Constraint - Aware Approach to Web Service Composition

C. B. Sai Krishnan[1], S. Brintha Rajakumari[2], M. Beema Mehraj[3.]

[1]UG Student, sairam.2610@yahoo.com, Department of CSE, Bharath University, Chennai.
[2]Assistant Professor, brintha.ramesh@gmail.com, Department of CSE, Bharath University, Chennai
[3]Assistant Professor, beemamehraj@gmail.com, Department of CSE, Bharath University, Chennai

Abstract **-**To improve the quality and composite the services with specific constraints using web services, Constraint aware approach based web service composition machines can automatically select, integrate and invoke different web services in-order to achieve the user specified task according to the user constraints. To implement this, first user has to register and get a unique username and password. Before that we have to compose all different web services together into one using web service composition method. The composition system has two kinds of participants, service provider and service requestor. The service providers propose web services for use. The service requestors consume information or services offered by service providers. If the service requestor requests for a service, then the provider search for a service that composes the available services in the service repository. Requestor can get the correct result as output if the input satisfies the constraints of the service.

Index Terms **-** *Clustering, Search Engine, Web Service.*

*****

I. INTRODUCTION

Accurately measuring the semantic similarity between words is an important problem in web mining, information retrieval, and natural language processing. Web mining applications such as, community extraction, relation detection, and entity disambiguation; require the ability to accurately measure the semantic similarity between concepts or entities. In information retrieval, one of the main problems is to retrieve a set of documents that is semantically related to a given user query. Efficient estimation of semantic similarity between words is critical for various natural language processing tasks such as word sense disambiguation (WSD), txtual entailment, and automatic text summarization.

Semantically related words of a particular word are listed in manually created general-purpose lexical ontology's such as WordNet. In WordNet, a synset contains a set of synonymous words for a particular sense of a word. However, semantic similarity between entities changes over time and across domains. For example, apple is frequently associated with computers on the web. However, this sense of apple is not listed in most general-purpose thesauri or dictionaries. A user, who searches for apple on the web, might be interested in this sense of apple and not apple as a fruit. New words are constantly being created as well as new senses are assigned to existing words. Manually maintaining ontology's to capture these new words and senses is costly if not impossible.

Here it proposes an automatic method to estimate the semantic similarity between words or entities using web search engines. Because of the vastly numerous documents and the high growth rate of the web, it is time consuming to analyze each document separately. Web search engines provide an efficient interface to this vast information. Page counts and snippets are two useful information sources provided by most web search engines. Page count of a query is an estimate of the number of pages that contain the query words. In general, page count may not necessarily be equal to the word frequency because the queried word might appear many times on one page.

II. RELATED WORKS

The paper [1] is measuring the semantic similarity between words is an important component in various tasks on the web such as relation extraction, community mining, document clustering, and automatic metadata extraction. Despite the usefulness of semantic similarity measures in these applications, accurately measuring semantic similarity between two words (or entities) remains a challenging task. They proposed an empirical method to estimate semantic similarity using page counts and text snippets retrieved from a web search engine for two words. Specifically, we define various word co-occurrence measures using page counts and integrate those with lexical patterns extracted from text snippets. To identify the numerous semantic relations that exist between two given words, we propose a novel pattern extraction algorithm and a pattern clustering algorithm. The optimal combination of page counts-based co-occurrence measures and lexical pattern clusters is learned using support vector machines. The proposed method outperforms various baselines and previously proposed web-based semantic similarity measures on three benchmark data sets showing a high correlation with human ratings. Moreover, the proposed method significantly improves the accuracy in a community mining task.

Here the paper [2] is about web ontology's provide shared concepts for describing domain entities and thus enable semantic interoperability between applications. To facilitate concept sharing and ontology reusing, we developed Falcons

Concept Search, a novel keyword-based ontology search engine. In this paper, They illustrate how the proposed mode of interaction helps users quickly find ontology's that satisfy their needs and present several supportive techniques including a new method of constructing virtual documents of concepts for keyword search, a popularity-based scheme to rank concepts and ontology's, and a way to generate query-relevant structured snippets. They also report the results of a usability evaluation as well as user feedback.

This paper [3] aims to answer two general questions: What contributes to search engine rankings? And what can web content creators and webmasters do to make their content and sites easier to find by audiences using search engines? Key concepts: Search engines' rankings are shaped by three classes of participants: search engine companies and programmers, search engine optimization practitioners, and search engine users. Key lessons: By applying three key lessons, professional communicators can make it easier for audiences to find their web content through search engines: consider their web content's audiences and website's competitors when analysing keywords; insert keywords into web text that will appear on search engine results pages, and involve their web content and websites with other web content creators. Implications: Because successful search engine optimization requires considerable time, professional communicators should progressively apply these lessons in the sequence presented in this tutorial and should keep up to date with frequently changing ranking algorithms and with the associated changing practices of search optimization professionals.

The paper[4] is mostly explaining the user behaviour while interacting with the web are based on a top-down approach, where the entire Web, viewed as a vast collection of pages and interconnection links, is used to predict how the users interact with it. A prominent example of this approach is the random-surfer model, the core ingredient behind Google's Page Rank. This model exploits the linking structure of the Web to estimate the percentage of web surfers viewing any given page. Contrary to the top-down approach, a bottom-up approach starts from the user and incrementally builds the dynamics of the web as the result of the users' interaction with it. The second approach has not being widely investigated, although there are numerous advantages over the top-down approach regarding personalization and decentralization of the required infrastructure for web tools. In this paper, they propose a bottom-up approach to study the web dynamics based on web-related data browsed, collected, tagged, and semi-organized by end users. The approach has been materialized into a hybrid bottom-up search engine that produces search results based solely on user provided web-related data and their sharing among users. They conduct an extensive experimental study to demonstrate the qualitative and quantitative characteristics of user generated web-related data, their strength, and weaknesses as well as to compare the search results of our bottom-up search engine with those of a traditional one. Our study shows that a bottom-up search engine starts from a core consisting of the most interesting part of the Web (according to user opinions) and incrementally(and measurably) improves its ranking, coverage, and accuracy. Finally, we discuss how our approach can be integrated with Page Rank, resulting in a new page ranking algorithm that can uniquely combine link analysis with users' preferences.

Here paper [5] is a relevance feedback (RF) is an iterative process, which refines the retrievals by utilizing the user's feedback on previously retrieved results. Traditional RF techniques solely use the short-term learning experience and do not exploit the knowledge created during cross sessions with multiple users. In this paper, we propose a novel RF framework, which facilitates the combination of short-term and long-term learning processes by integrating the traditional methods with a new technique called the virtual feature. The feedback history with all the users is digested by the system and is represented in a very efficient form as a virtual feature of the images. As such, the dissimilarity measure can dynamically be adapted, depending on the estimate of the semantic relevance derived from the virtual features. In addition, with a dynamic database, the user's subject concepts may transit from one to another. By monitoring the changes in retrieval performance, the proposed system can automatically adapt the concepts according to the new subject concepts. The experiments are conducted on a real image database. The results manifest that the proposed framework outperforms the traditional within-session and log-based long-term RF techniques.

III. CONSTRAINT AWARE APPROACH IN WEB

Here the existing system defines user search goals as the information on different aspects of a query that user groups want to obtain. Information need is a user's particular desire to obtain information to satisfy his/her need. User search goals can be considered as the clusters of information needs for a query. The inference and analysis of user search goals can have a lot of advantages in improving search engine relevance and user experience.

There are many issues in the systems. They are, what users care about varies a lot for different queries, finding suitable predefined search goal classes is very difficult and impractical. Analysing the clicked URLs directly from user click-through logs to organize search results. However, this method has limitations since the number of different clicked URLs of a query may be small. Since user feedback is not considered, many noisy search results that are not clicked by any users may be analysed as well. Therefore, this kind of methods cannot infer user search goals precisely. Only identifies whether a pair

of queries belongs to the same goal or mission and does not care what the goal is in detail.

Here, it proposes a tendency to aim at discovering the quantity of various user search goals for a question and portraying every goal with some keywords mechanically. We have a tendency to initial propose a completely unique approach to infer user search goals for a question by agglomeration our projected feedback sessions. Then, it has a tendency to propose a completely unique optimisation technique to map feedback sessions to pseudo-documents which may with efficiency mirror user info wants. At last, it has a tendency to cluster these pseudo documents to infer user search goals and depict them with some keywords. The projected feedback session consists of each clicked and un-clicked address's and ends with the last URL that was clicked in an exceedingly single session we have a tendency to propose this novel criterion "Classified Average Precision" to judge the reconstitute results. Supported the projected criterion, we have a tendency to conjointly describe the strategy to pick the simplest cluster variety.

It proposes a framework to infer different user search goals for a query by clustering feedback sessions. We demonstrate that clustering feedback sessions is more efficient than clustering search results or clicked URLs directly. Moreover, the distributions of different user search goals can be obtained conveniently after feedback sessions are clustered. It proposes a novel optimization method to combine the enriched URLs in a feedback session to form a pseudo-document, which can effectively reflect the information need of a user. Thus, we can tell what the user search goals are in detail. It proposes a new criterion CAP to evaluate the performance of user search goal inference based on restructuring web search results. Thus, we can determine the number of user search goals for a query.

IV. IMPLEMENTATION OF CONSTRAINT AWARE APPROACH

The framework of our approach consists of two parts divided by the dashed line. In the upper part, all the feedback sessions of a query are first extracted from user click-through logs and mapped to pseudo-documents and depicted with some keywords. We can see that in the figure1. Since we do not know the exact number of user search goals in advance, several different values are tried and the optimal value will be determined by the feedback from the bottom part.

In the bottom part, the original search results are restructured based on the user search goals inferred from the upper part. Then, we evaluate the performance of restructuring search results by our proposed evaluation criterion CAP. And the evaluation result will be used as the feedback to select the optimal number of user search goals in the upper part.

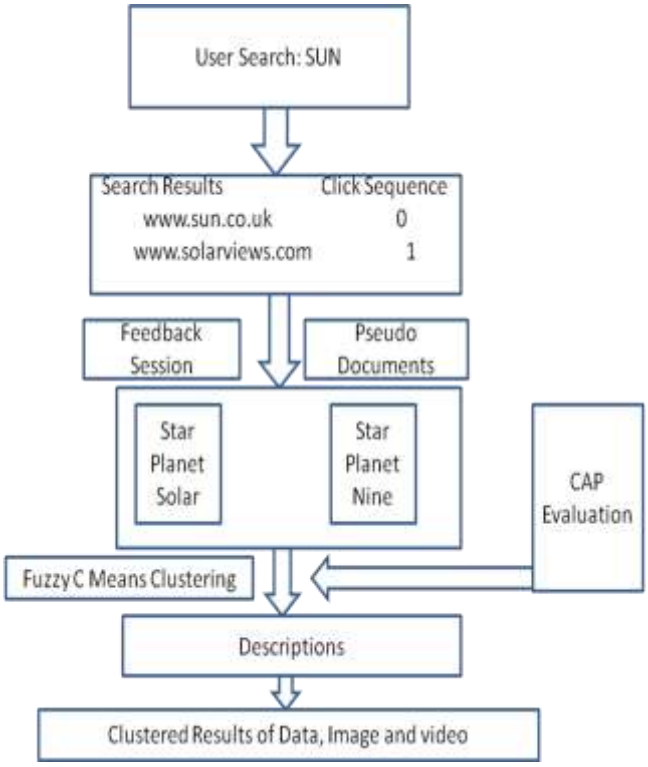

Figure 1: Architecture Diagram

V. RESULT AND ANALYSIS

The running time is usually short. In reality, our approach can discover user search goals for some popular queries offline at first. Then, when users submit one of the queries, the search engine can return the results that are categorized into different groups according to user search goals online. Thus, users can find what they want conveniently is shown in the figure 2.

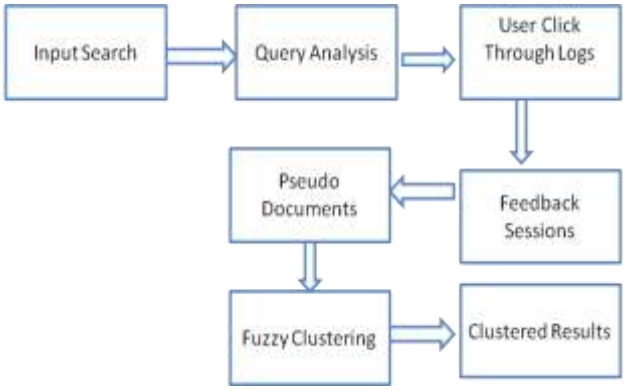

Figure 2: Web model diagram

Our approach can be enhancing by applying both offline and online process in this concept. Where offline approach can be reached by our default searched user logs along with the feedback sessions. Which will proves lots of efficiency in our approach. Privacy preserved personalized search also can be enhanced for secure search. Ontology will be a best technique for future work of this implementation

The graph search-based algorithm can be applied to the general WSC problem, and generates all the possible solutions.

For the two pre-processing methods, despite their differences, they both introduce conditional branch structures (explicitly or implicitly) into the process model of a solution in order to solve the problems brought by service constraints, and ensure the correct execution of the resulting composite service. Extensive experiments are conducted via a publicly available test set, and experimental results show the effectiveness and admissibility of our approach.

The main homepage of the search engine is in figure 3 we have few menus such as home, registration, and login and about us. A new user who is going to search needs to register and that will be stored in the database. In this page the user will register for using the login facility. It collects some details of the user. Once the user is registers successfully, it shows a pop alert registered successfully. In the login page now the user can login with username and password.

Once the user login it asks for option such as (a) google search : This is a normal search which redirects to google website (b) customize search: In this search we can have a specfic search and get what exaclty the user needs. Once the user selects the google search, user needs to specify the URL. According to the URL requested the output is given. In the customised search the user can give a query and search and get relevant links in a clustered format is in the figure 4.

According to the user need the search is made and related websites are displayed, so that the user can avoid the unwanted URLS which or not relevant. This method of customised search saves time for the user. The user has a separate login in home page as user login after searching. The user login contains pseudo documents and feedback sessions.

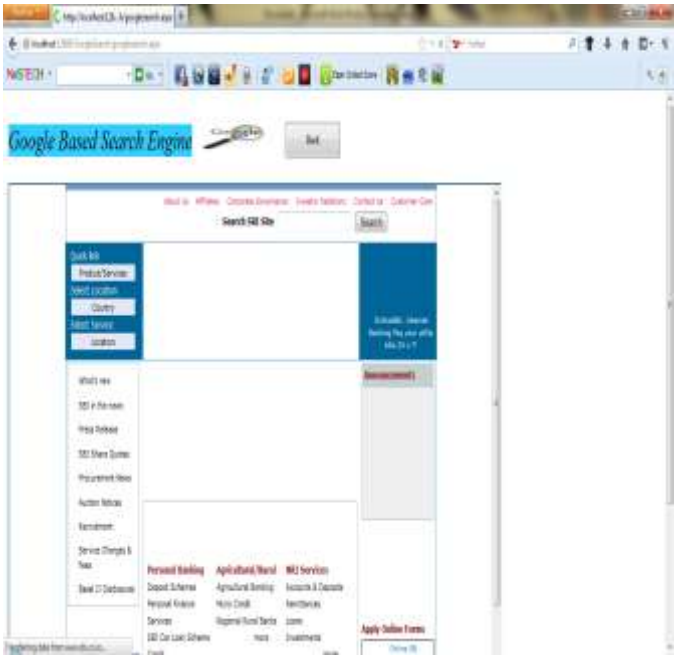

Figure3: Google Search Option

Figure 4: Customised Search Engine Results

VI. CONCLUSION

Service constraints are used to ensure the correct execution of the service or proper interaction with other services, thus having a significant impact on service composition. This work deals with automated service composition while considering them. Through some real-world business scenarios, we show that some available services in a specific business scenario have the same inputs and outputs, and can perform the same task, but have different constraints. This is common, but has been largely ignored by previous efforts for service composition. A novel solution to tackle this problem is proposed in this paper, which includes a graph search-based algorithm and two different pre-processing methods.

REFERENCES

[1] DanushkaBollegala, Yutaka Matsuo, and Mitsuru Ishizuka *IEEE Transactions on Knowledge and Data Engineering,* 2010.

[2] YuzhongQu and Gong Cheng Systems, Man and informatics, half A: Systems and Humans, *IEEE Transactions on Knowledge and Data Engineering*, 2011.

[3] JOHN B. KILLORAN skilled Communication, *IEEE Transactions on Knowledge and Data Engineering*, 2013.

[4] Athanasios Papagelis and Christos Zaroliagis Systems, Man and informatics, half A: Systems and Humans, *IEEE Transactions on Knowledge and Data Engineering* ,2012

[5] Peng-Yeng rule, BirBhanu, Fellow, Kuang-Cheng Yangtze Kiang, and Anlei Dong., *IEEE Transactions on Knowledge and Data Engineering*, 2008.

[6] D. Ardagna and B. Pernici, "Adaptive service composition in flexible processes," *IEEE Trans. Softw. Eng.*, vol. 33, no. 6, pp. 369–384, Jun. 2007.

[7] M. Beek, A. Bucchiarone, and S. Gnesi, "A survey on service composition approaches: From industrial standards to formal methods," IEEE Comput. Soc. Press, New York, NY, USA, Tech. Rep. 2006TR-15, Istituto, 2006, pp. 15–20.

[8] D. Benslimane, Z. Mavamar, Y. Taher, M. Lahkim, M. C. Fauvet, and M. Mrissa, "A multilayer and multiperspective approach to compose web services," in *Proc. AINA*, May 2007, pp. 31–37.

[9] D. Berardi, D. Calvanese, G. De Giacomo, M. Lenzerini, and M. Mecella, "Automatic composition of e-services that export their behaviour," in *Proc. ICSOC*, Dec. 2003, pp. 43–58.

[10] D. Berardi, D. Calvanese, G. De Giacomo, M. Lenzerini, and M. Mecella, "Synthesis of underspecified composite e-services based on automated reasoning," in *Proc. ICSOC*, Nov. 2004, pp. 105–114.

[11] D. Berardi, D. Calvanese, G. De Giacomo, M. Lenzerini, and M. Mecella, "Automatic service composition based on behavioural descriptions," *Int. J. Coop. Inf. Syst.*, vol. 14, no. 4, pp. 333–376, Dec. 2005.

[12] A. Brogi, S. Corfini, and R. Popescu, "Composition-oriented service discovery," in *Proc. SC*, Apr. 2005, pp. 15–30.

[13] A. Brogi and S. Corfini, "Behaviour-aware discovery of web service compositions," *Int. J. Web Serv. Res.*, vol. 4, no. 3, pp. 1–25, Jul.–Sep. 2007.

[14] A. Brogi, S. Corfini, and R. Popescu, "Semantics-based composition oriented discovery of web services," *ACM Trans. Internet Technol.*, vol. 8, no. 4, article 19, Sep. 2008.

[15] Z. J. Ding, J. L. Wang, and C. J. Jiang, "An approach for synthesis Petri nets for modelling and verifying composite web services," *J. Inf. Sci. Eng.*, vol. 24, no. 5, pp.