

# ECADS: An Efficient Approach for Accessing Data and Query Workload

**Rakesh Malvi**

M.Tech (COMPUTERAPPLICATION &TECHNOLOGY), UIT, BHOPAL, INDIA

E-mail: malviyarakeshsh1@yahoo.in

**Dr. Ravindra Patel, Dr. Nishchol Mishra**

M.Tech (COMPUTERAPPLICATION &TECHNOLOGY), UIT, BHOPAL, INDIA

E-mail: ravindra@rgtu.net, nishchol@rgtu.net

**Abstract**—In current scenario a huge amount of data is introduced over the web, because data introduced by the various sources, that data contains heterogeneity in nature. Data extraction is one of the major tasks in data mining. In various techniques for data extraction have been proposed from the past, which provides functionality to extract data like Collaborative Adaptive Data Sharing (CADS), pay-as-you-go etc. The drawbacks associated with these techniques is that, it is not able to provide global solution for the user. Through these techniques to get accurate search result user need to know all the details whatever he want to search. In this paper we have proposed a new searching technique “Enhanced Collaborative Adaptive Data Sharing Platform (ECADS)” in which predefined queries are provided to the user to search data. In this technique some key words are provided to user related with the domain, for efficient data extraction task. These keywords are useful to user to write proper queries to search data in efficient way. In this way it provides an accurate, time efficient and a global search technique to search data. A comparison analysis for the existing and proposed technique is presented in result and analysis section. That shows, proposed technique provide better than the existing technique.

**Index Terms**—CADS (Collaborative Adaptive Data Sharing Platform), Dynamic Forms, Dynamic Query, ECADS (Enhanced Collaborative Adaptive Data Sharing Platform), Nepal Dataset, QCV.

## I. INTRODUCTION

In last decade, enormous amount of data is flooded over the internet and that data is collected and stored in databases in all over the world. That data comes from the various social networking sites and some other media. To get that data use it in useful mean information and knowledge about that data is required. Data mining is the field where all the data and information about that data is fetched and useful information is extracted. In data mining various techniques are provided to extract useful information from that data. That data can be used in various fields like telecommunication, banking, any crisis

situation or some other means [3].

There are lots of software domains the place customers create and share information; for example, information blogs, scientific networks, social networking companies, or catastrophe management networks. Present understanding sharing instruments, like content administration software (e.g., Microsoft Share-point), permit customers to share documents and annotate (tag) them in an ad-hoc approach. In a similar way, Google Base permits customers to outline attributes for their objects or decide on from predefined templates. This annotation procedure can facilitate subsequent information discovery. Many annotation techniques allow best “untyped” keyword annotation: for example, a consumer may annotate a climate record utilizing a tag equivalent to “Storm category 3.” Annotation approaches that use attribute-worth pairs are on the whole more expressive, as they are able to contain extra know-how than untyped tactics. In such settings, the above know-how may also be entered as (Storm category, 3). A recent line of work towards utilizing more expressive queries that leverage such annotations, is the “pay-as-you-go” querying approach in Databases: In Databases, customers provide knowledge integration pointers at query time. The belief in such techniques is that the info sources already contain structured expertise and the challenge is to match the question attributes with the source attributes. Many methods, although, do not even have the elemental “attribute-worth” annotation that will make a “pay-as you- go” querying possible. Annotations that use “attribute value” pairs require customers to be more principled of their annotation efforts. Users should understand the underlying schema and discipline varieties to use; they will have to also know when to make use of each and every of those fields. With schemas that usually haven tens and even thousands of available fields to fill, this assignment becomes complicated and cumbersome. This results in information entry users ignoring such annotation capabilities. Even supposing the procedure permits users to arbitrarily annotate the information with such attribute-worth pairs, the users are commonly unwilling to participate in this project: The undertaking no longer simplest requires gigantic effort but it also has doubtful usefulness for subsequent searches sooner or later: who is going to make

use of an arbitrary, undefined in a normal schema, attribute type for future searches? However even when making use of a predetermined schema, when there are tens of potential fields that can be utilized, which of these fields are going to be priceless for looking the database sooner or later?

In data mining various predefined tools are provided to the user to extract valuable information from that data. There are various methods like machine learning, or some other techniques are used to provide a way automated extraction of the data. Data mining is a knowledge discovery process which used to extract information from the various type of resource data and that data can be used for various applications. Generally classification, regression, clustering are used to classify that data into different classes which provides a better identification for the data and also help in the process of extraction [4].

In Classification data classified into different classes and then process of extraction is applied to extract data on the basis of that classification. In that way data is classified into different classes and groups. That data used as a training data too train untrained data and classify untrained data into different classes. In that that data can be used for various type of decision making [5].

There are various type of techniques are presented by the various researchers to provide better classification for that data. Like there are various type of platforms are provided by the Google to manage and define measures for objects. Thus that helps to classify data into various classes [7].

CADS [1] provides cost effective and good solution to help efficient search result. The goal of CADS is to support a process that creates nicely annotated documents that can be immediately useful for commonly issued semi-structured queries of end user.

Annotation methods that use attribute-value dyads square measure usually additional communicatory, as they'll contain additional data than untyped approaches. A recent lines of labor towards utilizing additional communicatory queries that leverage such annotations, is that the "pay-as-you-go" querying strategy in knowledge areas: In knowledge spaces, users give knowledge integration hints at question time. The position such systems is that the data supplies already contain structured information and therefore the quandary is to match the question attributes with the source attributes.

Even if the system permits users to annotate the information with such attribute-value pairs, the users square measure typically unwilling to perform the task. Such difficulties results in terribly basic annotations that is typically restricted to straight forward keywords. Such straight forward annotations create the analysis and querying of the info cumbersome. User's square measure typically restricted to plain keyword searches, or has access to terribly basic annotation fields, like "creation date" and "size of document", we tend to propose CADS (Collaborative adaptive information Sharing) platform that is associate "annotate-as-you-create" infrastructure that facilitates fielded information annotation. A key con-

tribution of our system is the direct use of the question employment to direct the annotation method, additionally to examining the content of the document. Our aim is to rank the annotation of documents towards generating attribute names and attribute values for attributes can that may typically employed by querying users and these attribute values will give best potential results to the user whereby users will ought to deal solely with relevant result.

In this paper a new technique called enhanced collaborative data sharing (ECADS) is presented. In that technique various predefined tools and key words are provided to the used to search data from a large dataset, in existing technique there is no global solution is provided to the user. In existing technique provide a mechanism to search data in which used need to have prior knowledge about the structure of the system thus a new technique is presented in this paper which provides more accurate search results as compare to the other technique.

A short description of proposed technique presented below:

First user need to select field which he want to search data.

Suggestions for the search keywords are provided to the user which he can used to form query for search data. An easy and flexible selection mechanism is provided to select keyword and an automatic query can be formed to provide better search result for the user. A comparison of the search results for existing and proposed techniques is presented in result and analysis section which shows that proposed technique provides accurate and time efficient search results in context of search.

The rest of the paper organizes as follows: in section II, a brief literature review over the techniques which used to provide solution to search data is presented. A related work which shows, a description over the CADS technique for search data is presented in section III. In section IV a description over the proposed technique is presented. In section V, experimental setup and result analysis for the proposed technique is presented. An evaluation of the results for proposed and existing technique is presented in VI. Section VII concludes the paper.

## II. LITERATURE REVIEW

In [2] pay as you go user feedback system based technique is presented to provide feedback for the data space system. A technique is provided which provides a pay-as-you-go query technique in data space is presented. That technique uses queries which take annotations to provide an efficient query technique to search data, in that data space user provides hints to user to search data. But in that technique a pre assumption consider that user have all the structured information of the searched data there is a problem is occurs which is to match the query attribute to the source attribute.

In [3] a technique is proposed to provide a framework for fast disaster recovery in the business community information network. In that a crisis management system and disaster recovery system is used to provide is used.

These systems provide a technique to deal situation where crisis occurs and merge it to the natural calamity. There are various techniques are used to provide a framework to recover and prevent these issues. A proper solution is required for such problem. The technique presents in this paper resolves the issues related to these problems.

In [6] an ensemble technique called random K-label-sets (RAKEL) is presented which provides a multi-label classification for the data. In that technique (RAKEL) first each ensemble divided into labeled subsets and then a single label classifier is learned to provide a combine results for these classifier as a power set of these subsets. By the use of label correlation, single label classifier and single label classifiers are applied to over the subsets in a controlled way, with limited no of data per label. In that technique correlation used between tags to annotate the label. But there is collaborative annotation mechanism is provided.

In [10]. Procedure to learn about the objects that require annotation by the use of studying from a set of until now annotated records. This as a rule requires the mark-up of a big assortment of files. The MnM system, for illustration, was developed to examine how this venture would be facilitated for domain experts. Comfortably marking a number of files just isn't ample; the items marked have obtained to be excellent examples of the types of contexts in which the gadgets are found out.

Discovering the right mixture of exemplar records is a more difficult venture for non IE specialists than the time-ingesting work of marking up a pattern of records. Melita addressed this predicament through suggesting the best mixture of records for annotation. Unsupervised programs, like Armadillo, are starting to deal with these challenges via exploiting unsupervised finding out approaches. PANKOW (used in OntoMat), for instance, demonstrates how the distribution of distinct patterns on the internet can be used as evidence so that you could approximate the formal annotation of entities in web pages through a principle of 'annotation through maximal (syntactic) proof'. For illustration, the number of events the phrase "cities reminiscent of Paris" occurs on web websites, would provide one piece of proof that Paris is a town, which would be considered inside the mild of counts of exceptional patterns containing "Paris".

In [14], a segmentation technique for the words in hand-written Gurumukhi script is presented. Segmentation of the words into characters is one of the challenging task to do that becomes more difficult when segmenting a handwritten script and touching characters. A water reservoir based technique is presented. If water is poured from top to the bottom of the character. That water stored at the cavity region of the character. Like if water poured from the top that region considered as top reservoir. To analyze that technique 300 hand written scripts are collected from the Gurumukhi script.

In [15], a data extraction framework which uses a natural language processing based feedback extraction technique is presented. In that data is extracted from the

feedbacks dataset to extract sematic relations in the domain ontology. Sematic analysis is the process understanding meaning of the words to which it represents, it the process of determining the meaning of the sentence on the basis of the word level meaning in the sentence.

III. RELATED WORK

Existing work is done in this work format where the annotation scheme is being improved by CADS technique; Fig. 1 shows the mechanism how the existing flow is working.

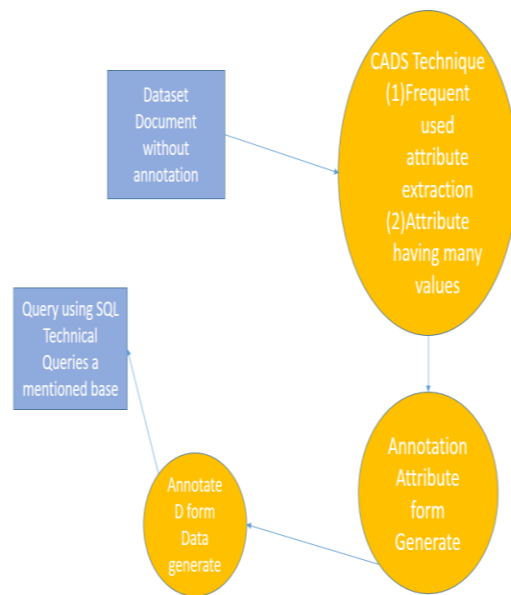


Fig.1. Flow-diagram for Existing technique.

QV: Query value is the Suggest attributes based on the querying value component, which is similar to ranking attribute based on their popularity in the workload.

CV: This is the content value attribute based on Suggest attributes based on the content value.

In [13] a CADS and USHER technique is presented to provide better performance to extract useful data from the raw data. CADS (Collaborative Adaptive Data Sharing Platform) is a technique which provides a mechanism to automatically generate query form and provide functioning to search data from the raw data. An USHER technique is used to enhance the quality of the results. Thus that technique resolves all the limitations of the existing technique and provide enhanced query to extract useful data from the raw data.

In [1], an enhanced technique to search data from the large scale data is presented. A structure of the metadata is generated by analyzing the document which contains useful information to extract data from a large scale data. A structured algorithm is presented that identifies the structured features which presents in the documents. Thus that technique called CADS provide4s an enhanced functionality to extract useful information from the large scale dataset that can be used in various applications.

A. Issues with previous technique

- In existing technique there is no suitable global solution is provided to the user to annotate and search data.
- In existing technique solution provided only can be handled by the user who is aware off the technique or belongs from technical stream. But in case of nontechnical person if they want to search data about the disaster management or crisis etc. it need to be known about the structure of the process which not efficient way to done that process.

An efficient technique is required to provide a global solution which enhance the efficiency of the on-demand Service which used to provide these information to the user. Some issues with the existing techniques are discussed below

- There is no efficient technique is provided to the user.
- Available techniques are more suitable for technical persons
- To get accurate results, a proper query is required.

IV. PROPOSED WORK

An improved CADS technique is presented, where an efficient query technique is provided to the user to search data. In that technique a frame work is provided by which any user can access the data efficiently. In that technique a predefined query generator is provided to the user to generate query for the search. Thus in that an accurate query is provided to the user which helps to generate proper results. Suppose a user want to search any information about any crisis, in that technique some predefined keywords are provided to provide proper solution for the user. In that in that technique user have no need to know about the structure of the whole process or remember the whole query to search data. Thus in that way that technique resolves all issues which occurs in previous techniques.

That technique provides better results than the other techniques. A description of the analysis result is presented which shows that proposed technique provides improved performance as compare to the existing techniques. Precision, Recall and computation time taken as a parameter.

In order to calculate the precision and recall we have various parameter values from the dataset progress:

Which are TP, FP, TN, and FN, which is getting in the form of confusion matrix?

True positive rate (or sensitivity):  $TPR=TP/ (TP+FN)$

False positive rate:  $FPR=FP/ (FP+TN)$

True negative rate (or specificity):  $TNR=TN/ (FP+TN)$

False negative rate  $FNR=1-TPR$

Recall and precision are generally used to evaluate the performance of the technique in that results of the every technique is observed and used to provide a comparison of the performance of the various techniques to evaluate the performance of the technique. There are various techniques are used to retrieved the image.

For query q,

$A(q)$  =A set of images in dataset

$B(q)$  =A set of images relevant to query

Precision =  $A(q) \cap B(q)/A(q)$

Recall= $A(q) \cap B(q)/B(q)$

Precision provides description about the images which are relevant to the query image, it is the ratio of no of relevant image to the retrieved image and recall is the ratio of the no. of relevant image to the query and to the no of relevant image in the database. Thus these techniques are used to evaluate the performance of the technique. Same in the case of text or any search data. In this fig. 2 show flow chart the step by step technique will be used. We will get less as less resources need to provide while the data need publish in public. This things we can perform with different retrieval technique on CADS.

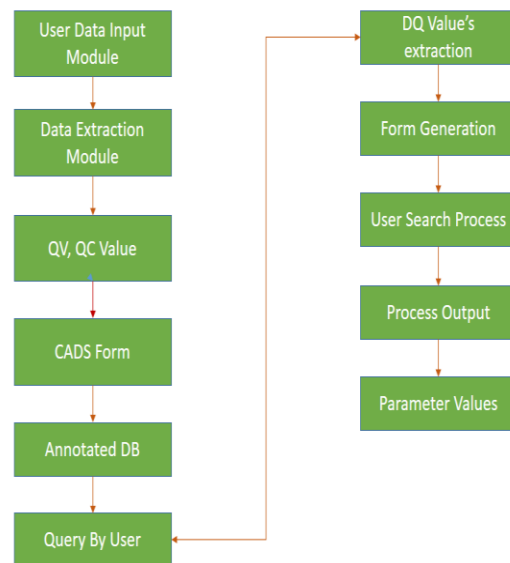


Fig.2. Flow Chart for ECADS Technique.

Ecads Algorithm:-

In order to perform the execution of the form based technique at querying side approach we have used the techniques which use the advantage of existing CADS approach. The Abbreviation usage in this document is described in this table 1:

Table 1. Shows the abbreviation and meaning

Abbreviation	Meaning and usage
CV	Content value, a data usage value from the existing dataset.
QV	Query value of the data, available attributes in dataset.
QQV	Query side or database side attribute value in the system.
QCV	A content value which is extracted from the database side.
T	Threshold value usage in proposed technique in order to generate form at query side based on the QQV.
R	The consider value output storage element.
Qn	Query table input. From first column to nth column in dataset.
Ci-n	Column indexing from i to n i.e. from i=0 or initialization to end of the column value from database end.
Sds	Sample dataset consider for the simulation environment.
A(Q)	It is the set of values in dataset.

[A]. ECADS Algorithm Pseudo CODE:

```

For (Qi-n)
{
Read data and process using technique Classification.
Go to next;
}
Obtain QCV <- QQV
Calculate the Score Waj;
QCV<Qaj & QQV
Content value=QCV= Value of (P (Aj|w))

Calculation T
Score value (T) =Maximum Average values Of (Value of
(P (Aj|w)))
Sequentially process data
T> value
R<- put value
Generate form and extraction data.
    
```

[B]. ECADS Algorithm:-

Input: SDS, Models.  
 Output: Query output & QCADS Form  
 Steps:  
 Retrieve next Qi from column 1-n.

In this step the first is going to load the dataset and then retrieve the column value from the initial point to the end of column, data is going to read from the first column up to the end of column in order to read the value and follow it for the further work to generate dynamic form according to the condition and algorithm.

Get the column and data value for each column ci Calculate QQV → QCV.

In this step first the calculation is to obtain the knowledge about the column value which is query query- ing value name is given, such that it is working at data

base end the query value mechanism described in the existing paper. [1] And based on the query value for each query value its content value is going to obtained in this step.

Here WAj be the set of queries in W that use Aj as one of the predicate conditions.

$$P (Aj|w) = (|Waj|+1) / (|Wa|+1) \tag{1}$$

Calculate QCV.

The calculation of content value from the QQV is going to determine in this step such that it can categorize for the form generation suitability.

$$\text{Content value} = \text{QCV} = \text{Value of } (P (Aj|w)) \tag{2}$$

Here with the help of QQV we extract the value of QCV.

Constant or Calculate T .Where QCV is the maximum possible extraction of the all unseen values.

Score value (T) is generated here –

$$\text{Score value (T)} = \text{Maximum Average values Of (Value of } (P (Aj|w))) \tag{3}$$

In this step a conclusion value for the form generation eligibility , a threshold value based on the content value is calculated , and the best available value are consider to the further process using this mechanism.

```

Sequentially process data
T> value
R<- put value
    
```

In this process we evaluate the value and read it and if it found the QCV according to the T, it is going to consider as final value for the form generation and storing in to a variable R, i.e. form generation attribute.

```

Apply QCV
Else
Go to Step 1
    
```

Finally in this step a query side form is generated in order to process the data available in form manner such that an efficient query technique can be apply in the dataset.

V. EXPERIMENTAL SETUP & RESULT ANALYSIS

In order to work on our proposed work we have performed our experimental setup and analysis using java API using JDK 8.0 where we have used swing api for design and other supporting api for implementing logical requirement for the algorithm and following results are produced while comparing the existing and proposed algorithm.

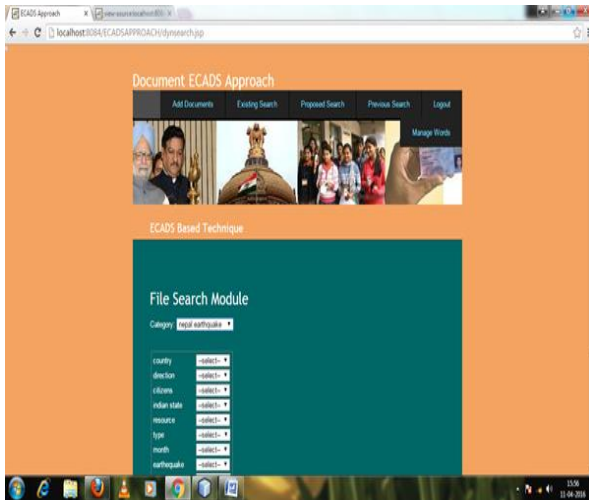


Fig.3. File search module for ECADS.

A NETBEANS IDE a simulator which provides development environment to develop projects in java is used.

In the proposed technique a global query architecture to provide an enhanced query to search data is presented. In fig. 3. A query generation mechanism is presented. That provide keywords to build query to search data and enhance the performance of the whole search mechanism.

In existing technique, there is query generation mechanism is provided which degrades the performance of the whole search mechanism. Because if user not aware about the technical aspects of the topic for which the search operation is performed, that is too difficult to build an accurate query for the search operation. That degrades the performance of the whole technique in terms of accuracy and efficiency. Thus an enhanced technique is presented in this paper.

To implement the proposed technique a dataset of the earthquake in Nepal and disaster management is used. And an analysis over that data is performed to analyse the performance of the proposed technique.

A statistical and graphical comparison for the search results provides by the existing and proposed technique is provided in this section. Precision, recall and execution time are taken as an evaluation parameter to evaluate the performance of the techniques.

Table 2. Comparison of the existing technique and Proposed technique.

Algo name	TP Value	FN Value	FP Value	TN Value
Existing Approach	53	8.0	67.0	22.0
Proposed Approach	53	37.0	38.0	22.0

Upon getting following parameter we have calculated actual required parameter to compare technique such as precision, recall, computation timing from the two different techniques and got the values mentioned in below table as statically analysis of result. In which we have got the less computation time of our proposed approach and required precision and recall value using Nepal dataset used in that project.

In table 2, a calculation for the True Positive (TP), True negative (TN), False Positive (FP) and False Negative (FN).

TP = correctly identified.

TN = incorrectly identified.

FP = correctly rejected.

FN= incorrectly rejected.

These values are used to calculated the precision and recall for the technique which taken as a parameter to evaluate the performance of the technique.

Precision = positive predictive values= TP/ (TP+FP).

Recall = True Positive Rates= TP/ (TP+FN).

A comparison of the precision, recall and computation time for the existing and propose technique is presented in Table 3. That shows proposed technique consumes less amount of time to provide search results for the query and comparison of the precision and recall shows that the proposed technique provides better search results as compare to the existing technique. That way proposed technique provides an enhanced mechanism to search data.

Table 3. A statistical analysis of results for the proposed technique.

Algorithm name	Computation time (in ms)	Precision (%)	Recall (%)
Existing approach	25022	44.16	88.88
Proposed approach	9142	58.24	58.88

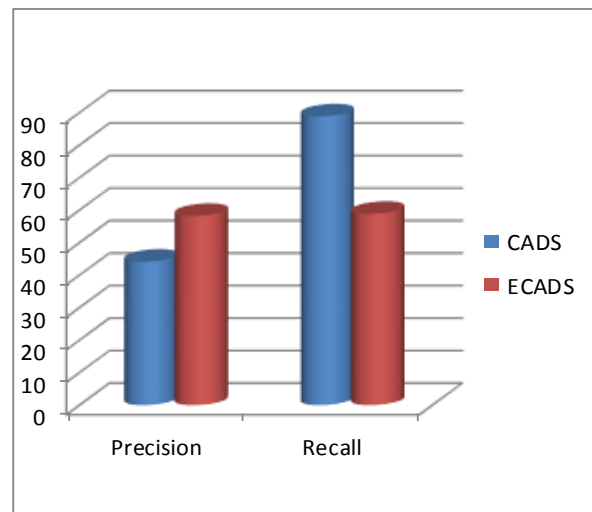


Fig.4. Result chart for Precision and Recall Value.

A graphical comparison for the existing and proposed technique over the parameters called precision and recall is presented in Fig. 4, which shows that proposed tech

nique provides accurate results in context of search. We generate the auto query which uses dynamic form based query and data accessing system. Recall and precision are generally used to evaluate the performance of the technique in that results of the every technique is observed and used to provide a comparison of the performance of the various techniques to evaluate the performance of the technique. So all the evaluation process of query the dataset are always done using these two calculation.

A graphical comparison over time is presented in Fig. 5, which shows that the proposed technique taken small span of time to provide search results. So that user can remove from the burden of query formation, thus computation time can be reduced. We can definitely get good results in all computation time.

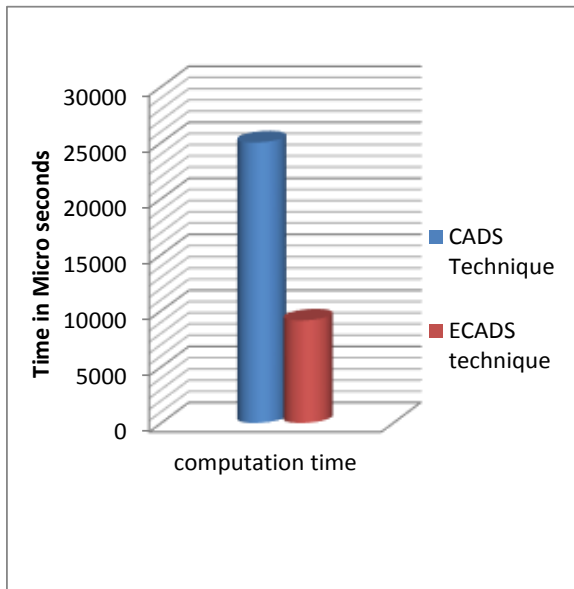


Fig.5. Graphical comparison over Execution time.

In that way a search technique is provided to the user which provides time efficient and accurate search result in context of search.

VI. EVALUATION OF EXPERIMENTS

To evaluate the performance of the technique precision and recall are generally used in case of search technique. A statistical and graphical analysis of the technique is presented in section V. That shows proposed technique provides better performance as compare to the other techniques.

For query q,

A (q) =A set of images in dataset

B (q) =A set of images relevant to query

Precision = A (q) ^B (q)/A (q)

Recall=A (q) ^B (q)/B (q)

Computation Time:

Computation time is the length of time required to perform a computational process. Computation time can be represented as a sequence of time slots for performing computation on the various available segments of the services. The computation time is proportional to the number of services.

A comparison analysis for the existing and proposed technique is presented in the section V. That shows proposed technique provides an accurate and efficient mechanism to build an accurate query to perform search operation data.

VII. CONCLUSION

There are various techniques which used to search data in large dataset scenario, but these technique suffers some defects like high computation time, no global solution is provided etc. in existing technique to get accurate result for the search data a proper structure for that data is required. That is difficult to every user to know the internal structure thus a new technique is proposed in this paper. That technique provides predefined keywords to search the data is provided these keywords contains the information which help user to search data and get result without knowing the structure of the whole system. In that way a simple common person who don't know anything about the system can perform their search operation and get accurate result. A result analysis is presented in section V. which shows that proposed technique provides better results as compare to the other techniques.

REFERENCES

- [1] Eduardo J. Ruiz, Vangelis Hristidis, and Panayiotis G. Ipeirotis: proposed a paper "CADS Technique Introduced to Generate Dynamic form to Maintain Data in Structure Format." IEEE, 2014
- [2] S.R. Jeffery, M.J. Franklin, and A.Y. Halevy: proposed a paper "Pay-as-You-Go User Feedback for Data space Systems," ACM, 2008.
- [3] K. Saleem, S. Luis, Y. Deng, S.-C. Chen, V. Hristidis, and T. Li: proposed a paper "Towards a Business Continuity Information Network for Rapid Disaster Recovery" 2008.
- [4] J. M. Ponte and W.B. Croft: proposed a paper "A Language Modeling Approach to Information Retrieval".
- [5] R. T. Clemen and R.L. Winkler: proposed a paper "Unanimity and Compromise among Probability Forecasters.
- [6] G. Tsoumakas and I. Vlahos's: propose a paper "Random Label sets: An Ensemble Method for Multilevel Classification.
- [7] P. Heymans, D. Ramage, and H. Garcia-Molina: proposed a paper "Social Tag Prediction".
- [8] Y. Song, Z. Zhuang, H. Li, Q. Zhao, J. Li, W.-C. Lee, and C.L. Giles: proposed a paper "Real-Time Automatic Tag Recommendation".
- [9] D. Eck, P. Lamere, T. Bertin-Mahieux, and S. Green: proposed a paper "Automatic Generation of Social Tags for Music Recommendation.
- [10] Victoria Uren, Philipp Cimiano, José Iria, Siegfried Handschuh, Maria Vargas-Vera, Enrico Motta, and Fabio Ciravegna "Semantic Annotation Systems for Knowledge

Management: A Survey of Requirements and State of the Art” MIAKT, 2004.

- [11] B. Russell, A. Torralba, K. Murphy, and W. Freeman: propose a paper “Label Me: A Database and Web-Based Tool for Image Annotation”.
- [12] C.D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval, first ed. Cambridge Univ.Press, <http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0521865719>, July 2008.
- [13] Anita L. Devkar, Dr. Vandana S. Inamdar “Combine Approach of CADS and USHER Interfaces for Document Annotation” IJIRCCCE, May, 2015.
- [14] Munish Kumar, M.K. Jindal, R.k. Sharma “Segmentation of Isolated and Touching Characters in Offline Handwritten Gurmukhi Script Recognition” IJITCS, 2014.
- [15] Pratik K. Agrawal, Avinash. J. Agrawal “Opinion Analysis Using Domain Ontology for Implementing Natural Language Based Feedback System” IJITCS, 2014.

**How to cite this paper:** Rakesh Malvi, Ravindra Patel, Nishchol Mishra, "ECADS: An Efficient Approach for Accessing Data and Query Workload", International Journal of Information Technology and Computer Science(IJITCS), Vol.8, No.12, pp.39-46, 2016. DOI: 10.5815/ijitcs.2016.12.05

### Authors' Profiles



**Rakesh Malvi**, is currently working toward the M.Tech degree in the department of School of Information Technology in Rajiv Gandhi Pradyogiki Vishwavidyalaya, Bhopal. He has also received the BE degree from the RKDF College of Technology & Research, Bhopal.



**Dr. Ravindra Patel**, is working as Associate Professor in Department of Computer Applications in Rajiv Gandhi Pradyogiki Vishwavidyalaya, Bhopal. He has done Master in Computer Applications from Barkatullah University, Bhopal, India in 1998. He has completed his Ph.D. in Computer Science from Rani Durgawati University, Jabalpur, India in 2005. He has published 32 research papers in various international journals (including Elsevier, Springer, IEEE index etc.) And around 11 research papers in proceedings of various peer-reviewed conferences in India and abroad. He has contributed chapters in edited books published by Elsevier, Springer. He has more than 14 years of teaching and research experience in post graduation. He has more than 8 years of administrative experience as Head of Department. His area of research includes but not limited to Data Mining, Big Data analytics, Cyber Security, Human-computer Interaction.



**Dr. Nishchol Mishra**, is working as Associate Professor in School of Information Technology in Rajiv Gandhi Pradyogiki Vishwavidyalaya, Bhopal. He has done his M.Tech & Ph.D in Computer Science & Engineering. He has published over 40 papers in various international journals of reputed. He has more than 14 year of research experience. His area of research includes Multimedia, Data Mining, and Image Processing & Data Analytics.