ISSN No. 0976-5697

Volume 5, No. 7, September-October 2014

International Journal of Advanced Research in Computer Science

TECHNICLA NOTE

Available Online at www.ijarcs.info

A Survey on Retrieval Strategy for Case-Based Reasoning using Modified - USIMSCAR for Object Oriented Cases

N.Adhithyaa
PG Scholar
Department of Information Technology
SNSCE, Coimbatore, India

P.Sumathi
Associate Professor
Department of Information Technology
SNSCE, Coimbatore, India

Abstract: Case based reasoning is the problem solving technique which is fully based on the solutions of similar past problems. In CBR system, retrieval is an important part since the overall performance is based on it. Aim of this paper is to retrieve the combined set of each cases and rules relevant to the target drawback, using the combined knowledge of similarity and association. This paper proposes the retrieval strategy in CBR System for object oriented cases using modified - USIMSCAR algorithm. It also argues the performance of the generated association rules using SCAR algorithm in CBR systems. Association rules are created using ARM techniques and OO-datasets are used for evaluation. Modified - USIMSCAR algorithm is applied in medical domain.

Keywords: CBR, Association knowledge (AK), Association rule mining (ARM), Case-Based Reasoning, Multilevel association rules, Modified-USIMSCAR(unified similarity and soft- matching class association rule)

I. INTRODUCTION

The fundamental premise of the case-based reasoning is that the expertise within the variety of past cases is leveraged to resolve new issues [1]. An individual expertise is termed a case, and its assortment is keep within the case based mostly. Typically, every case is delineated by the matter description and also the corresponding answer description. Among the four typical phases in CBR area unit the retrieval, reuse, revise, and retain, the retrieval may be a key to CBR, since the success of CBR system is heavily dependent on the performance of retrieval [2]. Thus, process SK continues to be terribly difficult, long, and onerous to practices. Second the similarity live is incredibly static so the definition is extremely potential to be applied to any or all target issues systematically.

In this paper, we propose the association analysis of cases for the item orientated cases. Association data represents powerfully evident, attention-grabbing relationships between glorious downside options and solutions shared by an outsized variety of cases.

For USIMSCAR to alter the retrieval method with object oriented cases, addressed the two issues: the primary is a way to formalize similarity encoding similarity data and also the second is a way to generate soft-matching class association rules (scars) encoding association data.

II. RELATED WORK

A. *Similarity Knowledge*

In the case based reasoning context, similarity information encoded via live computing between the target downside Q and cases. The higher the similarity between Q and case C is that the additional helpful case C for down side Q. It is the native– world principle that decomposes the similarity measure by the local similarities for individual attributes of cases and therefore the world similarity aggregating these similarities [3]. A global similarity operate are often willy-nilly complicated, however easy functions square measure typically used like weighted average aggregation [4].

B. *Background of AK*

AK aims to represent patently attention-grabbing relationships shared by the big variety of relevant hold on cases, using the combination of various DM techniques. These are the ARM [5], class ARM [6], and soft-matching ARM (SARM) [7].

1) ARM: ARM aims to mine certain relationships, associations, in the transaction database [8]. It focuses on discovering the set of extremely co-occurred options shared by sizable amount of records within the database. In the Case based Reasoning context, ARM will be accustomed discover interesting relationships from the given case base. A transaction and the item can be seen as the case and an attribute–value pair, respectively. Apriori [9] is one of the traditional algorithms for the ARM. Interestingness measures are helpful to judge the standard and rank the big range of ARs extracted [10].Generating association rules that have bigger support as compared to user outlined minimum threshold and confidence bigger than user outlined minimum confidence is that the main downside of Association Rule Mining.

2) Class ARM: Class association rules (CARs) [11] are the special set of ARs whose consequents are restricted to the one target variable. In the Case based Reasoning context, the CAR is seen as associate AR whose results holds the item fashioned, because the combine of the answer attribute and also the worth of it.

3) *SARM:* Consider a rule X → Y. A limitation of the normal ARM algorithms (e.g., Apriori) is that itemsets X and Y are discovered by the equality relation. when addressing items the same as one another, these algorithms might perform poorly. For example, in the supermarket sales database, Apriori cannot find rules like 80% of the customers who buy products similar to milk (e.g., cheese) and products similar to eggs (e.g., mayonnaise) also buy bread. To address this issue, the soft-matching criterion was planned [12], where the antecedents and consequents of ARs found by similarity assessment.Using this criterion, the problem of SARM is to search out all rules of the shape X → Y, wherever the soft support and soft confidence of every rule aren't but minsupp and minconf, severally. The definitions of support and soft confidence are generalized by victimisation support and confidence. This generalization is finished by permitting things to match, as long as their similarity exceeds the user-specified minimum similarity (minsim).

III. EXISTING SYSTEM

The case is structured into the two main elements. The first half is that the downside part that contains the outline, characterizing the past problem. The second half is that there solution half that contains the outline of the appropriate resolution for the delineated problem.

To represent the cases formally, several CMB systems typically adopt well-known data illustration attribute-value pairs. In this method two algorithms is generated that soft-matching class association rule (SCAR) and unified knowledge of similarity and soft- matching class association rule (USIMSCAR). In the SCAR rule, generate the soft-matching class association rules for the frequent item sets. In the second algorithm, we combine the similarity knowledge and the association knowledge for improve the performance of the similarity Based Retrieval (SBR).

IV. PROPOSED SYSTEM

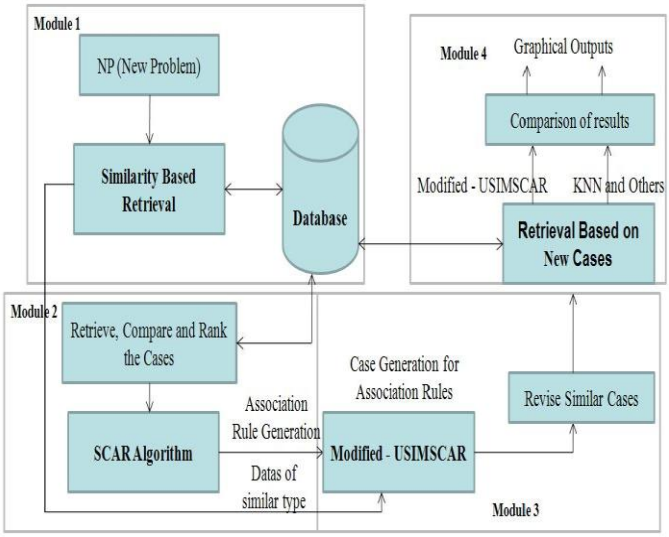
This section explains the proposed system. Here, same method for the Case-Based Reasoning (CBR), however use the structural illustration for the case base that's the object oriented representation. The object oriented representation represents each case at multiple levels of abstraction. The object oriented representation is the straight forward extension of the attribute- worth combine illustration permits for the outline of the cases with the complicated object adjusted structure[13].

To represent the Object oriented cases with the Modified - USIMSCAR, need to address the following two issues: first is how to formalize similarity measures encryption similarity data and second is how to generate SCARs encryption association data. The similarity measure for the item headed cases has got to be able to adequately figure the similarity between the same- level cases or the various level cases.

The generation of the SCARs from the item orienting cases primarily needs the mechanism that discovers frequent itemsets from the cases at the various levels. Used the Support- Count and Bit-from multilevel (SC-BF) Algorithm, using these algorithmic rule finding the frequent itemsets at the highest most level then increasingly deepening the mining method into their frequent descendants at lower conception levels. By integrating SC-BF algorithmic rule and the soft-matching criterion, generate the SCARs from the object oriented cases.

Mine multilevel association rules efficiently using concept hierarchies, and the soft-matching criteria. Object oriented algorithm defines sequence of mappings from the set of low-level ideas to higher-level [14]. Using the concept hierarchies, 1st retrieve the frequent itemsets from the case base at a similar level or totally different level so combined the soft-matching criteria to come up with the soft-matching class association rules.

4.1SystemArchitecture



V. RESULTS AND ANALYSIS FOR SBR(MODULE 1)

This section provides the comparisons between the existing method and the proposed method for Module 1(SBR) in term of the accuracy and the time complexity. For medical evaluation database, take the total time for retrieve frequent itemsets from the object oriented case and generating the soft-matching rules from the frequent itemsets and finally generate the association rule, is less compared to the existing method. Accuracy is also improved in the object oriented cases. Below figure show the Accuracy between the existing method and proposed method.

TABLE I
Attribute Vs Accuracy (%) for Medical Dataset
Medical Dataset

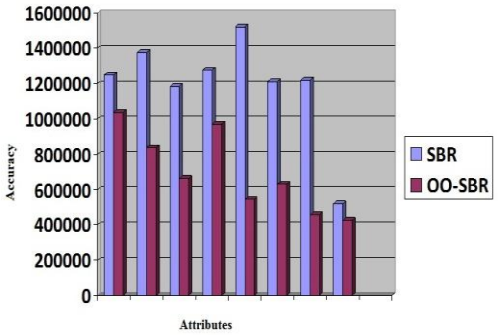
Attributes	Existing Algorithm(SBR)	Modified Method (Object oriented USIMSCAR)
1	15	20
2	15	23
3	24	33
4	24	30
5	30	38
6	30	37
7	60	65
8	24	29

Table II Contains the information about execution time for running the existing method and the proposed method. Proposed method take less time to run the algorithm for the number of attributes enter by the user.

TABLE II
Attributes Vs Time for Medical Dataset

Attributes	Existing Algorithm(SBR)	Modified Method (Object oriented SBR)
1	1248426	1035337
2	1376311	0837051
3	1182619	0661093
4	1275880	0970096
5	1518538	0543652
6	1211486	0627721
7	1218748	0457310
8	0518805	0425308

Fig. 2. Execution Time for Medica Dataset



5.1 Advantages Of Proposed System

Takes less memory to store the entire data because the data was grouped at the branch level and reduces the execution time. It also improves its accuracy.

VI. CONCLUSIONS AND FUTURE WORK

This paper present the SBR in case based reasoning for the object oriented structure, in this process first retrieve frequent item sets from the object oriented case using the SC-BF algorithm after that applied Soft-matching criteria for generating the SCARs. Second combined the Association knowledge and the similarity knowledge for the improvement of the similarity based retrieval. Advantages of using the SC-BF algorithm is the less memory to store the entire data because the data was grouped at the branch level and reduce the execution time. It improves its accuracy. So it improves the total performance of the case-based reasoning.

As future work, implementation of the modified-USIMSCAR can be processed.

VII. REFERENCES

[1] R. Lopez De Mantaras, D. McSherry, D. Bridge, D. Leake, B. Smyth, S. Craw, B. Faltings, M. L. Maher, M. T. Cox, K. Forbus, M. Keane, A. Aamodt, and I. Watson, "Retrieval, reuse, revise and retention in CBR," *Knowledge. Eng. Rev.*, vol. 20, no. 3, pp. 215–240, 2005.

[2] Y. Guo, J. Hu, and Y. Peng, "Research on CBR system based on mining," *Appl. Soft Comput.*, vol. 11, no 5014, 2011.

[3] H. Ahn and K.-J. Kim, "Global optimization of case-based reasoning or breast cytology diagnosis," Expert Syst. Appl. , vol. 36, no. 1, pp. 724–734, 2009.

[4] B. Pandey and R. Mishra, "Case-based reasoning and data mining integrated method for the diagnosis of some neuromuscular disease," *Int. J. Med. Eng. Informat.*, vol. 3, no. 1, pp. 1–15, 2011.

[5] Y.-B. Kang, A. Zaslavsky, S. Krishnaswamy, and C. Bartolini, "A knowledge-rich similarity measure for improving IT incident resolution process," in *Proc. ACM Symp. Appl. Comput.*, 2010, pp. 1781–1788.

[6] F. Lorenzi and F. Ricci, "Case-based recommendersystems: A unifying view," in *Intelligent Techniquesfor WebPersonalization*, vol. 3169. Berlin, Germany: Springer, 2005,pp. 89–113.

[7] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues,methodological variations, and system approaches," *AICommun.*, vol. 7,pp. 39–59, Mar. 1994.

[8] G. R. Beddoe and S. Petrovic, "Selecting and weighting features using a genetic algorithm in a case-based reasoning approach to personnel rostering," *Eur. J. Oper. Res.*, vol. 175, no. 2, pp. 649–671, 2006.

[9] K. Bradley and B. Smyth, "Personalized information ordering: A case study in online recruitment," *Knowl.-Based Syst.*, vol. 16, nos. 5–6, pp. 269–275, 2003.

[10] C. M. Vong, P. K. Wong, and W. F. Ip, "Case-based classification system with clustering for automotive engine spark ignition diagnosis," in *Proc. 9th Int. Conf. Comput. Inf.Sci.*, Aug. 2010, pp. 17–22.

[11] F. Azuaje, W. Dubitzky, N. Black, and K. Adamson, Discovering relevance knowledge in data: A growing cell structures approach," *IEEE Trans. Syst., Man, Cybern. B,Cybern.*, vol. 30, no. 3, pp. 448–460, Jun.2000.

[12] Z. Y. Zhuang, L. Churilov, F. Burstein, and K. Sikaris, "Combining data mining and CBR for intelligent decision support for pathology ordering by general practitioners," *Eur. J.Oper. Res.*, vol. 195, no. 3, pp. 662–675, 2009.

[13] L.Tong and D.Wu," Research on optimization of case-based reasoning system, " 3rd international conference on control, automation and system engineering ,2013.

[14] A. Stahl, "Learning of knowledge-intensive similarity measures in casebased reasoning," Ph.D. dissertation, Artificial Intelligence nowledge- Based Systems Research Group, Tech. Univ. Kaiserslautern, Kaiserslautern,Germany, 2003.