.t. our application under construction, and then to manage our own data set using the resulting schema. Recent work in description logics provides different solutions to achieve such a reuse of a reference ontology-based DMS. Indeed, modern ontological languages like the W3C recommendations RDFS, OWL, and OWL2 are actually XML based syntactic variants of well known DLs. All those solutions consist in extracting a module from an existing ontological schema such that all the constraints concerning the relations of interest for the application under construction are captured in the module.

Existing definitions of modules in the literature basically in this system, revisit the reuse of reference ontology based DMS in order to build a new DMS with specific needs. It goes one step further by not only considering the design of a module based DMS (i.e., how to extract a module from an ontological schema) also study how a module based DMS can benefit from the reference DMS to enhance its own data management skills. Our contribution is to introduce and study novel properties of robustness for modules that provide means for checking easily that a robust module based DMS evolves safely w.r.t. both the schema and the data of the reference DMS.

From a module robust to consistency checking, for any data update in a corresponding module-based DMS, we show how to query the reference DMS for checking whether the local update does not bring any inconsistency with the data and the constraints of the reference DMS. from a module robust to query answering, for any query asked to module-based DMS, It shows how to query the reference DMS for obtaining additional answers by also exploiting the data stored in the reference DMS.

II RELATED WORK

Conventional Data Management Strategies

In conventional information management principles, the stored records are normally identified by sets of key words or index terms, and requests for information are expressed by using Boolean combinations of index terms. The retrieval strategy is normally based on an auxiliary inverted-term index that lists the corresponding set of document references for each allowable index term. The Boolean retrieval system is designed to retrieve all stored records exhibiting the precise combination of key words included in the query: when two query terms are related by an and connective, both terms must be present in order to retrieve a particular stored record; when an or connective is used, at least one of the query terms must be present to retrieve a particular item.

Content-based Methods

In content-based recommendation methods, the utility $u(c, s)$ of item s for user c is estimated based on the utilities $u(c, si)$ assigned by user c to items $si \in S$ that are "similar" to item s. For example, in a movie recommendation application, in order to recommend movies to user c, the content-based recommender system tries to understand the commonalities among the movies user c has rated highly in the past (specific actors, directors, genres, subject matter, etc.). Then, only the movies that have a high degree of similarity to whatever user's preferences are would get recommended. The content-based approach to recommendation has its roots in information retrieval and information filtering research.

Limited content analysis

Content-based techniques are limited by the features that are explicitly associated with the objects that these systems recommend. Therefore, in order to have a sufficient set of features, the content must either be in a form that can be parsed automatically by a computer (e.g., text), or the features should be assigned to items manually. While information retrieval techniques work well in extracting features from text documents, some other domains have an inherent problem with automatic feature extraction. For example, automatic feature extraction methods are much harder to apply to the multimedia data, e.g., graphical images, audio and video streams. Moreover, it is often not practical to assign attributes by hand due to limitations of resource Another problem with limited content analysis is that, if two different items are represented by the same set of features, they are indistinguishable.

III PROPOSED DESIGN

This proposed system Hybrid extraction of robust model (GA) is very efficient and reuse of e-learning resources in a distributed environment like the Web for better result. This proposed system argues that to achieve the on-demand semantic-based resource management for Web-based e-learning, one should go beyond using domain ontology's statically.This research mechanism includes 1) improving the proposed evolution approach by making use of and comparing different evolutionary algorithms, 2) applying the proposed approach to support more applications, and 3) extending to the situation with multiple e-learning systems or services.

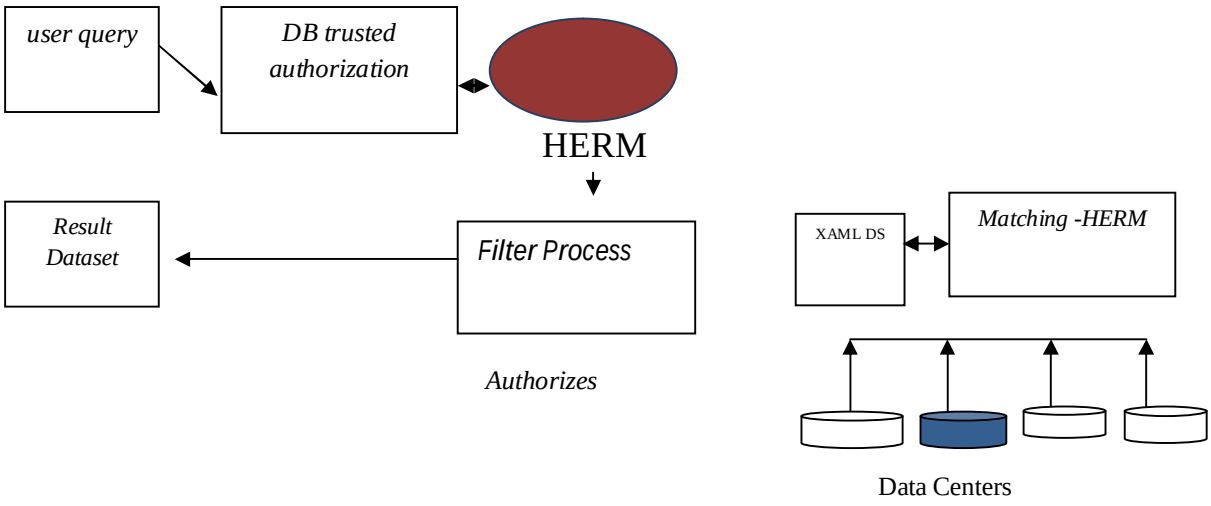

Fig 1: System Architecture of GA

ISSN: 2349-7017(O)
ISSN: 2349-7009(P)
www.ijiris.com

GA Algorithm

The extraction and matching process has been implemented step by step with initial requirements.

```
1.  String[][] fillCLabMatrix(Tree of Nodes source,target);
2.  String[][]cLabsMatrix;
3.  String[] matchers;
4.  int i,j;
5.  matchers=getMatchers();
6.  for each sourceAtomicConceptOfLabel in source
7.  i=getACoLID(sourceAtomicConceptOfLabel);
8.  for each targetAtomicConceptOfLabel in target
9.  j= getACoLID(targetAtomicConceptOfLabel);
10. cLabsMatrix[i][j]=getRelation(matchers,
11. sourceAtomicConceptOfLabel,targetAtomicConceptOfLabel);
12. return cLabsMatrix
13. String getRelation(String[] matchers,
14. AtomicConceptOfLabel source, target)
15. String matcher;
16. String relation Idk ;
17. int i=0;
18. while ((i<sizeof(matchers))&&(relation== Idk))
19. matcher= matchers[i];
20. relation=executeMatcher(matcher,source,target);
21. i++;
22. return relation;
```

The major work of this module is to perform the data matching by getting the user query the data request can be done in optimized fashion, that is, the data request can come from the data agent of any type of user query. Data Agent can request for explicit request with a specific condition, where in this application, it is modeled as patient information, here the agent can request for specific name. Agent can also request for sample request, that is, they can request for specific number of clustering by sequence. These requests are raised by the data agents during runtime and the open datasets are also allocated uniquely for each query during runtime. This Dataset accompanied by necessary additions depending upon the requirement.Opendata set used to read,write, Matching the datas from more than one different database services.

The dataset information is subdivided in several sections, "content", "Meta information", "Exchanges", and "Allocations" (in multi datasets).client queries are move to the respective section of the dataset. The details of the dataset are structured according to the stored in xaml format. The process details consist of two main sections: meta information and the actual flow data . The items may vary, depending on the type of dataset. The dataset is designed for use in research related to data mining in [disease] robust . Datasets are objects that contain data tables where you can temporarily store the data for use in your application. If your application requires working with data, you can load the data into a dataset, which provides your application with a local in-memory cache of the data to work with. You can work with the data in a dataset even if your application becomes disconnected from the database. The dataset maintains information about changes to its data so updates can be tracked and sent back to the database when your application becomes reconnected. A dataset contains no actual data by default. Filling a dataset with data actually refers to loading data into the individual Data Table objects that make up the dataset. You fill the data tables by executing Table Adapter queries, or executing data adapter (for example, Sql Data Adapter) commands. When you fill a dataset with data, various events are raised, constraints are checked, and so on. For more information about loading data into a dataset.

IV CONCLUSION

The proposed technique GA have given better results for solving the problem of safe personalization of modules built from an existing reference DMS. This raises new issues to check easily that a module-based DMS evolves independently but coherently w.r.t. the reference DMS from which it has been built. it have introduced two notions of module robustness that make possible to build locally the relevant queries to ask to the reference database in order to check global consistency (possibly upon each update), and to obtain global answers for local queries. It have provided polynomial time algorithms that extract minimal and robust modules from a reference ontological schema expressed as a data query. It proposes an alternative to our result about global query answering, which applies under the severe constraints that the data set of the reference DMS has to be modified. Contrarily to recent works in

distributed databases, data replication can be avoided while guaranteeing global consistency. This approach is a good tradeoff between the NoSQL approaches and the SQL approaches for managing distributed data stores. While most of the NoSQL approaches are schema less, our approach makes possible to handle useful schema constraints. It provides efficient means to check global consistency, a stronger property than eventual consistency that is prevalent in distributed data stores. On the other hand, and are more flexible than the SQL approaches since global consistency is checked periodically and not at each update of the reference DMS.

V. FUTURE ENHANCEMENT

In future this technique GA will be planned to evaluate in wide range of distributed data base servers in particular to compare the size of the modules extracted by our algorithm to the results provided by existing systems. It also plan to apply our algorithms to the real use case of the MyCorporisFabrica DMS, mentioned in the introduction, which has been developed manually as a personalization of the (reference) Foundational Model of Anatomy DMS. Finally, this system plan to extended approach to distributed module-based DMSs, where answering queries combines knowledge of several modules associated with possibly several reference range DMSs.

REFERENCES

[1] The Description Logic Handbook: Theory, Implementation, and Applications, F. Baader, D. Calvanese, D. McGuinness, D. Nardi, P.F. Patel-Schneider, eds. Cambridge Univ. Press, 2003.
[2] Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization, H. Stuckenschmidt, C. Parent, S. Spaccapietra, eds. Springer, 2009.
[3] S. Ghilardi, C. Lutz, and F. Wolter, "Did I Damage My Ontology? A Case for Conservative Extensions in Description Logics," Proc. 10th Int'l Conf. Principles of Knowledge Representation and Reasoning (KR), 2006.
[4] R. Kontchakov, L. Pulina, U. Sattler, T. Schneider, P. Selmer, F. Wolter, and M. Zakharyaschev, Minimal Module Extraction from DL-Lite Ontologies Using QBF Solvers," Proc. 21st Int'l Joint Conf.Artificial Intelligence (IJCAI), 2009.
[5] Z. Wang, K. Wang, R.W. Topor, and J.Z. Pan, "Forgetting concepts in DL-Lite," Proc. Fifth European Semantic Web Conf. Semantic Web: Research and Applications (ESWC), 2008.
[6] B. Konev, D. Walther, and F. Wolter, "Forgetting and Uniform Interpolation in Extensions of the Description Logic EL," Proc. 22nd Int'l Workshop Description Logics, 2009.
[7] B. Konev, C. Lutz, D. Walther, and F. Wolter, "Semantic Modularity and Module Extraction in Description Logics," Proc. 18th European Conf. Artificial Intelligence (ECAI), 2008.
[8] B. Konev, D. Walther, and F. Wolter, "Forgetting and Uniform Interpolation in Large-Scale escription Logic Terminologies," Proc. 21st Int'l Joint Conf. Artifical intelligence (IJCAI), 2009.
[9] B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler, "Just the Right Amount: Extracting Modules from Ontologies," Proc. 16th Int'l Conf. World Wide Web (WWW), 2007.
[10] K. Wang, Z. Wang, R.W. Topor, J.Z. Pan, and G. Antoniou, "Concept and Role Forgetting in ALC Ontologies," Proc. Eighth Int'l Semantic Web Conf. (ISWC), 2009.
[11] D. Calvanese, G.D. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati, "Tractable Reasoning and Efficient Query Answering in Description Logics: The DL-Lite Family," J. Automated Reasoning,vol. 39, no. 3, pp. 385-429, 2007.
[12] O. Palombi, G. Bousquet, D. Jospin, S. Hassan, L. Reve´ret, and F. Faure, "My Corporis Fabrica: A Unified Ontological, Geometrical and Mechanical View of Human Anatomy," Proc. Second Workshop 3D Physiological Human (3DPH), 2009.
[13] S. Abiteboul, R. Hull, and V. Vianu, Foundations of Databases. Addison-Wesley, 1995.
[14] M.Y. Vardi, "The Complexity of Relational Query Languages," Proc. 14th Ann. ACM Symp. Theory of Computing (STOC), 1982.
[15] A. Cali, G. Gottlob, and T. Lukasiewicz, "Datalog+-: A Unified Approach to Ontologies and Integrity Constraints," Proc. Int'l Conf. Database Theory (ICDT), 2009.
[16] R. Cattell, "Scalable Sql and Nosql Data Stores," SIGMOD Record, vol. 39, no. 4, pp. 12-27, 2010.
[17] B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler, "Extracting Modules from Ontologies: A Logic-Based Approach," Proc. Third Int'l Workshop OWL Experiences and Directions (OWLED), 2007.
[18] B. Cuenca Grau, I. Horrocks, Y. Kazakov, and U. Sattler, "Modular Reuse of Ontologies: Theory and Practice," J. Artificial Intelligence Research, vol. 31, pp. 273-318, 2008.

