Journal of Theoretical and Applied Information Technology
31st January 2015. Vol.71 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN EEFICIENT MANAGEMENT OF SEMANTIC METADATA REPOSITORIES USING TEMPORAL BASED KALMAN FILTER ON WEB ARCHITECTURE

[1]A.ABDUL RAHUMAN, [2]Dr.C.CHANDRASEKAR
[1]Research scholar, Department of Computer Science, Karpagam University, Coimbatore, Tamil Nadu, India
[2]Associate Professor and Reader, Department of Computer Science, Periyar University, Salem, Tamil Nadu, India.
E-mail: [1]a.r.rahuman@gmail.com, [2]ccsekar@gmail.com

ABSTRACT

A major concern incorporating data accesses for web search classification is the need for accurate user result, optimality solution on semantic ontology search with combination of keyword. These concerns often make it difficult to produce results relevant to the user query with maximal metadata repositories. Moreover, with multiple pages produced as a result of current standard search on web domain, harness the augmented filtering potential relevant to the user query. To present an effective metadata managing repositories on web domain, Temporal-based Kalman Filter on Web Architecture (TKF-WA) is proposed in this paper. TKF-WA addresses a metadata based semantic ontology search on web domain. Initially, semantic learning object link class is developed to classify the set of keywords automatically for effective management of metadata repositories. The metadata links the related learning objects internally and manages maximal metadata repositories. The managed metadata repositories on Web Architecture present a Temporal-based Kalman Filtering algorithm to remove the irrelevant pages which are not related or insignificant to the user query. Temporal-based Kalman Filtering performs the series of measurement based on time with metadata repositories to satisfy the user query set. Kalman Filtering is temporal based in the proposed web architecture to not only handle the regular (i.e., sentence-based) but also the irregular (i.e., keyword-based) query set. Experiment conducted on IRIS2 dataset with factors such as system utility rate on filtering, search retrieval efficiency and average response time.

Keywords: *Temporal based Kalman Filter Design, Semantic Metadata Repositories, Ontology, Web Architecture, Semantic Learning Object, User Query*

1. INTRODUCTION

Separation of keyword from query is one of the most interesting and important topics in web search mining. The aim of web search mining is to find the relevant pages significant to the user query in web structure, and organize them into related learning objects for further study and analysis. There have been many web search algorithms published so far. Satisfy User Profiles (SUP) [1], with the aid of prior constraints, potentially maximized the user utility. To achieve this, optimality conditions were derived through which effectively utility improvement and probe rate of filtering were increased with little bit of additional resources. However, with the increase in user query, satisfaction of user profiles was not addressed.

Database-driven Reverse Dictionary (DRD) [2] with the help of algorithmic design not only provided scalability of records with minimum response time but also improved the quality of results produced. But return of relevant queries with maximal data repositories remained unaddressed. Automatic Discovery of Personal Name using Lexical Pattern-based (ADPN-LP) [3] approach efficiently extracted candidate item sets from snippets. Also the information retrieval rate was improved using ranking support vector machines. But, with, multiple pages produced as output, filtering was affected and as a result, relevancy of user query was not derived.

To develop efficient applications designed in a scattered way, web services can be used with internet and business intranets. QoS parameter for Web service composition was addressed in [4] due

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to the difficulty in organization. But scalability during runtime remained unaddressed. Bayesian learning and Expectation Maximization was introduced in [5] to address the issues related to scalability by minimizing the human intervention based on the extraction pattern. The efficiency of extraction was achieved only with single database.

Unsupervised Duplicate Detection (UDD) [6] measured the presence of duplicates in query results from multiple databases. With this, the record matching efficiency was improved. Though precision and recall value were improved, concepts of the sentence were least considered than the frequency captures. Concept based mining model was designed in [7] with the aim of considering sentence concept rather than the frequency with which the query appears. With the aid of concept based similarity measure, precision and recall were improved.

With the increase in web based information, one of the most challenging and significant tasks of the user is how to obtain the most relevant information given the query as input. Personalized ontology model [8] was designed for representation of knowledge and the basics behind the concept over user profiles providing both world knowledge and from the repositories. Automatic Template Extraction (TEXT) [9] effectively separated unknown queries in the form of clusters and template extraction was performed to improve the scalability. However, the method was not full proof with respect to security. Single server and two round solution was introduced in [10] to address the issues related to security. The method used matrix additive coding for efficient encryption. Also the user relevant results were increased with the aid of searchable encryption technique.

Our contributions are in four folds. (i) To present an effective metadata based semantic ontology search and manage repositories on web domain using Temporal-based Kalman Filter on Web Architecture (TKF-WA), (ii) To perform effective management of metadata repositories for keywords classification using semantic learning object link class, (iii) To manage maximal metadata repositories based on time by linking metadata with the related learning objects to satisfy user query set and (iv) Design of Kalman Filtering to not only handle the regular (i.e., sentence-based) but also the irregular (i.e., keyword-based) query set and to remove the irrelevant pages that are not related to the user query with the help of Temporal-based Kalman filtering algorithm.

The paper is organized as follows: Section 2 provides related research with respect to user query based on web search engine. In Section 3 the detailed explanation of Temporal-based Kalman Filter on Web Architecture (TKF-WA) is described with the help of a neat framework using algorithmic description. Section 4 provides a detail description about the experimental setup required to design TKF-WA. Section 5 discusses in detail about the experimental settings provided for TKF-WA. Finally, Section 6 concludes the work.

2. RELATED WORKS

Certain results obtained using the commercial search engines does not consider the interest of the user and as a result same results are retrieved. Search engine personalization was introduced in [11] that not only included the positive preferences but also the negative preferences improving the precision and similarity of values being obtained. However, user interests were not ranked. Keyword and content based image retrieval was introduced in [12] to improve the retrieval accuracy and rank them according to the qualitative nature. Model Based on Cocitation (MBCC) [13] used ranking as the base to obtain uniqueness of page being ranked.

Semantic Web is widely used as a phenomenon to determine and satisfy the user requests in order to efficiently use the web content. Rule based method was introduced in [14] that used glossaries with the aid of inference method. It not only increased the coverage but also the recall rate. However, distributed nature of data was unaddressed. Two level cache structures were structured in [15] to address the issues related to distributed nature of information retrieval using distribution strategy. Though hit rate was improved, with the repeated occurrences, information retrieval rate was minimized. Multimedia search engines may contain occurrences of repeated events that results in minimizing the efficiency of multimedia search. In [16], hash algorithms were introduced to improve the efficiency of multimedia search.

A web tool called as GeoTxt [17], addressed three important issues related to extraction, disambiguation, and geolocation of places in an unstructured text format. This significantly increased microblog posts. Bias in

Journal of Theoretical and Applied Information Technology
31st January 2015. Vol.71 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

retrieval system was addressed in [18] with the aid of language modeling and term selection value for query expansion. This in turn improved the entropy value.

On the basis of above said methods and techniques, a design view of temporal based kalman filter on web architecture for managing metadata repositories is provided in the section given below.

3. DESIGN OF TEMPORAL BASED KALMAN FILTER ON WEB FOR MANAGING METADATA REPOSITORIES

In this section, the design of temporal based kalman filter on web architecture for managing metadata repositories is presented with the aid of architecture diagram. The work starts with the description about semantic learning object link class. Next, the Temporal based Kalman filtering operation is described with the help of an algorithm.

The main objective of the proposed work is to manage the metadata repositories using semantic learning object link class and retrieve the user result through effective filtering method. The Temporal based Kalman Filtering is introduced to provide the foundation for a semantic ontology search engine and obtain the user result by removing multiple unrelated page results. The TKF-WA design on web architecture manages the metadata repository database, where the database contains the actual keywords. The related keywords are linked to one another in TKF-WA design using Semantic Learning Object Link Class. The Link class between the semantics describes the structure of keywords stored in the metadata repositories. The link class in the metadata repository on web domain is illustrated in Figure 1.

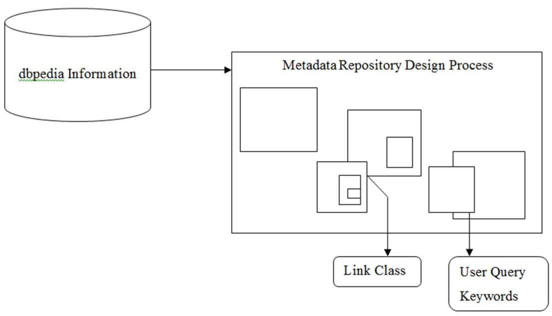

Figure 1: Metadata Repository Link Class

The figure 1 clearly describes the metadata repository design process and generates the objective repository framework through different set of semantic keywords. The information from dbpedia is used as a sample work in the design of TKF-WA design to develop the link classes. For instance, the user keyword is unique or related to other combination of keywords and fetches the accurate result. The keyword of one class is related to the other class of objects, and as a result the link is established in the metadata repository. The design of metadata repository in TKF-WA stores data about other data and ensure consistency for maximal set of class objects.

The design of TKF-WA web architecture filters specific user results from metadata repository by removing the irrelevant information. The removal of irrelevant information improves the precision rate in TKF-WA design. The TKF-WA design on web semantic domain filter out bad procedures by removing the semantic learning objects irrelevant links. The filtering operation in TKF-WA design is illustrated in Figure 2.

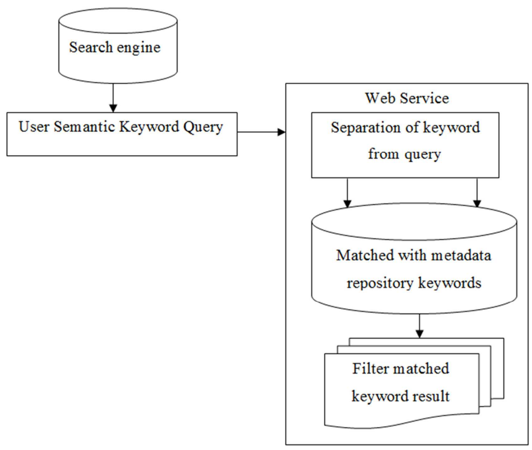

Figure 2: Design of filtering operation in TKF-WA

As illustrated in Figure 2, the filtering process obtains the user query with keywords for processing. The web service initially separates the keywords from the user query and then performs the matching operation. The matching operation matches the user keyword with the stored metadata repository keyword and filters out the accurate user result. Web Architecture with Temporal based Kalman Filter Design is illustrated in Figure 3.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

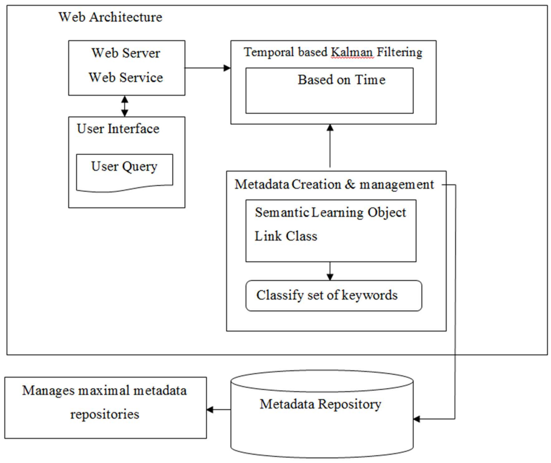

Figure 3: Design of Temporal based Kalman Filter on Web Architecture

As depicted in Figure 3, the design of temporal based kalman filter on web architecture contains the web server which offers the web services. The web server performs the process based on the request placed by the user through the http (hypertext transfer protocol). The basic network protocol is used to distribute the user request query keywords from user (i.e., client side) to the server side via search engine. The user request is matched with the metadata repository information. The metadata repository is created using the Semantic Learning Object Link Class. The semantic learning objects of different users are linked with the repository (i.e.,) database using the link class.

The design of TKF-WA classifies the keywords according to the overall class information. The linking of class in the metadata repositories helps to manage maximal number of semantic keywords which are used for performing filtering operation. The Temporal based Kalman filtering is used to filter user relevant keywords based on time. The temporal based Kalman filtering is developed because the web domain handles not only regular semantic keywords but also irregular semantic keywords. The forthcoming subsections discusses in detail about the metadata creation and management using semantic learning object link class. Managing of maximal metadata repositories using Temporal based Kalman filtering with respect to time is also discussed in detail.

3.1 Semantic Learning Object Link Class

An initial step is presented in TKF-WA design to classify the keywords in the metadata repository using the semantic learning object link class. The semantic keywords provided by the user are placed in the repository by adding the metadata feature. The metadata in TKF-WA design contains the common, technical and relational objects. These objects are subjected to keywords and related to attain effective classification on web architecture.

Learning Object =
{Common, Technical, Relational} (1)

Eqn (1) carries obtains the collection of objects to perform the classification of semantic keywords. The semantic keywords are interrelated with one class to another, and as a result the link class is created. Let us consider the keywords k1, k2, k3….kn based on the user query, where the keywords are separated based on the learning objects on web domain. The link class ‘c’ refers to the single keyword or group of keywords which are interrelated with each other. The keywords interrelated are link together and named with new class name. The Learning object link class with the semantic keyword is defined as,

Link Class =
{c1(k1),c2(k2,K3),c3(k3),…,cn(kn)} (2)

The design of TKF-WA generates the link class to classify the user keywords for effective creation of metadata repositories. The metadata holds different classes ‘c1, c2, c3…cn’. For instance, the keywords like k1 -> fruits, k2 -> apple, k3 -> orange are obtained and performs the class link using the semantic learning object class link. ‘k1’ is the common object, whereas the ‘k2’ and ‘k3’ are linked in a way such that they belong to the classification of fruits. In a similar manner, all the user query keywords are classified in using the design of TKF-WA in order to manage maximal metadata repositories.

3.2 Temporal based Kalman Filtering Operation

Once the user query keywords are classified, Temporal based Kalman filtering is used in the proposed work to filter the user query in a precise answer by removing the irrelevant result pages. The Kalman filtering handles only the regular statement or query issued on the side of user whereas Temporal based Kalman filtering handles both the regular and irregular user query. Kalman filter operates recursively with metadata repository information to attain the optimality solution. The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

algorithmic step involved in the design of Temporal based Kalman filtering is given below,

//Temporal based Kalman Filtering
Begin

Input: User query 'UQ' with set of keywords
Output: Filtered User results with precise answer

//Separation of User keywords
Step 1: While (Temporal based Kalman Filter estimates the regular and irregular user query)
Step 2: Regular and irregular keyword separated based on learning object frequency analysis
Step 3: Relative frequency of each keyword is measured based on time
Step 4: Once outcome measured, Goto step 6
Step 5: End While

//Matched with metadata repositories
Step 6: Regular and irregular keyword frequency used to identify the outcome of the link class
Step 7: Link class frequency of keyword is matched with user semantic keyword
Step 8: Repeat step for each user query

End

The Temporal based Kalman filtering algorithm works in two consecutive processes. The separation of user keyword set through frequency measure based on time. Frequency analysis is briefly explained in section 3.2.1. Once the outcome is measured, the next step is proceeded to perform match with the metadata repositories. The metadata repository matches the user keyword with the link class which holds the classified metadata keyword. As the Temporal based Kalman filtering is recursive in nature it handles different user query semantic search on web domain.

3.2.1 Separation of user keywords

The separation of different user query keywords in TKF-WA uses the frequency quantitative analysis measure. The frequency quantitative analysis measure provides with the different number of previously occurred search query term with learning objects and defined as,

$$Frequency\ Measure = \log \frac{(Current\ k_{t,\alpha} + Previous\ k_{t,\alpha})}{|\alpha|} + 1 \qquad (3)$$

From (2), we can observe that the frequency measure is arrived in a quantitative manner using the current and previous keywords based on time measure. 'α' is the number of terms in the user query. The frequency measure is used as the basis to match the result with the metadata repositories.

3.2.2 Matching with metadata repository

Once the frequency measure is obtained, the metadata repository is used on filtering the user correct result with class link. Matching operation is defined as,

$$M = \{(Frequency\ measure(UQ))\} = \{Link\ Class\ search\} \qquad (4)$$

Eqn (4) denotes the matching rate of semantic ontology search query on web. The relevant frequency value of the user is matched with the link class search to filter out the exact query result by removing the unrelated result pages.

4. EXPERIMENTAL EVALUATION

Temporal based Kalman Filter Design is developed in our proposed work to manage the metadata repositories. The experiment is conducted on JAVA platform using the IRIS2 dataset. Iris2 dataset is specifically used to assess the performance of filtering algorithms in semantic ontology search. The knowledge model is developed to manage the metadata repositories using the linked data principle. IRIS2 is a medium-sized dataset with high inner connectivity amongst entities and consists of instances related to scientific research domains. The scientific research on web domain are publications, authors, research topics, journals, conferences and publishers as well as their relationships cites, writes, broader, and co-author.

Experimental work also takes the information for managing of metadata repositories using dbpedia.org. dbpedia.org/ontology/spots team season, dbpedia.org/ontology/SportsSeason, and dbpedia.org/ontology/SportsTeamSeason are used. A season with respect to a specific sports team is related in the metadata repositories and the same is followed for the entire league. A season for the particular sports team is clearly specified with the subclasses. Temporal based Kalman Filter Design on Web Architecture (TKF-WA) for Managing Web Semantic Metadata Repositories is compared against the existing Satisfy User Profiles (SUP) [1], Database-driven Reverse Dictionary (DRD) [2] and

Journal of Theoretical and Applied Information Technology
31st January 2015. Vol.71 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Automatic Discovery of Personal Name using Lexical Pattern-based (ADPN-LP) [3] approach. The experiment is conducted on factors such as effective system utility rate based on precision, search retrieval efficiency based on recall, probe rate of filtering and average response time.

The system utility rate SUR using TKF-WA is obtained by measuring the ratio of relevant pages Rel_P to the retrieved pages Ret_P as given below

$$SUR = \frac{Rel_P \cap Ret_P}{Ret_P} \quad (5)$$

The search retrieval value SRV on TKF-WA is obtained using the ratio of pages that are relevant to the user query that are successfully retrieved.

$$SRV = \frac{Rel_P \cap Ret_P}{Rel_P} \quad (6)$$

The probe rate of filtering is obtained using the Temporal based Kalman filtering algorithm. It is measured in terms of percentage (%). The average response time ART in TKF-WA is obtained by the difference between the average time for obtaining previous keywords Average Time_{PK} and average time for obtaining current keywords Average Time_{CK}. It is measured in terms of milliseconds.

$$ART = (\,Average\ Time_{PK} - Average\ Time_{CK}) \quad (7)$$

5. RESULTS ANALYSIS OF TKF-WA

Temporal based Kalman Filter Design on Web Architecture (TKF-WA) is compared against the existing Satisfy User Profiles (SUP), Database-driven Reverse Dictionary (DRD) and Automatic Discovery of Personal Name using Lexical Pattern-based (ADPN-LP). Table 1 evaluates the System Utility Rate based on precision is measured in terms of percentage achieved with different number of probes ranging from 50 to 35- and comparison is made with the three existing schemes namely, Satisfy User Profiles (SUP) [1], Database-driven Reverse Dictionary (DRD) [2] and Automatic Discovery of Personal Name using Lexical Pattern-based (ADPN-LP) [3].

Table 1: Tabulation for system utility rate

Number of probes	System Utility Rate based on precision (%)			
	TKF-WA	SUP	DRD	ADPN-LP
50	58.32	53.30	48.28	42.27
100	65.45	60.43	55.41	49.40
150	68.73	63.71	58.69	52.68
200	72.45	67.43	62.41	56.40
250	66.35	61.33	56.31	50.30
300	75.85	70.83	65.81	59.80
350	78.88	73.86	68.84	62.83

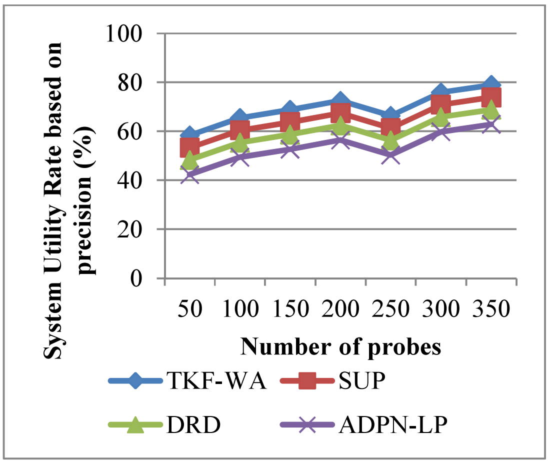

Figure 4 Measure of system utility rate

Figure 4 describes the system utility rate based on the number of probes on web architecture. The system utility rate for each number of probes placed by the user is measured with the relevant and retrieved pages on web architecture. As a result, the system utility rate based on precision using the TKF-WA is improved by effective filtering by 6 – 9 % compared to SUP [1]. Moreover with the removal of irrelevant information from metadata repository based on the specific user results, the result of precision rate is improved by 12 – 17 % and 20 – 27 % compared to DRD [2] and ADPN-LP [3] respectively.

Journal of Theoretical and Applied Information Technology
31st January 2015. Vol.71 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2 Tabulation For Search Retrieval Value Based On Recall

Number of probes	Search Retrieval Value based on recall (%)			
	TKF-WA	SUP	DRD	ADPN-LP
50	61.34	58.32	51.28	45.29
100	68.47	65.45	57.41	52.42
150	71.75	68.74	60.69	55.70
200	75.47	72.45	64.41	59.42
250	69.37	66.35	58.31	53.32
300	78.87	75.85	67.81	62.28
350	81.90	78.88	70.84	65.85

The Search Retrieval Value based on recall of our work TKF-WA with the existing two schemes namely Satisfy User Profiles (SUP) [1], Database-driven Reverse Dictionary (DRD) [2] and Automatic Discovery of Personal Name using Lexical Pattern-based (ADPN-LP) [3] is provided in table 2.

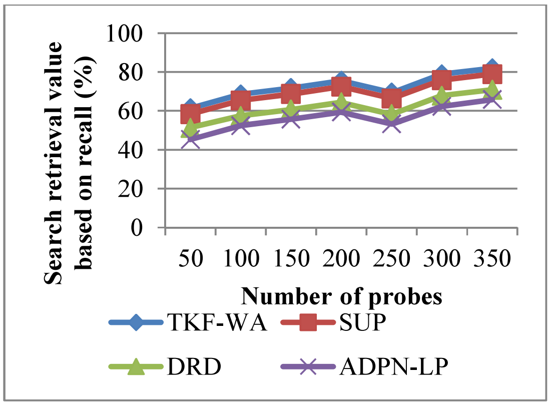

Figure 5 Measure Of Search Retrieval Value

Figure 5 illustrates the search retrieval value based on recall with respect to the probes issued by the user in the form of regular and irregular queries. The TKF-WA with the design of common, technical and relational objects is executed for a maximum number of probes provided by the user as query to effectively search and retrieve the value and therefore improve the search retrieval value by 3 – 4 % when compared with the SUP [1]. Moreover, with the application of semantic learning object link that effectively performs the classification of semantic keywords improves the search retrieval value by 13 – 16 % when compared to DRD [2]. With this, the semantic learning object class link generates the link class that classifies the user keywords using the metadata repositories resulting in the increased search retrieval values by 19 – 26 % compared to ADPN-LP [3].

Table 3 Tabulation For Probe Rate Of Filtering

Number of probes	Probe Rate of Filtering (%)			
	TKF-WA	SUP	DRD	ADPN-LP
50	0.235	0.211	0.200	0.190
100	0.265	0.244	0.232	0.215
150	0.285	0.262	0.251	0.225
200	0.425	0.423	0.410	0.380
250	0.485	0.461	0.450	0.425
300	0.335	0.316	0.304	0.295
350	0.385	0.363	0.351	0.331

The probe rate of filtering of our scheme and comparison made with two other existing schemes namely, Satisfy User Profiles (SUP) [1], Database-driven Reverse Dictionary (DRD) [2] and Automatic Discovery of Personal Name using Lexical Pattern-based (ADPN-LP) [3] is listed in table 3.

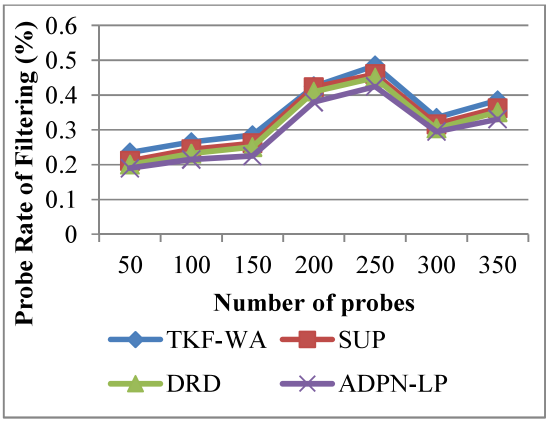

Figure 6 Measure Of Probe Rate Of Filtering

Figure 6 describes the probe rate of filtering on the probe count in the range of 50 to 350 probes taken for experimental purpose and implemented in JAVA. The probe count is obtained from the web server that offers the web services. The probe rate of filtering is improved using Temporal based Kalman filtering that handles both the regular and irregular user query and obtain the precise answer improving the probe rate of filtering by 5 – 10 % compared to SUP [1]. Also, the Temporal based Kalman filtering operates recursively using the metadata repository information to arrive at the optimality solution improving the probe rate of filtering by 3 – 14 %

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and 11 – 21 % compared to DRD [2] and ADPN-LP [2] respectively.

Table 4 Tabulation For Average Response Time

Number of probes	Average Response Time (ms)			
	TKF-WA	SUP	DRD	ADPN-LP
50	0.315	0.311	0.323	0.301
100	0.365	0.361	0.353	0.341
150	0.385	0.381	0.373	0.361
200	0.290	0.286	0.278	0.266
250	0.512	0.508	0.500	0.500
300	0.533	0.529	0.521	0.510
350	0.585	0.581	0.573	0.562

More accurately the influence of average response time based on web architecture with respect to the probe is listed in table 4 and comparison is made with three other existing schemes. It can also be seen that the average response time increases with the increase in the probe size.

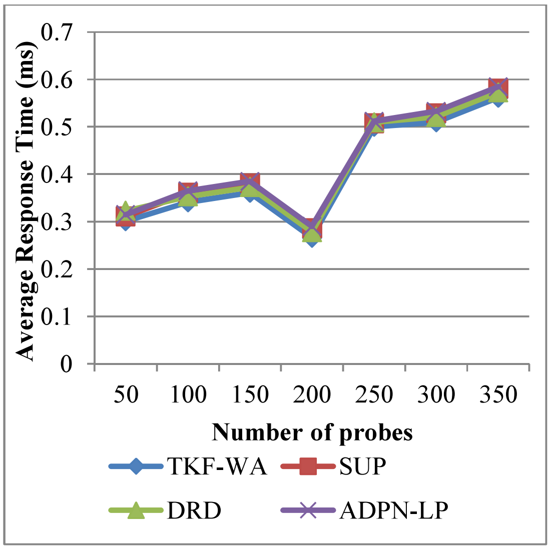

Figure 7 Measure Of Average Response Time

Fig 7 describes the average response time based on the probe size measured in terms of milliseconds (ms). From the figure it is illustrative that the average response time is not linear because of the nature of query which includes both regular and irregular types. This is turn make the response time to be not consistent with number of probes. With the application of frequency quantitative analysis measure, the average response time is improved by 1 to 7 % compared to SUP [1]. This is because in TKF-WA, current and previous keywords are used as the basis to measure frequency with respect to time which is then matched with the metadata repositories. Instead of performing the matching operations prior to identifying the frequency measure, the matching operations are performed after evaluation of frequency measure that results in the improvement of average response time by 1 – 7 % and 2 – 9 % compared to DRD [2] and ADPN-LP [2] respectively.

6. CONCLUSION

The inconsistence user query results on web architecture and the presence of regular and irregular query set have motivated us to propose an integrated solution for increasing the system utility rate on filtering and improve the search retrieval efficiency. To achieve this goal, Temporal-based Kalman Filter on Web Architecture (TKF-WA) is designed for metadata based semantic ontology search on web domain. Classification of regular and irregular query set is computed using the well-known semantic learning object link class. We also analyzed the performance of managing the maximal metadata repositories that removes the irrelevant pages which are not of higher significant to the user query. To implement our mechanism, we devised a filtering algorithm called Temporal based Kalman filtering to handle both the regular and irregular user query set. Experiments conducted on IRIS2 dataset proves the efficiency of our method that increases the system utility rate based on precision and search retrieval value based on recall among different probes made by the user through web server. Moreover, the probe rate of filtering was increased with relatively lesser average response time. The TKF-WA attains the performance improvements over other web search engine model with the experiments conducted with various conditions using JAVA platform results in average response time, maximal probe rate filtering and search retrieval value.

REFRENCES:

[1] Haggai Roitman., Avigdor Gal., and Louiqa Rasc., "A Dual Framework and Algorithms for Targeted Online Data Delivery," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 1, JANUARY 2011

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[2] Ryan Shaw., Anindya Datta., Debra VanderMeer., and Kaushik Dutta., "Building a Scalable Database-Driven Reverse Dictionary," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 3, MARCH 2011

[3] Danushka Bollegala., Yutaka Matsuo., and Mitsuru Ishizuka., "Automatic Discovery of Personal Name Aliases from the Web.," IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 6, JUNE 2011

[4] M. Rajeswari, G. Sambasivam, N. Balaji, M.S. Saleem Basha, T. Vengattaraman, P. Dhavachelvan," Appraisal and analysis on various web service composition approaches based on QoS factors", Journal of King Saud University – Computer and Information Sciences, Elsevier, Sep 2013

[5] Tak-Lam Wong and Wai Lam," Learning to Adapt Web Information Extraction Knowledge and Discovering New Attributes via a Bayesian Approach", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 22, NO. 4, APRIL 2010

[6] Weifeng Su, Jiying Wang, and Frederick H. Lochovsky," Record Matching over Query Results from Multiple Web Databases", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 22, NO. 4, APRIL 2010

[7] Shady Shehata, Fakhri Karray, and Mohamed S. Kamel," An Efficient Concept-Based Mining Model for Enhancing Text Clustering", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 22, NO. 10, OCTOBER 2010

[8] Xiaohui Tao, Yuefeng Li, and Ning Zhong," A Personalized Ontology Model for Web Information Gathering", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 4, APRIL 2011

[9] Chulyun Kim and Kyuseok Shim," TEXT: Automatic Template Extraction from Heterogeneous Web Pages", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 4, APRIL 2011

[10] Ningduo Peng, Guangchun Luo, Ke Qin, and Aiguo Chen," Query-Biased Preview over Outsourced and Encrypted Data", Hindawi Publishing Corporation The ScientificWorld Journal Volume 2013

[11] Kenneth Wai-Ting Leung and Dik Lun Lee," Deriving Concept-Based User Profiles from Search Engine Logs", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 22, NO. 7, JULY 2010

[12] Sohail Sarwara, Zia Ul Qayyuma, Saqib Majeedb," Ontology Based Image Retrieval Framework using Qualitative SemanticImage Descriptions", Elsevier, 17thInternational Conference in Knowledge Based and Intelligent Information and Engineering Systems - KES2013

[13] Yue Xie and Ting-Zhu Huang," A Model Based on Cocitation for Web Information Retrieval", Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2014

[14] Myunggwon Hwang, Student Member, IEEE, Chang Choi, Student Member, IEEE, and Pankoo Kim," Automatic Enrichment of Semantic Relation Network and Its Application to Word Sense Disambiguation",IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 23, NO. 6, JUNE 2011

[15] Weizhe Zhang, Hui He, and Jianwei Ye," A Two-Level Cache for Distributed Information Retrieval in Search Engines", Hindawi Publishing Corporation The ScientificWorld Journal

Volume 2013

[16] Saed Alqaraleh and Omar Ramadan," Elimination of Repeated Occurrences in Multimedia Search Engines", IAJIT, Dec 2011

[17] Morteza Karimzadeh, Wenyi Huang, Siddhartha Banerjee, Jan Oliver Wallgrün, Frank Hardisty, Scott Pezanowski, Prasenjit Mitra and Alan M. MacEachren," GeoTxt: A Web API to Leverage Place References in Text", ACM, Nov 2013

[18] Sanam Noor1 and Shariq Bashir," Evaluating Bias in Retrieval Systems for Recall Oriented Documents Retrieval", IAJIT, Sep 2013