



Neighborhood based Framework Active Learning

Dr. K.P.Kaliyamurthi¹, K.G.S. Venkatesan², S. Sriram³, N.Vijay⁴, Richard Solomon⁵

Professor & Head, Dept. of C.S.E., Bharath University, Chennai, India¹

Associate Professor, Dept. of C.S.E, Bharath University, Chennai, Tamil Nadu, India²

Assistant Professor, Dept. of C.S.E, Bharath University, Chennai, Tamil Nadu, India³

Dept of C.S.E, Bharath University, Chennai, Tamil Nadu, India⁴

Dept of C.S.E, Bharath University, Chennai, Tamil Nadu, India⁵

ABSTRACT: Our proposed approach tries to decrease the impact of SLA violation through a user discrimination approach and avoid entering the complex procedure of negotiation. There will be no chance to get the unknown result or loss of result. We can measure the SLA violation in long term measuring. For simulations, we considered five different user types and five classes of virtual machines with different reliability level labels. A learning automaton is an adaptive decision-making unit located in a stochastic environment and the action selection is based on a specific Probability distribution which is updated according to environment the automaton receives by employing a particular action which hence increases the efficiency by reducing the maximum cost and by improving Quality of Service (QoS).

KEYWORDS: QoS(Quality of Service), SLA(Service Level Agreement)

I. INTRODUCTION

In this project User satisfaction as a significant antecedent to user loyalty has been highlighted by many researchers in market based literatures. SLA violation as a significant factor can decrease user's agreement level. The amount of this reduce depends on user's characteristics. So we have to present a new approach to reduce the impact of SLA violations on user's satisfaction level, but not by decreasing number of SLA violations. Instead, we try to use characteristics of users to decrease the impact of SLA violations on user satisfaction level [1]. But some of them are unknown to the service provider and selfish users are not interested to reveal them truly. Most of the works of literature ignore considering such characteristics and treat others just based on SLA parameters. So, two users with dissimilar characteristics but similar SLAs have equal significance for the service provider.

using pair-wise queries to make informative points and it may take multiple queries to resolve the uncertainty about a data point. This system as uncertainty based sampling for supervised learning; an active learner queries the instance about which the label uncertainty is maximized [9]. The probability making an instance learned clusters. Here to select an instance such that knowing its neighborhood will allow us to gain maximal information about the underlying clustering data. In particular instance focusing only on its relationship with the labeled instances.

Clearly the user based approach tries to satisfy more risk, unfavourable users pay a lesser amount of penalty, while this is not the case for random approach, the total number of paid penalties in random approach is greater than user based approach. This simulation contains user types with different Willingness To Pay (WTP), but similar risk aversion [5]. The average WTP of the users which their requests are assigned to virtual machines from different classes using the user based approach. Comparing the results, we can see the requests from users with higher WTP are assigned to virtual machines with higher reliability. This helps service provider to satisfy the users with higher WTP and improve its profitability. Supply and dependable management of SLA agreements is of principal importance for both Cloud providers and consumers [6].



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

II. RELATED WORK

In this [1], we propose a discriminative batch mode active learning approach that formulates the instance selection task as a continuous optimization problem over auxiliary instance selection variables. [2] present a framework for batch mode active learning" that applies the Fisher information matrix to select a number of informative examples simultaneously. The main computational task is how to efficiently identify the subset of unlabeled examples that can result in the largest reduction in the Fisher information. [3] we focus on the problem of performing spectral clustering under an economical constraint, where there is a boundary on the total number of entries which can be queried from the parallel matrix.

We propose two main algorithms for this problem, allow tradeoff among computational efficiency and actual performance, and are also important for the problematic of speeding up standard spectral clustering [4] CorePhrase works by first constructing a list of candidate key phrases for each cluster, scoring each candidate key phrase according to its features, ranking the key phrases by score, and finally selecting a number of the top ranking key phrases for output. [5] We present six Impact Challenges to explicitly focus the field's energy and attention, and we then converse existing difficulties which must be addressed.

We aim to inspire on-going talk and focus on ML that matters. [6] Hierarchical clustering initializes a cluster system as a set of singleton clusters (agglomerative case) or a single cluster of all points (divisive case) and proceeds iteratively with merging or splitting of the most appropriate cluster until the stopping criterion is achieved [7] they used MALLET tool for searching the topic of documents [8] Kernel mode semi supervised algorithm is used. [9] investigate the power of these citation-specific word features, and equate them with the original document's textual representation in a document clustering task on two collections of labeled scientific journal papers from two distinct domains. [10] The data clustering process is broken down into three steps: sampling of a small representative subset of the points, clustering of the tested data, and populating the first clusters with the remaining data followed by refinements. [11] clustering with similarity information. [12] we evaluate the effects of 9 semantic similarity measures with a term re-weighting method on document clustering of PubMed document sets [21].

III. EXISTING SYSTEM

In existing, using pair-wise queries to make informative points and it may take multiple queries to resolve the uncertainty about a data point. This system as uncertainty based sampling for supervised learning; an active learner queries the instance about which the label uncertainty is maximized. The probability making an instance learned clusters. Here to select an instance such that knowing its neighborhood will allow us to gain maximal information about the underlying clustering data. In particular instance focusing only on its relationship with the labeled instances [18].

IV. PROPOSED SYSTEM

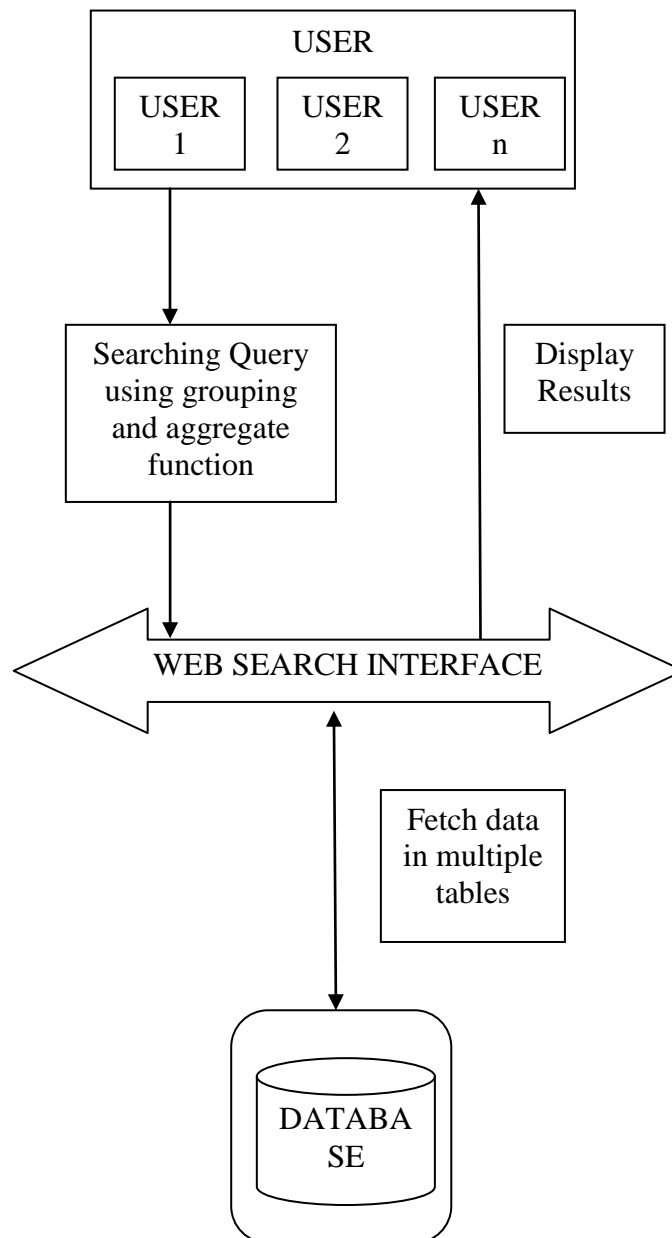
In Proposed System, Clearly the user based approach tries to satisfy more risk, unfavourable users pay a lesser amount of penalty, while this is not the case for random approach, the total number of paid penalties in random approach is greater than user based approach. This simulation contains user types with different Willingness To Pay (WTP), but similar risk aversion [19]. The average WTP of the users which their requests are assigned to virtual machines from different classes using the user based approach. Comparing the results, we can see the requests from users with higher WTP are assigned to virtual machines with higher reliability. This helps service provider to satisfy the users with higher WTP and improve its profitability. Supply and dependable management of SLA agreements is of principal importance for both Cloud providers and consumers [25].

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

V. PROPOSED SYSTEM ARCHITECTURE



System architecture can comprise system components, the visibly properties of those components, the relationships (e.g. the behavior) between them. It can deliver a plan from which products can be procured, and systems developed, that will work organized to implement the overall system [29].

Two phase technique such as Explore and Consolidate.

Explore (traverse in order to learn about it) or Neighbourhood based Framework



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

Input: We should provide sets of data points (D); the total number of classes c ; the maximum number of pair-wise queries T .

Output: a clustering of D into c clusters.

Initialization:

$C = \Phi$; $N_1 = \{x\}$, where x is a random point in D ; $N_b = N_1$; $l = 1$; $t = 0$;

repeat

$I =$ Semi supervised- Clustering (D, C)

$X^* =$ Most informative (D, I, C)

for each $N_i \in N_b$ in decreasing order of $p(X^* \in N_i)$

do

Query X^* against any data point $X_i \in N_i$;

$t++$;

then $l++$; $N_i = \{X^*\}$; $N_b = N \setminus N_i$;

until $t > T$

return Semi-Supervised-Clustering (D, C)

Consolidate or Most Informative

Input: We should provide set of data instances (D); the cluster then assigns I ; A set of neighbourhoods

Output: The most possible data point X^*

Learn a random forest classifier on and compute the similarity Matrix M ;

VI. MODULES

A. Authentication :

In this module presents users a form with username and Password fields. The user need to enters a valid username/password then the user will be granted to access the information. If the user types in invalid username and password that user will be considered as unauthorized user and denied access to that user. If the user has not account yet, goes to register then re login [30].

User: - Search Documents

In this module presents a user who gives the query to search particular documents from the database or server.

User: - View Documents

In this module presents a user who views particular documents from the database or server [36].

User:- Query based clustering

User has to select one option from given options. Both options are cannot link and must link .If user is selecting must link He will get the related documents for particular data. If he selects cannot link, he will gets the documents filtered by users and stored in database. By this way we can improve the performance [35].

User: - File Download

If user wants to download some documents first, he entered the query related to the documents in the given textbox then, just clicks the download option on the page.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

Admin: - View/Edit Documents

In this module admin wants to view all the documents stored in database. The documents from databases he verify that it is a must link or cannot link [37].

Admin: - File Upload

In this module presents admin wants to add new file on database. He just enters the documents details in given textbox and clicks the download option.

Admin: - Clusters as manually

In this module presents when admin upload a file that time he need to verify the documents details. Then, make a clustering as manually [38].

VII. RESULT ANALYSIS

- The user need to login at the home page or user can create an account



Fig 1 : Improve your document by typing the word and searching

- The user can search for a particular word by typing the word and searching



Fig 2 : Admin can view and edit the document

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

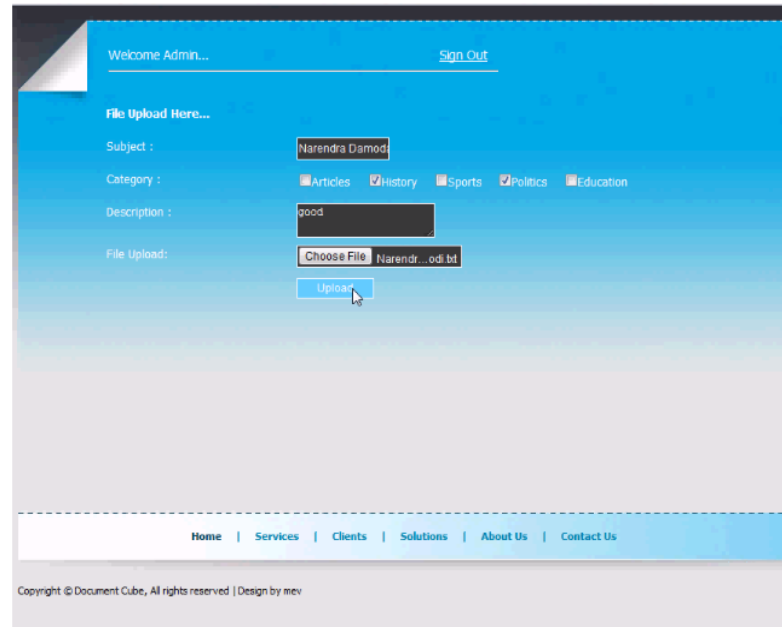


Fig 3 : search for a particular word by typing the word and searching

VII. CONCLUSION & FUTURE ENHANCEMENTS

In this paper, we studied an extend SQL by set predicates to support set-level comparisons. Such predicates, combined with grouping, allow selection of dynamically formed groups by comparison between a group and a set of values. We presented two evaluation methods to process set predicates. Comprehensive experiments on synthetic and TPC-H data show the effectiveness of both the aggregate function-based approach and the bitmap index-based approach. For optimizing the multipredicate queries, we planned a histogram-based probabilistic method to estimate the selectivity of set predicates. The approximation oversees the evaluation order of multiple predicates, producing efficient query plans.

We plan to implement in future, optimization strategy for the handle of queries with multiple set predicates connected by logic operations (AND, OR, NOT) and also investigated how Bitmap performs under different table sizes and query complexities.

VIII. ACKNOWLEDGEMENT

The author would like to thank the Vice Chancellor, Dean-Engineering, Director, Secretary, Correspondent, HOD of Computer Science & Engineering, Dr. K.P. Kaliyamurthi, Bharath University, Chennai for their motivation and constant encouragement. The author would like to specially thank **Dr. A. Kumaravel, Dean, School of Computing**, for his guidance and for critical review of this manuscript and for his valuable input and fruitful discussions in completing the work and the Faculty Members of Department of Computer Science & Engineering. Also, he takes privilege in extending gratitude to his parents and family members who rendered their support throughout this Research work.



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

REFERENCES

1. M. Arlitt and T. Jin, "A Workload Characterization Study of the 1998 World Cup Web Site," IEEE Network, Vol. 14, No. 3, pp. 30-37, May/June 2000.
2. D. Chamberlin, M. Astrahan, K. Eswaran, P. Griffiths, R. Lorie, J. Mehl, P. Reisner, and B. Wade, "Sequel 2: A Unified Approach to Data Definition, Manipulation, and Control," IBM J. R & D, Vol. 20, No. 6, PP. 560-575, 1976.
3. C.Y. Chan and Y.E. Ioannidis, "An Efficient Bitmap Encoding Scheme for Selection Queries," Proc. ACM SIGMOD Int'l Conf. Management of Data, 1999.
4. A.K. Chandra and P.M. Merlin, "Optimal Implementation of Conjunctive Queries in Relational Data Bases," Proc. Ninth Ann. ACM Symp. Theory of Computing (STOC), 1977.
5. A. Cuomo, G. Di Modica, S. Distefano, A. Puliafito, M. Rak, O. Tomarchio, S. Venticinque, and U. Villano, "An SLA-Based Broker for Cloud Infrastructures," J. Grid Computing, Vol. 11, PP. 1-25, 2013.
6. A.K. Chandra and P.M. Merlin, "Optimal Implementation of Conjunctive Queries in Relational Data Bases," Proc. Ninth Ann. ACM Symp. Theory of Computing (STOC), 1977. M. Esnaashari and M.R. Meybodi, "Data Aggregation in Sensor Networks Using Learning Automata," Wireless Networks, Vol. 16, No. 3, PP. 687-699, 2010.
7. K.G.S. Venkatesan. Dr. V. Khanna, Dr. A. Chandrasekar, "Autonomous System(AS) for mesh network by using packet transmission & failure detection", Inter. Journal of Innovative Research in computer & comm. Engineering, Vol. 2, Issue 12, PP. 7289 – 7296, December - 2014.
8. K.G.S. Venkatesan and M. Elamurugaselvam, "Design based object oriented Metrics to measure coupling & cohesion", International journal of Advanced & Innovative Research, Vol. 2, Issue 5, PP. 778 – 785, 2013.
9. Teerawat Issariyakul • Ekram Hoss, "Introduction to Network Simulator NS2".
10. S. Sathish Raja and K.G.S. Venkatesan, "Email spam zombies scrutinizer in email sending network Infrastructures", International journal of Scientific & Engineering Research, Vol. 4, Issue 4, PP. 366 – 373, April 2013.
11. G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," IEEE J. Sel. Areas Communication., Vol. 18, No. 3, PP. 535-547, Mar. 2000.
12. K.G.S. Venkatesan, "Comparison of CDMA & GSM Mobile Technology", Middle-East Journal of Scientific Research, 13 (12), PP. 1590 – 1594, 2013.
13. P. Indira Priya, K.G.S. Venkatesan, "Finding the K-Edge connectivity in MANET using DLTRT, International Journal of Applied Engineering Research, Vol. 9, Issue 22, PP. 5898 – 5904, 2014.
14. K.G.S. Venkatesan and M. Elamurugaselvam, "Using the conceptual cohesion of classes for fault prediction in object-oriented system", International journal of Advanced & Innovative Research, Vol. 2, Issue 4, PP. 75 – 80, April 2013.
15. Ms. J. Praveena, K.G.S. Venkatesan, "Advanced Auto Adaptive edge-detection algorithm for flame monitoring & fire image processing", International Journal of Applied Engineering Research, Vol. 9, Issue 22, PP. 5797 – 5802, 2014.
16. K.G.S. Venkatesan. Dr. V. Khanna, "Inclusion of flow management for Automatic & dynamic route discovery system by ARS", International Journal of Advanced Research in computer science & software Engg., Vol. 2, Issue 12, PP. 1 – 9, December – 2012.
17. Needhu. C, K.G.S. Venkatesan, "A System for Retrieving Information directly from online social network user Link ", International Journal of Applied Engineering Research, Vol. 9, Issue 22, PP. 6023 – 6028, 2014.
18. K.G.S. Venkatesan, R. Resmi, R. Remya, "Anonymizing Geographic routing for preserving location privacy using unlinkability and unobservability", International Journal of Advanced Research in computer science & software Engg., Vol. 4, Issue 3, PP. 523 – 528, March – 2014.
19. Selvakumari. P, K.G.S. Venkatesan, "Vehicular communication using Fvmr Technique", International Journal of Applied Engineering Research, Vol. 9, Issue 22, PP. 6133 – 6139, 2014.
20. K.G.S. Venkatesan, G. Julin Leeya, G. Dayalin Leena, "Efficient colour image watermarking using factor Entrenching method", International Journal of Advanced Research in computer science & software Engg., Vol. 4, Issue 3, PP. 529 – 538, March – 2014.
21. D. Chatziantoniou and K.A. Ross, "Groupwise Processing of Relational Queries," Proc. 23rd Int'l Conf. Very Large Databases (VLDB), pp. 476-485, 1997.
22. K.G.S. Venkatesan. Kausik Mondal, Abhishek Kumar, "Enhancement of social network security by Third party application", International Journal of Advanced Research in computer science & software Engg., Vol. 3, Issue 3, PP. 230 – 237, March – 2013.
23. Annapurna Vemparala, Venkatesan.K.G., "Routing Misbehavior detection in MANET'S using an ACK based scheme", International Journal of Advanced & Innovative Research, Vol. 2, Issue 5, PP. 261 – 268, 2013.
24. K.G.S. Venkatesan. Kishore, Mukhtar Hussain, "SAT : A Security Architecture in wireless mesh networks", International Journal of Advanced Research in computer science & software Engineering, Vol. 3, Issue 3, PP. 325 – 331, April – 2013.
25. Annapurna Vemparala, Venkatesan.K.G., "A Reputation based scheme for routing misbehavior detection in MANET'S ", International Journal of computer science & Management Research, Vol. 2, Issue 6, June - 2013.
26. K.G.S. Venkatesan, "Planning in FARS by dynamic multipath reconfiguration system failure recovery in wireless mesh network", International Journal of Innovative Research in computer & comm. Engineering, Vol. 2, Issue 8, August - 2014.
27. D. Chatziantoniou and E. Tzortzakakis, "Asset Queries: A Declarative Alternative to Mapreduce," ACM SIGMOD Record, vol. 38, no. 2, pp. 35-41, Oct. 2009.
28. K.G.S. Venkatesan, AR. Arunachalam, S. Vijayalakshmi, V. Vinotha, "Implementation of optimized cost, Load & service monitoring for grid computing", International Journal of Innovative Research in computer & comm. Engineering, Vol. 3, Issue 2, PP. 864 – 870, February - 2015.
29. R. Karthikeyan, K.G.S. Venkatesan, M.L. Ambikha, S. Asha, "Assist Autism spectrum, Data Acquisition method using Spatio-temporal Model", International Journal of Innovative Research in computer & communication Engineering, Vol. 3, Issue 2, PP. 871 – 877, February - 2015.
30. K.G.S. Venkatesan, B. Sundar Raj, V. Keerthiga, M. Aishwarya, "Transmission of data between sensors by devolved Recognition",



ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 3, March 2015

International Journal of Innovative Research in computer & comm. Engineering, Vol. 3, Issue 2, PP. 878 – 886, February -2015.

31. K.G.S. Venkatesan, N.G. Vijitha, R. Karthikeyan, “Secure data transaction in Multi cloud using Two-phase validation”, International Journal of Innovative Research in computer & comm. Engineering, Vol. 3, Issue 2, PP. 845 – 853, February -2015.
32. G. Antoshenkov, “Byte-Aligned Bitmap Compression,” Proc. Conf. Data Compression, 1995.
33. K.G.S. Venkatesan, “Automatic Detection and control of Malware spread in decentralized peer to peer network”, International Journal of Innovative Research in computer & comm. Engineering, Vol. 1, Issue 7, PP. 15157 – 15159, September -2013.
34. Sathish Raja, S.K.G.S. Venkatesan, “Electronic Mail spam zombies purify in email connection”, International Journal of Advanced Research in Computer Science Engineering & Information Technology, Vol. 1, Issue 1, PP. 26 – 36, June –2013.
35. K.G.S. Venkatesan. Dr. V. Khanna, S.B. Amarnath Reddy, “Providing Security for social Networks from Inference Attack”, International Journal of Computer Science Engineering & Scientific Technology, March –2015.
36. K.G.S. Venkatesan, Dr. Kathir. Viswalingam, N.G. Vijitha, “ Associate Adaptable Transactions Information store in the cloud using Distributed storage and meta data manager”, International Journal of Innovative Research in computer & communication Engineering, Vol. 3, Issue 3, PP. 1548 – 1555, March -2015.
37. A.K. Chandra and P.M. Merlin, “Optimal Implementation of Conjunctive Queries in Relational Data Bases,” Proc. Ninth Ann. ACM Symp. Theory of Computing (STOC), 1977.
38. “Jaql: Query Language for Javascript Object Notation (Json),”<http://code.google.com/p/jaql/>, 2013.