ISSN No. 0976-5697

Volume 3, No. 3, May-June 2012

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Framework for Context Ontology Driven Indexing, Ranking and Search in Search Engines

Parul Gupta*
Department of Computer Science and Engineering
YMCA University of Science and Technology
Faridabad
parulgupta_gem@yahoo.com

Dr. A.K.Sharma
Department of Computer Science and Engineering
YMCA University of Science and Technology
Faridabad
Ashokkkale2@rediffmail.com

Abstract: The size of the publicly indexable World Wide Web (WWW) has probably surpassed 14.3 billion documents and as yet growth shows no sign of leveling off. As more information becomes available on the Web it is more difficult to provide effective search services for Internet users. This paper integrates the proposed architecture for hierarchical clustering based indexing, multilevel indexing, context based index which further leads to the creation of ranked context based index using ontology. Context-ontology is a shared vocabulary to share context information in a pervasive computing domain and include machine-interpretable definitions of basic concepts in the domain and relations among them. Further ontology driven conjunctive query expansion based on mining user logs and architecture for relevant searching of web documents using data mining techniques such as clustering and association rules has been presented. Context Ontology has also been applied to ranking of the search results. An attempt has been made to measure the performance and compare the results of the prevailing approach and the new approach.

Keywords: Context Ontology, multilevel indexing, query expansion, clustering, context repository

I. INTRODUCTION

Search Engines processing steps are carried out by three distinct and cooperating modules [1]. The *Crawler* gathers Web documents and stores them into a huge repository after being compressed. Every Web page has an associated ID number called document identifier, which is assigned whenever a new URL is parsed out of a web page. The *Indexer* parses the document & abstracts the data contained within the document repository, and creates an *Index* granting fast access to document contents and structure. Finally, the *Searcher* module accepts user queries, searches the index for the documents that best match each query, and returns the *references* to these documents in an understandable form. Its goal is to provide quality search results.

The indexing phase can be viewed as a Web Content Mining process. Starting from a collection of unstructured documents, the indexer extracts a large amount of information like the list of documents, which contain a given term. In the current scenario, the indexing structure contains the context and hence the document identifiers of the documents belonging to that context. This context represents the theme of the document that has been extracted using context repository, thesaurus and ontology repository. This is the final index that is constructed after extracting the context of the document. Rather than being formed on the term basis, the index is constructed on the context basis with context as first field and finally the document identifiers of the relevant documents.

The transfer of irrelevant information in the form of documents retrieved by an information retrieval system [2] and that are of no use to the user simply wastes network bandwidth and frustrates users. This condition occurs due to improper and time consuming ranking of the documents after the documents have been retrieved by the searcher from the index on the basis of keyword matching. Therefore, the major issue to be addressed in information selection is the development of a search mechanism that will help in getting maximum relevant documents. The possible solution to this problem can be ranking of documents beforehand and better ranking of the documents retrieved in response to the query.

Due to the extremely large volume of documents on the web, analysis of the entire collection is not feasible. In addition, web queries are very short, often consisting of only two or three words. Techniques that have proved to be successful with longer queries may not prove to be very effective with short queries. A single-term query that a normal user formulates often retrieves many irrelevant articles as well as fails to find hidden knowledge or relationships buried in content of the articles. Query expansion is often intended to overcome a vocabulary mismatch between the query and the document collection. Through query expansion, the effects of the word mismatch problem are reduced, resulting in a higher ratio of relevant documents in the retrieval results (precision) and a higher ratio of the relevant documents from the collection that are retrieved (recall). The main problem of query expansion is that in some cases the expansion process worsens the query performance. The ontology's can be used for query expansion in which firstly the query terms must be disambiguated so that they map to a unique ontology concept and then terms related in the ontology to the disambiguated concepts are added to the query.

Search engines uses ontology to find pages with words that are syntactically different but semantically similar. Ontology [3] is used for knowledge sharing and reuse. Ontology building from data mining can be achieved in two phases. The data-mining phase is related to data mining process including data mapper, rule formulator and extraction of knowledge. The ontology-building phase is related to the process of building the ontology from the extracted knowledge, which represents the output of the data mining. As the web becomes more pervasive considering the context in the query matching process can improve the quality of the retrieved results. In context-ontology based Web mining, the instances of concepts and relationships in a given ontology, or using them to discover other useful knowledge based on contextual information are discovered. The returned results are relevant according to the needs of user and less non-relevant results are filtered off. So the semantics of the user's query and of the contextual information that is considered relevant in the matching process can be captured to improve the quality of returned results.

II. RELATED WORK

The work in [4] presents an algorithm for hierarchical clustering based indexing. In this the clustering of the documents is done at various levels to build a compressed index. [5] presents multilevel indexing in which index is created at multiple levels thus directing the search to a specific path. The work in [6] is based on context based indexing in which the index is constructed on the basis of context rather than on the basis of terms. In [7], context driven Pre ranking has been proposed which pre ranks the documents under the similar context. It also shows an algorithm to rank the search results on the basis of context ontology. The work in [8] presents the relevant search using data mining techniques.

III. PROPOSED ARCHITECTURE

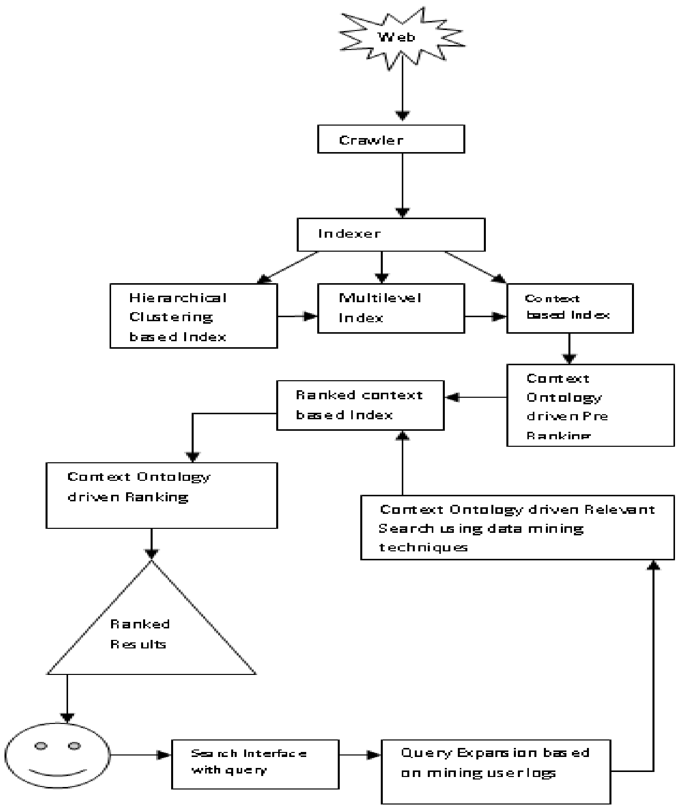

Web
Crawler
Indexer
Hierarchical Clustering based Index
Multilevel Index
Context based Index
Context Ontology driven Pre Ranking
Ranked context based Index
Context Ontology driven Ranking
Context Ontology driven Relevant Search using data mining techniques
Ranked Results
Search Interface with query
Query Expansion based on mining user logs

Figure 1 Framework for Ontology driven Indexing, Ranking and Search in Search Engines

A. *Hierarchical Clustering based Index:*

Clustering based Index [4] is based on a clustering algorithm that aims at partitioning the set of documents into ordered clusters so that the documents within the same cluster are similar and are being assigned the closer document identifiers. The extension of this clustering algorithm to be applied for the hierarchical clustering in which similar clusters are clubbed to form a mega cluster and similar mega clusters are then combined to form super cluster. Thus the different levels of clustering optimizes the search process by directing the search to a specific path from higher levels of clustering to the lower levels i.e. from super clusters to mega clusters, then to clusters and finally to the individual documents so that the user gets the best possible matching results in minimum possible time.

This algorithm picks the first document as cluster representative, then selects the most similar document to it and puts it in the cluster, it further selects document which is most similar to the currently selected document and repeats until the first cluster becomes full with n/k documents. The same process is then repeated to form the rest of the clusters. Thus the most similar documents are accumulated in the same cluster and are assigned consecutive document identifiers. Thus the algorithm is also efficient in compression of the index.

The framework for the hierarchical clustering is shown in figure.

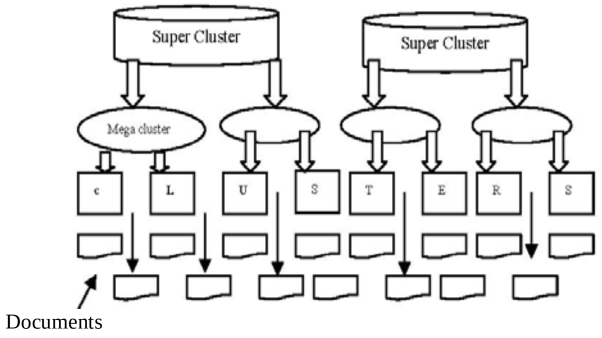


Figure 2 Hierarchical Clustering

B. *Multilevel Index:*

A global single large size index takes a long time to search the relevant documents in response to the user queries. So multilevel indexing was proposed i.e. creating the indexes at the multiple levels so as to reduce the search time by directing the search to a specific path from higher level indexes to the lower level indexes. The multilevel indexes [5] are created after the hierarchical clusters of the documents have been formed on the basis of document similarity. At each level of documents clusters hierarchy, a separate index is created and the search for a document proceeds from the higher level indexes to the lowest level index which is the document level index and is stored as inverted file.

At the lowest level that is at the document level, the index consists of the union of the terms in all the documents in different clusters. This index is a descriptive index whose structure is similar as that of the global large size index that exists in the current architecture. This document level index is stored as array of posting lists containing the term as well the document identifiers of the documents containing the given term. The second level index that exists at the mega cluster level consists of the terms that come after the intersection of the words of the similar clusters. The index at this level is a brief index which contains only the terms with no description at all. The highest level index which exists at the super cluster level consists of the intersection of the terms of the similar mega clusters. This highest level index is also a small size index containing only the terms. The search proceeds from the highest level index to the document level index due to which there is reduction in the search space.

a. *Structure of Multilevel Indexes:*

The structure of the index at the lowest level is in the form of posting lists as shown in figure:

Term	**No. of docs in which term appears**	**Docid of docs in which term appears**
Indexing	50	12,34,45,49…
Search engine	59	15,20,34,55…
Pipeline	15	3,6,9,12…

Figure 3 Structure of Cluster level index

However the indexes at the higher level consist of only the document terms with no extra information stored in them. The structure of the higher level indexes is shown in fig.

Term
Indexing
Crawling
Search engine
Pipeline
Parallel
Distributed

Figure 4 Structure of higher level indexes

b. *Implementation Of Proposed Work:*

For indexing the documents, firstly we have to parse the documents. After that similarity matrix is created and then k means algorithm is applied for creating the clusters. Clusters will be created at first level .For creating clusters at second level same procedure is applied again and then finally hierarchical clustering is done for indexing.

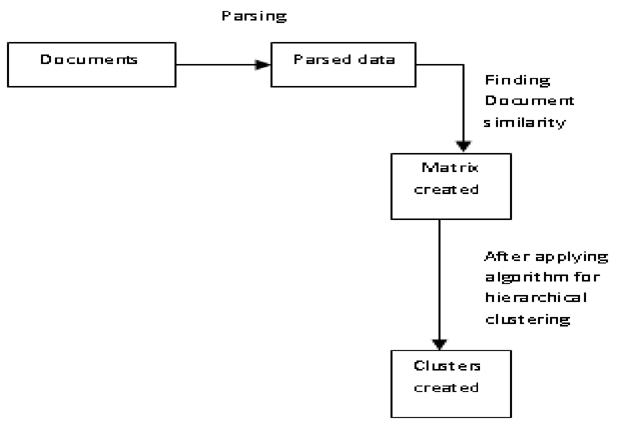
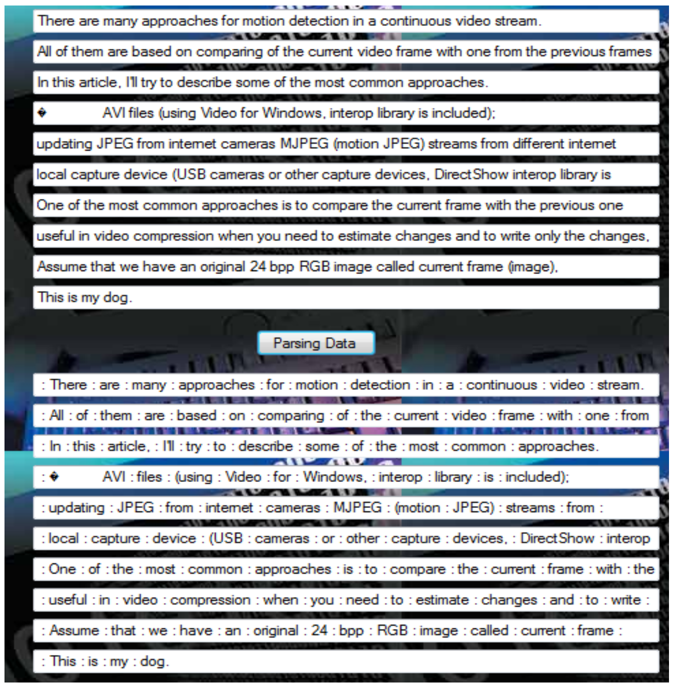

Parsing
Documents
Parsed data
Finding Document similarity
Matrix created
After applying algorithm for hierarchical clustering
Clusters created

Figure 5 Work Flow of Implementation

c. *Snapshots Of Implemented Work:*

a) *Given input & parsed data:*

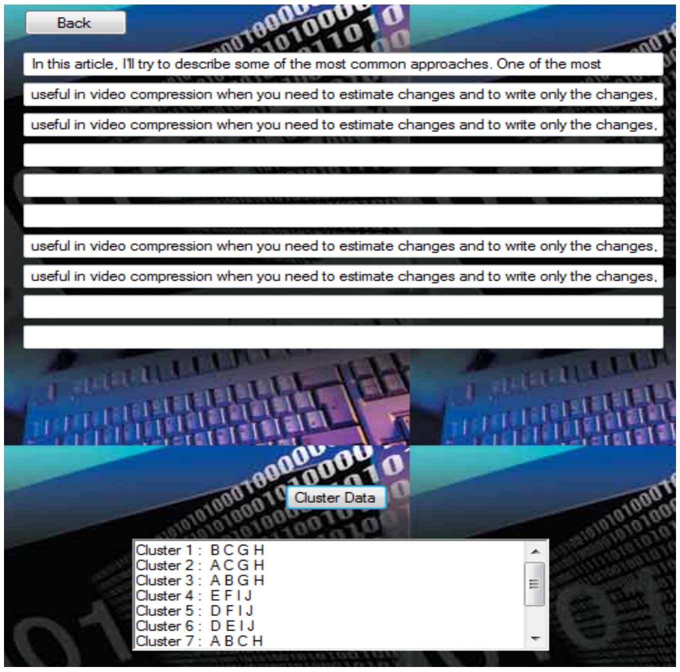
The following snapshot represents the parsed data which is the initial step for indexing [6] the data.


There are many approaches for motion detection in a continuous video stream.
All of them are based on comparing of the current video frame with one from the previous frames
In this article, I'll try to describe some of the most common approaches.
◆ AVI files (using Video for Windows, interop library is included);
updating JPEG from internet cameras MJPEG (motion JPEG) streams from different internet
local capture device (USB cameras or other capture devices, DirectShow interop library is
One of the most common approaches is to compare the current frame with the previous one
useful in video compression when you need to estimate changes and to write only the changes,
Assume that we have an original 24 bpp RGB image called current frame (image),
This is my dog.

: There : are : many : approaches : for : motion : detection : in : a : continuous : video : stream.
: All : of : them : are : based : on : comparing : of : the : current : video : frame : with : one : from
: In : this : article, : I'll : try : to : describe : some : of : the : most : common : approaches.
: ◆ AVI : files : (using : Video : for : Windows, : interop : library : is : included);
: updating : JPEG : from : internet : cameras : MJPEG : (motion : JPEG) : streams : from :
: local : capture : device : (USB : cameras : or : other : capture : devices, : DirectShow : interop
: One : of : the : most : common : approaches : is : to : compare : the : current : frame : with : the
: useful : in : video : compression : when : you : need : to : estimate : changes : and : to : write :
: Assume : that : we : have : an : original : 24 : bpp : RGB : image : called : current : frame :
: This : is : my : dog.

Figure 6 Given & Parsed Data

b) *Clustered data:*



In this article, I'll try to describe some of the most common approaches. One of the most
useful in video compression when you need to estimate changes and to write only the changes,
useful in video compression when you need to estimate changes and to write only the changes,
useful in video compression when you need to estimate changes and to write only the changes,
useful in video compression when you need to estimate changes and to write only the changes,

Cluster 1 : B C G H
Cluster 2 : A C G H
Cluster 3 : A B G H
Cluster 4 : E F I J
Cluster 5 : D F I J
Cluster 6 : D E I J
Cluster 7 : A B C H

Figure 7 Clustered Data

The data is now clustered according to the similarity of words. The above given figure shows the clusters created that are created after matching the similarity of document with each other.

d. Results:

a) Graph Representing the Created Clusters at First Level:

At first level, less no of clusters are created. As at first level the similarity between two documents is less.

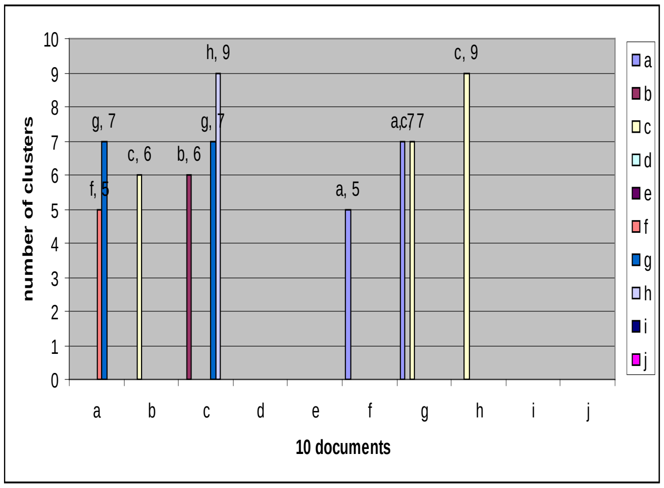


Figure 8 Clusters Created at First Level

b) Graph Representing the Created Clusters at Second Level:

At second level, more no of clusters are created in comparison to first level as the similarity between two documents is more.

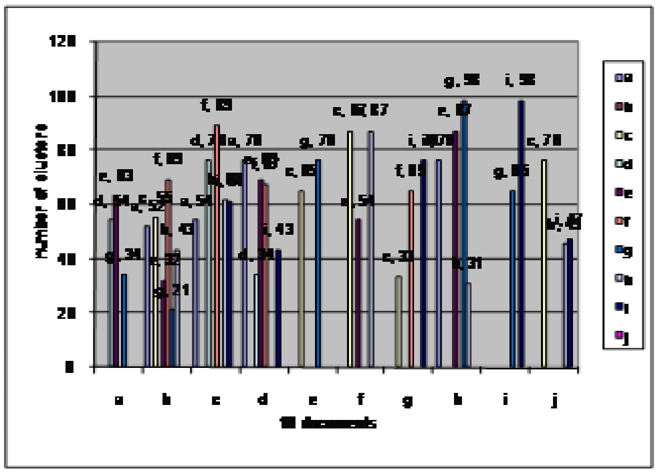


Figure 9 Clusters Created at Second Level

The results obtained with from the base version [7] approach are as under:

Column A = Contain experiment number.
Column B = Contain the updated page(s).
Column C = Contain URL of pages visited by crawler
Column D = Contain the start time of Crawler (Millisecond)
Column E = Contain the time to reach that page. (Millisecond)
Column F = Time spend to visit particular page (Millisecond)

A	B	C	D	E	F
1	Index	http://localhost:9254/CrawlWeb/update.jsph	1287490380435	1287490380669	234
		http://localhost:9254/CrawlWeb/index.jsp		1287490380747	312
2	P1	http://localhost:9254/CrawlWeb/update.jsp	1287490577650	1287490577837	187
		http://localhost:9254/CrawlWeb/P1.jsp		1287490577915	265
3	P23	http://localhost:9254/CrawlWeb/update.jsp	1287490730645	1287490730817	172
		http://localhost:9254/CrawlWeb/p23.jsp		1287490730895	250
4	P11 and P23	http://localhost:9254/CrawlWeb/update.jsp	1287477109254	1287477109426	172
		http://localhost:9254/CrawlWeb/p11.jsp		1287477109722	468
		http://localhost:9254/CrawlWeb/p23.jsp		1287477109987	733
5	P11, P22 and P33	http://localhost:9254/CrawlWeb/update.jsp	1287477421879	1287477422237	358
		http://localhost:9254/CrawlWeb/p33.jsp		1287477422549	670
		http://localhost:9254/CrawlWeb/p22.jsp		1287477422783	904
		http://localhost:9254/CrawlWeb/p11.jsp		1287477422846	967

Figure 10 Old Crawler timings

A	B	C	D	E	F
1	Index	http://localhost:9254/CrawlWeb/update.jsp	1287490380435	1287490380639	204
		http://localhost:9254/CrawlWeb/index.jsp		1287490380727	292
2	P1	http://localhost:9254/CrawlWeb/update.jsp	1287490577650	1287490577827	167
		http://localhost:9254/CrawlWeb/P1.jsp		1287490577890	240
3	P23	http://localhost:9254/CrawlWeb/update.jsp	1287490730645	1287490730787	142

		http://localhost:9254/CrawWeb/p23.jsp		1287490730875	230
4	P11 and P23	http://localhost:9254/CrawWeb/update.jsp	1287477109254	1287477109386	132
		http://localhost:9254/CrawWeb/p11.jsp		1287477109702	448
		http://localhost:9254/CrawWeb/p23.jsp		1287477109947	693
5	P11, P22 and P33	http://localhost:9254/CrawWeb/update.jsp	1287477421879	1287477422197	318
		http://localhost:9254/CrawWeb/p33.jsp		1287477422509	630
		http://localhost:9254/CrawWeb/p22.jsp		1287477422753	874
		http://localhost:9254/CrawWeb/p11.jsp		1287477422816	967

Figure 11 Proposed Crawler timings

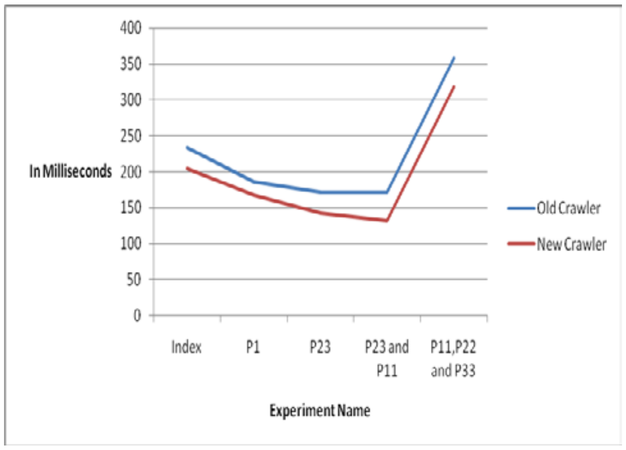


Figure 12 Graphical Time Difference

e. Components List and description:

a) Home Page

First Screen of our implemented project , It provide the Links to move to another screens of our project as well as it displays the data in the form of table which contains indexed data from our proposed algorithm.

b) Crawl_Old.cs:

Another screen, i.e Crawl Old which following the base algorithm of our project, it crawls the websites on the basis of old algorithm and display the data in the form of table like structure.

c) Crawl_New.cs

Another screen, i.e Crawl New This Screen Performs crawling on the basis of our proposed algorithm, it splits the different keywords and initializes them root and stores in the database.

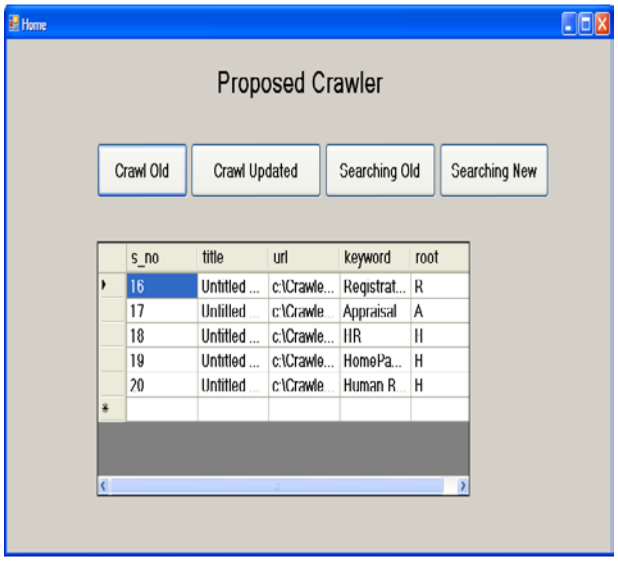
d) Searching.cs

Searching Screen is the form in which we can place our query keywords and perform searching on them. Searching will be done using our proposed algorithm it reduces the time of searching.

e) Searching Old.cs

Searching old screen is based on searching data gathered by base algorithm. It searches the complete database to search the keywords. That's why it take more time.

f. Snapshots of Proposed Crawler Application:



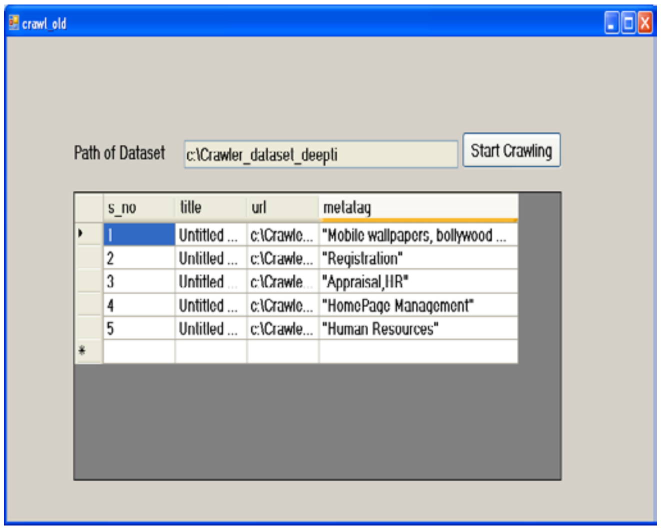
s_no	title	url	keyword	root
16	Untitled ...	c:\Crawle...	Registrat...	R
17	Untitled	c:\Crawle	Appraisal	A
18	Untitled ...	c:\Crawle...	HR	H
19	Untitled ...	c:\Crawle...	HomePa...	H
20	Untitled	c:\Crawle	Human R	H

Figure 13 Home Page

Navigation Details

The below section describes the navigation of all the screens for the 'Proposed Crawler application'

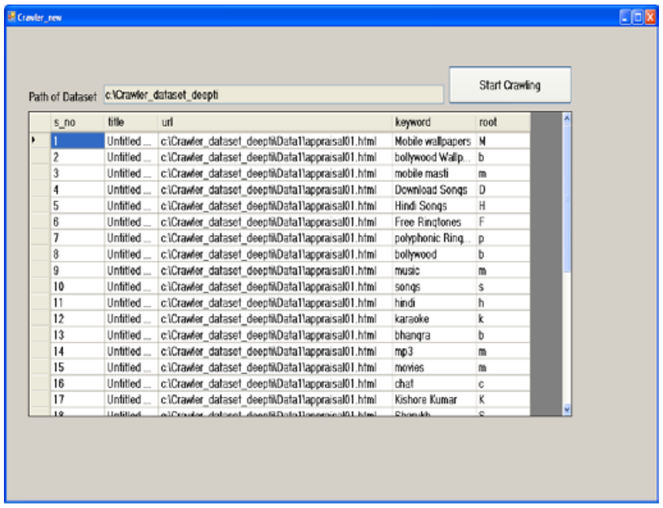
a) Crawl Old → Crawl Old.cs (Base Crawler Page)
b) Crawl New → Crawl New.cs (Proposed Crawler Page)
c) Searching Old → Searching Old.cs (Base Crawler Searching Page)
d) Searching New → Searching New.cs (Proposed Crawler Searching Page)



Path of Dataset: c:\Crawler_dataset_deepti

s_no	title	url	metatag
1	Untitled ...	c:\Crawle...	"Mobile wallpapers, bollywood ...
2	Untitled ...	c:\Crawle...	"Registration"
3	Untitled	c:\Crawle	"Appraisal,HR"
4	Untitled ...	c:\Crawle...	"HomePage Management"
5	Untitled ...	c:\Crawle...	"Human Resources"

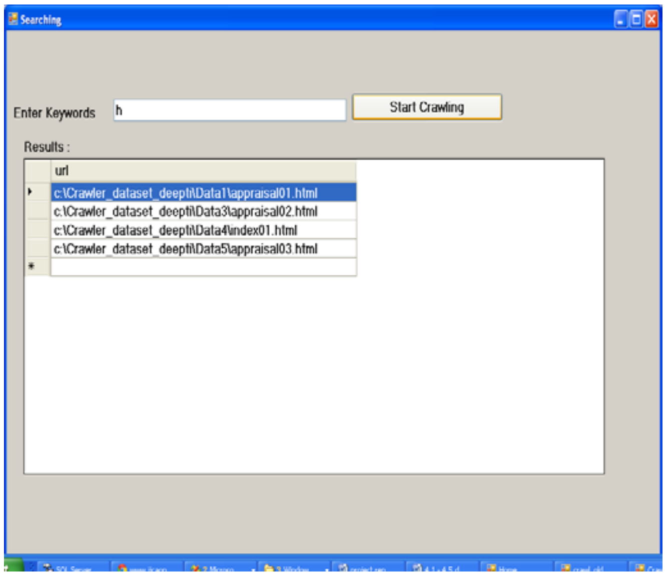
Figure 14 Base Crawler



Path of Dataset: c:\Crawler_dataset_deepti

s_no	title	url	keyword	root
1	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	Mobile wallpapers	M
2	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	bollywood Wallp...	b
3	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	mobile masti	m
4	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	Download Songs	D
5	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	Hindi Songs	H
6	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	Free Ringtones	F
7	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	polyphonic Ring...	p
8	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	bollywood	b
9	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	music	m
10	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	songs	s
11	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	hindi	h
12	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	karaoke	k
13	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	bhangra	b
14	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	mp3	m
15	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	movies	m
16	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	chat	c
17	Untitled ...	c:\Crawler_dataset_deepti\Data1\appraisal01.html	Kishore Kumar	K

Figure 15 Proposed Crawler

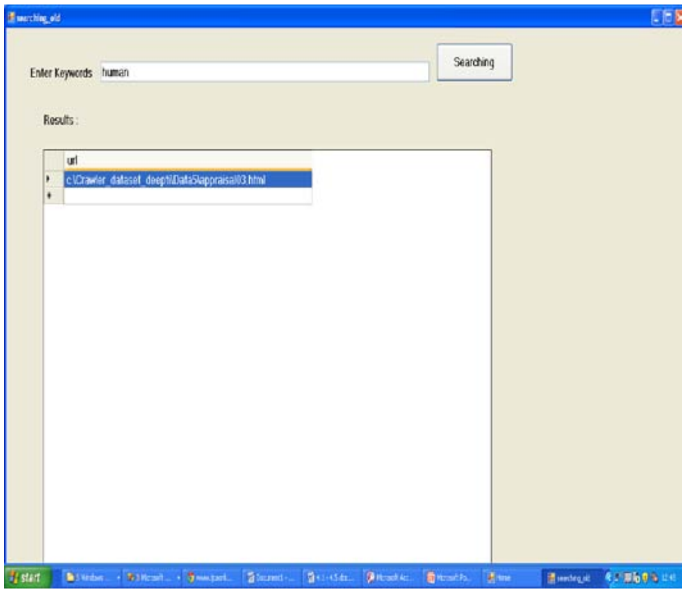


Enter Keywords: h

Results :

url
c:\Crawler_dataset_deepti\Data1\appraisal01.html
c:\Crawler_dataset_deepti\Data3\appraisal02.html
c:\Crawler_dataset_deepti\Data4\index01.html
c:\Crawler_dataset_deepti\Data5\appraisal03.html

Figure 16 Searching



Enter Keywords: human

Results :

url
c:\Crawler_dataset_deepti\Data5\appraisal03.html

Figure 17 Search Results

C. Context Based Index:

This is the indexing structure that contains the context and hence the document identifiers of the documents belonging to that context. This context represents the theme of the document that has been extracted using context repository, thesaurus and ontology repository [8]. This is the final index that is constructed after extracting the context of the document. Rather than being formed on the term basis, the index is constructed on the context basis with context as first field and finally the document identifiers of the relevant documents

In construction of context based index, once the document preprocessing is complete, the term with the maximum frequency matched with the title is extracted from the document. Then the maximum frequency keyword is being searched in the thesaurus (thesaurus can be taken online from thesaurus.com) and the context repository.

This step helps in extracting the context of the document but a keyword may have multiple contexts. So the multiple contexts are extracted. Now the next step is to extract the specific context of the document from these multiple contexts. The multiple contexts and the terms of the document are compared with the ontology repository. Thus by matching the keywords of the document and the multiple contexts with the concepts and the relationship terms in the ontology repository, the context of the document gets extracted. Now the posting list in the index consist of two columns, the one containing the context, the second one containing the lists of documents that contain the term with that specific context.

D. Ranked Context based Index:

The search engine provides results in response to the user queries. The query posed by the user is matched against the terms in the index and hence the relevant documents are retrieved in response to the user's query. However the returned documents may contain different order of relevance with the query. Thus the returned results are ranked according to their relevance with the user's request. So to attain the maximum level of the document relevancy ranking, ontology is applied to the ranking procedure. Moreover a pre ranking [9] of documents done beforehand may be useful in reducing the time taken to rank the documents after being searched in response to user query. The pre ranking is done by matching the keywords in the documents under the same context with the terms extracted from the domain and sub domain of the context of the document. The query fired by the user is expanded using the query expansion module.Query expansion (QE) is the process of reformulating a seed query to improve retrieval performance in information retrieval operations. In the context of web search engines, query expansion involves evaluating a user's input and expanding the search query to match additional documents. Query expansion involves techniques such as searching for the synonyms as well finding all the various morphological forms of words by stemming each word in the search query. It also deals in fixing spelling errors and automatically searching for the corrected form or suggesting it in the results. The terms in the expanded query are then matched one by one with documents keywords and the document with the highest matching value is being ranked the highest. In this way, the post rank of the documents is calculated and hence final rank is calculated as the sum of pre rank and post rank.

E. Query Expansion based on mining User Logs:

Query expansion is useful in reducing this query/document mismatch by expanding the query using words or phrases with a similar meaning or some other statistical relation to the set of relevant documents. An ontology driven conjunctive query expansion based approach on mining user logs has been proposed. In this, the clustered indexed documents and information extracted from user log are mined. Also, the association rules and ontology have been used for inferring rules and for identifying the relationship between the concepts. Query expansion has improved the effectiveness of ranked retrieval by automatically adding additional terms to a query. It attempts to overcome mismatch in query and document vocabularies by adding terms with similar statistical properties to those in the original query. Conjunctive queries can express a large part of queries, which are frequently issued on relational databases. In this query log lists the time that each search occurred, the IP address of the web user performing the search, the number of hits for the search, and the user's query. Moreover ontology has been used to identify relationship between various concepts. This approach proves to be more efficient than existing approaches because the query expansion module infer rules both from the user query log and clustered documents. Also, the performance has been proved to be better as compared to various existing search engines.

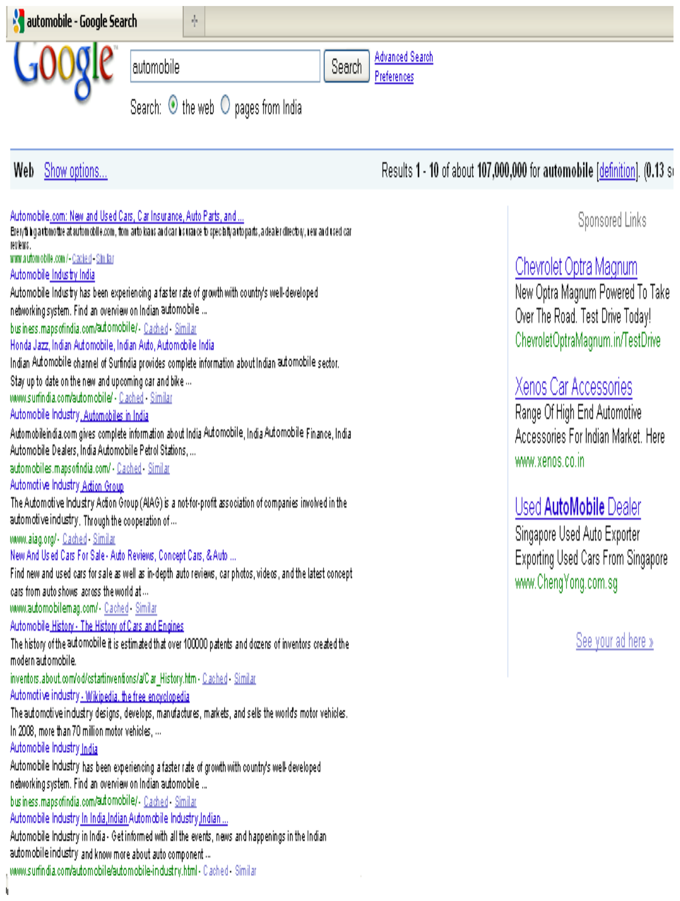
F. Context Ontology driven relevant search Using data Mining Techniques:

Query given by user is saved in database and then data mapper [10] maps this query into appropriate cells. For example :- Initially a URL or any .gif , .bmp etc. graphic image is there, then no need to create its ontology, directly query executes on the query processor from database. Otherwise stemming and stop listing is done on the query given by user and keywords are retrieved from the query.

Then the rules are formed from these keywords using the context repository. For eg. if mouse is a keyword, then rules are formed with animal as well as computer. After formation of rules, inference is retrieved in knowledge base for all rules and then ontology builder checks whether the ontology exists for these keywords or not. If yes then changes are made and if no, then ontology is created and saved in ontology repository. After that SQL query is formed from the keywords as Select mouse '||' animal from database and query processor executes this query on database using its ontology repository. The architecture is mainly divided into two phases. The initial phase i.e. passive phase is related to data mining process including data mapper, rule formulator and extraction of knowledge. The ontology-building phase is concerned with the process of building the ontology from the extracted knowledge. Finally, the SQL query is matched with the context and ontology repository and useful documents are returned.

a. Results:

Extensive experiments have been conducted over 2 search engines like Google, Alta Vista and compared with the proposed approach and it has been assumed that the our approach results in best performance as compared to Google and Alta Vista. This can be shown with the help of the following graph in Figure.



Results **1 - 10** of about **107,000,000** for **automobile** [definition]. (**0.13** s

Automobile.com: New and Used Cars, Car Insurance, Auto Parts, and ...
Everything automotive at automobile.com, from auto loans and car insurance to specialty auto parts, a dealer directory, new and used car reviews.
www.automobile.com/ - Cached - Similar

Automobile Industry India
Automobile Industry has been experiencing a faster rate of growth with country's well-developed networking system. Find an overview on Indian automobile ...
business.mapsofindia.com/automobile/ - Cached - Similar

Honda Jazz, Indian Automobile, Indian Auto, Automobile India
Indian Automobile channel of Surfindia provides complete information about Indian automobile sector. Stay up to date on the new and upcoming car and bike ...
www.surfindia.com/automobile/ - Cached - Similar

Automobile Industry, Automobiles in India
Automobileindia.com gives complete information about India Automobile, India Automobile Finance, India Automobile Dealers, India Automobile Petrol Stations, ...
automobiles.mapsofindia.com/ - Cached - Similar

Automotive Industry Action Group
The Automotive Industry Action Group (AIAG) is a not-for-profit association of companies involved in the automotive industry. Through the cooperation of ...
www.aiag.org/ - Cached - Similar

New And Used Cars For Sale - Auto Reviews, Concept Cars, & Auto ...
Find new and used cars for sale as well as in-depth auto reviews, car photos, videos, and the latest concept cars from auto shows across the world at ...
www.automobilemag.com/ - Cached - Similar

Automobile History - The History of Cars and Engines
The history of the automobile it is estimated that over 100000 patents and dozens of inventors created the modern automobile.
inventors.about.com/od/cstartinventions/a/Car_History.htm - Cached - Similar

Automotive industry - Wikipedia, the free encyclopedia
The automotive industry designs, develops, manufactures, markets, and sells the world's motor vehicles. In 2008, more than 70 million motor vehicles, ...

Automobile Industry India
Automobile Industry has been experiencing a faster rate of growth with country's well-developed networking system. Find an overview on Indian automobile ...
business.mapsofindia.com/automobile/ - Cached - Similar

Automobile Industry In India,Indian Automobile Industry,Indian ...
Automobile Industry in India - Get informed with all the events, news and happenings in the Indian automobile industry and know more about auto component ...
www.surfindia.com/automobile/automobile-industry.html - Cached - Similar

Sponsored Links

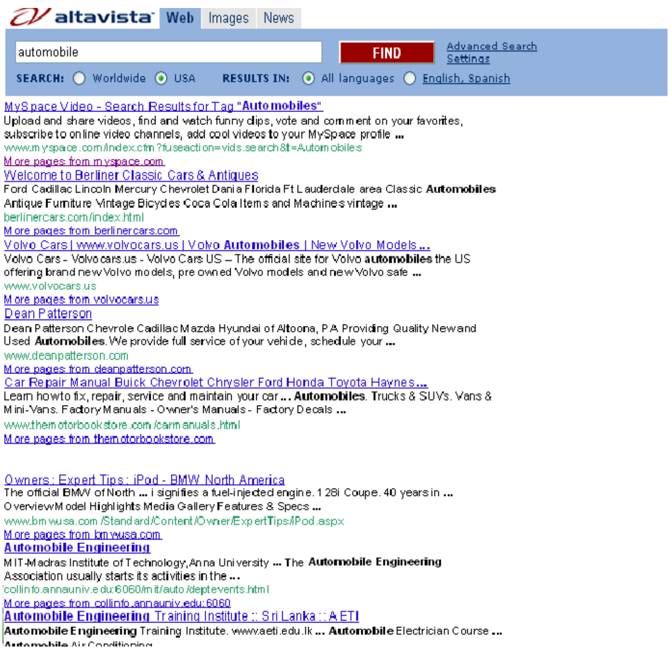
Chevrolet Optra Magnum
New Optra Magnum Powered To Take Over The Road. Test Drive Today!
ChevroletOptraMagnum.in/TestDrive

Xenos Car Accessories
Range Of High End Automotive Accessories For Indian Market. Here
www.xenos.co.in

Used **AutoMobile** Dealer
Singapore Used Auto Exporter Exporting Used Cars From Singapore
www.ChengYong.com.sg

See your ad here »

Figure 18 Google Search Results



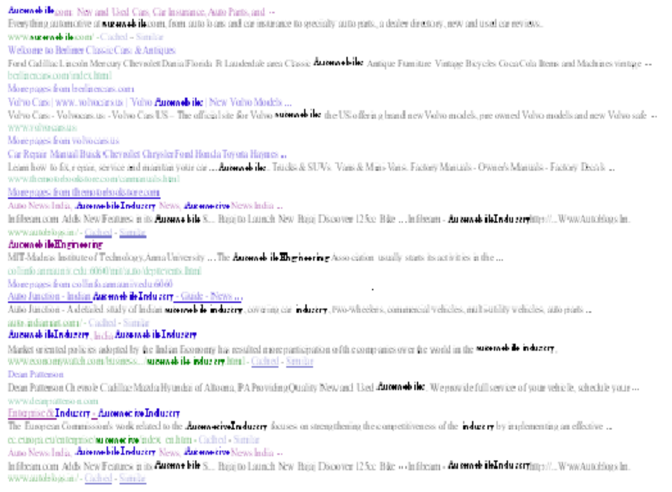
automobile

SEARCH: Worldwide [✓ USA] RESULTS IN: [✓ All languages] English, Spanish

MySpace Video - Search Results for Tag "**Automobiles**"
Upload and share videos, find and watch funny clips, vote and comment on your favorites, subscribe to online video channels, add cool videos to your MySpace profile **...**
www.myspace.com/index.cfm?fuseaction=vids.search&t=Automobiles
More pages from myspace.com
Welcome to Berliner Classic Cars & Antiques
Ford Cadillac Lincoln Mercury Chevrolet Dania Florida Ft Lauderdale area Classic **Automobiles** Antique Furniture Vintage Bicycles Coca Cola Items and Machines vintage **...**
berlinercars.com/index.html
More pages from berlinercars.com
Volvo Cars | www.volvocars.us | Volvo **Automobiles** | New Volvo Models **...**
Volvo Cars - Volvocars.us - Volvo Cars US – The official site for Volvo **automobiles** the US offering brand new Volvo models, pre owned Volvo models and new Volvo safe **...**
www.volvocars.us
More pages from volvocars.us
Dean Patterson
Dean Patterson Chevrole Cadillac Mazda Hyundai of Altoona, PA Providing Quality New and Used **Automobiles**. We provide full service of your vehicle, schedule your **...**
www.deanpatterson.com
More pages from deanpatterson.com
Car Repair Manual Buick Chevrolet Chrysler Ford Honda Toyota Haynes **...**
Learn how to fix, repair, service and maintain your car **... Automobiles**. Trucks & SUVs. Vans & Mini-Vans. Factory Manuals - Owner's Manuals - Factory Decals **...**
www.themotorbookstore.com/carmanuals.html
More pages from themotorbookstore.com

Owners : Expert Tips : iPod - BMW North America
The official BMW of North **...** i signifies a fuel-injected engine. 128i Coupe. 40 years in **...** Overview Model Highlights Media Gallery Features & Specs **...**
www.bmwusa.com/Standard/Content/Owner/ExpertTips/iPod.aspx
More pages from bmwusa.com
Automobile Engineering
MIT-Madras Institute of Technology,Anna University **...** The **Automobile Engineering** Association usually starts its activities in the **...**
collinfo.annauniv.edu:6060/mit/auto/deptevents.html
More pages from collinfo.annauniv.edu:6060
Automobile Engineering Training Institute :: Sri Lanka :: AETI
Automobile Engineering Training Institute. www.aeti.edu.lk **... Automobile** Electrician Course **...**
Automobile Air Conditioning

Figure 19 Alta Vista Search Results



Automobile.com: New and Used Cars, Car Insurance, Auto Parts, and ...
Everything automotive at **automobile**.com, from auto loans and car insurance to specialty auto parts, a dealer directory, new and used car reviews.
www.**automobile**.com/ - Cached - Similar
Welcome to Berliner Classic Cars & Antiques
Ford Cadillac Lincoln Mercury Chevrolet Dania Florida Ft Lauderdale area Classic **Automobiles** Antique Furniture Vintage Bicycles Coca Cola Items and Machines vintage ...
berlinercars.com/index.html
More pages from berlinercars.com
Volvo Cars | www.volvocars.us | Volvo **Automobiles** | New Volvo Models ...
Volvo Cars - Volvocars.us - Volvo Cars US – The official site for Volvo **automobiles** the US offering brand new Volvo models, pre owned Volvo models and new Volvo safe ...
www.volvocars.us
More pages from volvocars.us
Car Repair Manual Buick Chevrolet Chrysler Ford Honda Toyota Haynes ...
Learn how to fix, repair, service and maintain your car ... **Automobiles**. Trucks & SUVs. Vans & Mini-Vans. Factory Manuals - Owner's Manuals - Factory Decals ...
www.themotorbookstore.com/carmanuals.html
More pages from themotorbookstore.com
Auto News India, **Automobile Industry** News, **Automotive** News India ...
Infibeam.com Adds New Features in its **Automobile** S... Bajaj to Launch New Bajaj Discover 125cc Bike ... Infibeam - **Automobile Industry** http://...WwwAutoblogs.In.
www.autoblogs.in/ - Cached - Similar
Automobile Engineering
MIT-Madras Institute of Technology,Anna University ... The **Automobile Engineering** Association usually starts its activities in the ...
collinfo.annauniv.edu:6060/mit/auto/deptevents.html
More pages from collinfo.annauniv.edu:6060
Auto Junction - Indian **Automobile Industry** - Guide - News ...
Auto Junction - A detailed study of Indian **automobile industry**, covering car **industry**, two-wheelers, commercial vehicles, multi-utility vehicles, auto parts ...
auto.indiamart.com/ - Cached - Similar
Automobile Industry, India **Automobile Industry**
Market oriented policies adopted by the Indian Economy has resulted more participation of the companies over the world in the **automobile industry**.
www.economywatch.com/business.../**automobile-industry**.html - Cached - Similar
Dean Patterson
Dean Patterson Chevrole Cadillac Mazda Hyundai of Altoona, PA Providing Quality New and Used **Automobiles**. We provide full service of your vehicle, schedule your ...
www.deanpatterson.com
Enterprise & **Industry** - **Automotive Industry**
The European Commission's work related to the **Automotive Industry** focuses on strengthening the competitiveness of the **industry** by implementing an effective ...
ec.europa.eu/enterprise/**automotive**/index_en.htm - Cached - Similar
Auto News India, **Automobile Industry** News, **Automotive** News India ...
Infibeam.com Adds New Features in its **Automobile** S... Bajaj to Launch New Bajaj Discover 125cc Bike ... Infibeam - **Automobile Industry** http://...WwwAutoblogs.In.
www.autoblogs.in/ - Cached - Similar

Figure 20 Proposed System Search Results

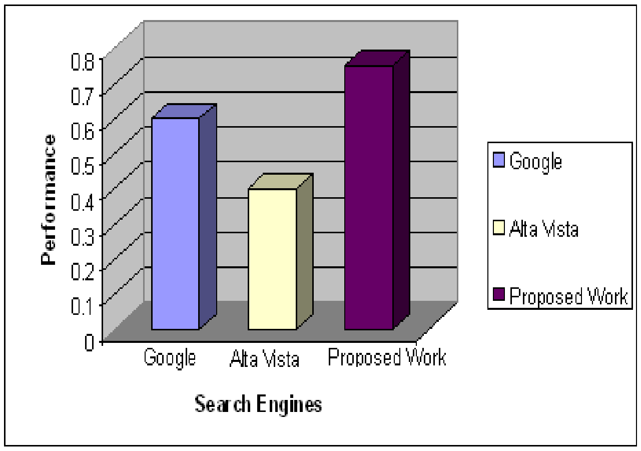


Figure 21 Performance Comparisons

IV. CONCLUSION

This paper discusses the framework for multilevel as well as context based indexing and also summarizes the use of context-ontology along with data mining technique for relevant searching and ranking of web documents. Ontology plays a vital role in ensuring the homogeneity of information through the sharing of semantics. Moreover ontology has been used to identify relationship between various concepts. This approach proves to be more efficient than existing approaches. Also, the performance has been proved to be better as compared to various existing search engines.

V. REFERENCES

[1] S.Brin and L.Page "The anatomy of a large scale hyper textual web search engine" In the 7th International WWW Conference, (WWW7), pp. 107117, Brisbane, Australia

[2] Van Rijsbergen C.J. Information retrieval. Butterworth 1979.

[3] Davies, J., Fensel, D., and van Harmelen, F., Towards the semantic web: ontology-driven knowledge management, John Wiley and Sons, Ltd., 2002.

[4] Parul Gupta, Dr. A.K.Sharma,"A framework for hierarchical clustering-based indecxing in search engines" International Conference on Parallel, Distributed and Grid Computing Conference (PDGC_2010), IEEE Xplore held at JUIT, Solan.

[5] Parul Gupta, Dr. A.K.Sharma, International Journal of Applied Engineering Research "A framework for multilevel indexing in search engines", Vol 4, No.8, pp 1423-1429, 2009.

[6] Parul Gupta, Dr. A.K.Sharma, International Journal of Applied Engineering Research "Implementation of multilevel indexing in search engines", Vol 6, No.15, pp 1873-1882, 2011.

[7] Deepika Sharma, Parul Gupta, Dr. A.K.Sharma, International Journal of Computer Applications "Crawler indexing using tree structure and its implementation", Vol 31, No.6, pp 34-39, 2011.

[8] Parul Gupta, Dr. A.K.Sharma, "Context based indecxing in search engines using ontology", International Conference on Futuristic Computer Applications (ICFCA 2010).

[9] Parul Gupta, Dr. A.K.Sharma, "Ontology driven pre and post ranking based information retrieval in web search engines", accepted in IJCSE, "in press".

[10] Parul Gupta, Nisha Pahal, Payal Gulati, "Context ontology driven relevant search using data mining techniques" International Conference on Methods and Models in Computer Science (ICM2CS 2010) held at JNU, Delhi (13th - 14th Dec, 2010).

[11] Parul Gupta, Dr. A.K.Sharma, "Indexing in search engines- a review", International Conference on Information Technolgy ICIT 09 held at PDM, Bahadurgarh, 18th – 19th June, 2009.