International Journal of Scientific & Engineering Research, Volume 5, Issue 5, May-2014
ISSN 2229-5518

Improvised Web Search through Web Services on Ontology Based Semantics

P. Nandhakumar[1] , Research Scholar, Department of Computer Science, Karpagam University, Coimbatore, Tamil Nadu, India. Email: nandhap@gmail.com[1]

M. Hemalatha[2] , Department of Computer Science, Karpagam University, Coimbatore, Tamil Nadu, India, Email: csresearchhema@gmail.com[2]

Abstract— The internet plays a major role in transforming huge amount of information day by day in World Wide Web. Search engine is the important tool source to interact user for information retrieval. The traditional way of finding answers to the search queries provide irrelevant result for the searched quires. Web 3.0 collaborates with semantic web gives better result for searched queries. The main focus of this paper is to give a framework using semantic web services constructed with service oriented architecture (SOA) for improving searching capacity. The proposed methodology uses depth first search algorithm for semantic search services. The semantic search uses RDF data set and work with responses constructed with Ontologies. The proposed study not only improves response time, it also provides ranking schemes based ontology relevant score of the responses observed.

Index Terms— Semantic Web, Semantic Web Search, SOA, Ontologies, Semantic Web Services, Semantic Web Service Composition.

1. INTRODUCTION

Large amount of data available are provided by World Wide Web, which made an unpredictable growth [1]. The information gathered by the means of search engines is mostly used. The search results provided by the search engines are most likely to be irrelevant from its search queries, because of abandoned number of information available in internet. The current human understandable data was extended by the foremost initiative of the semantic web [6], because it encodes some of the semantics of resources in a machine understandable form. Newer avenues of more additional functionalities, advanced application on the web are provided by the www data, which is in the machine understandable form.

Semantic web applications provide a meaningful search result for a searched query by integrating processed information and information searched.W3C Consortium's semantic web provides a database for Resource Description Framework (RDF) [2]. Best machine understandable format can be formulated in RDF database by introducing Web Ontology Language (OWL) [3], [4], [9]. The semantic web is imagined to encourage coordination of information accessible crosswise over different web requisitions, web servers through semantic web services [6] [7]. The semantic data accessible with the service providers is in the RDF form. Also the ontologies based dependent upon the RDF information give imparted ideas and relations utilizing either OWL1 or OWL2. We characterize such benefit service as semantic providers. To encourage extra enlightening rationales around the elements the schema depicted in this paper recognizes Owl2 when contrasted with its antecedent Owl1 which gives restricted backing. The interoperability of the semantic data accessible with the web requisitions is a major prerequisite which prompts giving semantic hunt as an administration separated from extra semantic web administrations for provisions. Semantic Web Services are like traditional web services which give machine to machine collaboration over the www between a web server and a web client. Semantic web services use markup dialects to give information in a machine meaningful structure with all the social portions (ontology details). The SSFSWS presented in this paper gives a semantic inquiry environment based on services oriented architecture (SOA). The structural planning of the SSFSWS is as indicated in Figure 1. The remaining original copy is composed as takes follows section 2 of the paper talks about the writing survey embraced and the work of individual analysts. The following area presents the SSFSWS in which the semantic search providers, the depth first search algorithm for semantic web service composition. Section 4 this paper talks about the test study attempted and the conclusion and future work is examined in the last area of the paper.

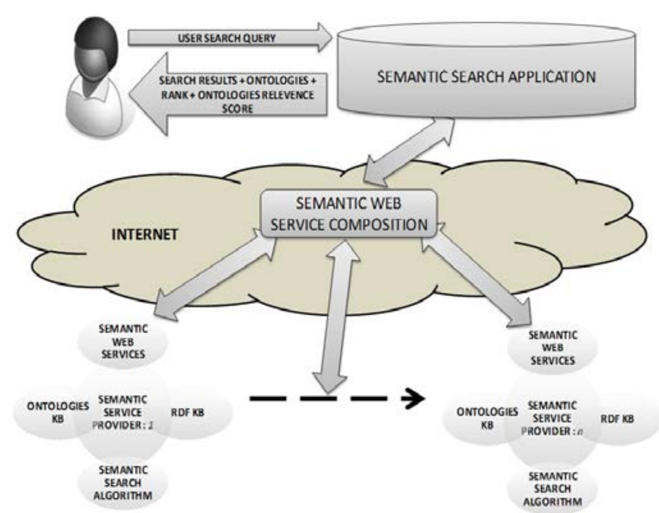

Figure 1: SSFSWS System Architecture

ISSN 2229-5518

2. LITERATURE REVIEW

The semantic web envisions the web content to be in a machine understandable form. Representation the web content in the RDF form has been standardized by W3C [2]. Searching the RDF web content put forth the existence of semantic search engines [21]. Early attempts by researchers [10] to semantically alter the search queries to provide better search responses were introduced. Researchers have proposed search engines based on the RDF knowledge bases [11][12][13]. Ontologies based semantic search engines proved to provide higher accuracy [14] [15] [16]. Hybrid techniques for providing better search results have also been studied during the course of the research work presented here. In hybrid techniques in addition to RDF data Ontologies extracted are also used to provide for better search environments. [17][18][19]. It could be concluded that hybrid techniques perform better than the other classes of semantic search engines hence the SSFSWS utilizes a hybrid technique of providing search responses based on the RDF data and the Ontologies extracted [20]. The framework discussed in this paper considers the semantic search provided through web services [22]. Ranking of the semantic search results have been achieved using various techniques like extended information retrieval techniques [23], interpretations [24], file rank matrices [25] and concept based ranking [26]. Form the study it can be concluded that ontology based ranking mechanism would be idyllic. The framework proposed in this paper considers an Ontology Relevance Score based ranking system. Image Search based on semantics is considered as a hot topic of research currently [27][28]. The future work of the SSFSWS proposed in this paper could be considered to support image based search.

3. SEMANTIC SEARCH FRAMEWORK USING SEMANTIC WEB SERVICES

SSFSWS framework discuss proposed in this section of the paper. The SOA chosen enable us to realize the framework using a modularized approach. The SOA shown in Fig. 1 could be considered as a complex system of $p : q$ dependencies. Where p represents the services offered and q represents the applications offered by the SOA system. In SOA an application may need multiple service offerings or varied applications need similar services or similar applications may be provided by varied services. The SSFSWS utilizes a similar application of semantic search provided by the n semantic search providers hence it could be said that the SSFSWS depends on the availability of the semantic search application offered by the n semantic search providers. Semantic web service management tends to be cumbersome if it is done manually. In order to automate the semantic web service management we need a common syntax and a common semantic service description to interoperate. The W3C have standardized the syntax definition through the Web Service Description Language (WSDL) [29]. Semantic Interface Description language adopted by the SSFSWS is OWL-S [30]. Let us consider a set of all semantic concepts S_C defined as

$$S_C = \{s_{c_1}, s_{c_2}, s_{c_3}, \dots\dots\dots s_{c_a}\}$$

Where s_{c_a} represents the a^{th} semantic concept. The s_{c_a} is a concept derived from the RDF $r_{c_a} \in R_C$ knowledge base (KB) and the ontologies $o_{c_a} \in O_C$ derived from the RDF knowledge base R_C. In other words

$$s_{c_a} = f_c(r_{c_a}, o_{c_a})$$

Where $o_{c_a} = f_O(R_C, r_{c_a})$ f_O Represents the ontology building function using $SROIQ\ (DL)$. The ontology building function extracts all the related concepts and axioms of r_{c_a} present in the RDF KB R_C .The SSFSWS represents a complex SOA hence the RDF R_C data set is available with n^{th} semantic service providers. The RDF data can be defined as

$$R_C = R_{C1} \cup R_{C2} \cup R_{C3} \cup \dots\dots R_{Cn}$$

Where $R_{C1} \neq R_{C2} \neq R_{Cn}$

$$R_C = \{r_{c_{11}}, r_{c_{12}}, \dots\dots r_{c_{1a}}\} \cup \{r_{c_{21}}, r_{c_{22}}, \dots\dots r_{c_{2a}}\} \cup \dots\dots \{r_{c_{n1}}, r_{c_{n2}}, \dots\dots r_{c_{na}}\}$$

The ontologies extracted or the ontology knowledge base could be defined as $O_C = O_{C1} \cup O_{C2} \cup \dots \cup \dots\dots O_{Cn}$ where O_{Cn} the ontology set available with n^{th} semantic service provider. The locally available ontologies could be defined as $O_{Cn} \propto R_C$ from the above definition it is clear the ontologies available with semantic service provider n may not contain all the possible concepts, relations and axioms as the complete RDF set R_C is unavailable with the n^{th} semantic service providers. This is the problem that exists in the current semantic search deployments available. The purpose of the SSFSWS is to overcome the short comings by using efficient searching algorithms and semantic web service compositions.

3.1 Semantic Search Providers

The semantic search providers in the SSFSWS provide semantic search web services which support the semantic search application. The semantic search providers are defined as

$$PR = \{Pr_1, Pr_2, Pr_3, \dots\dots\dots Pr_n\}$$

where Pr_n is the n^{th} semantic search providers. The system architecture of the semantic search providers is shown in Figure 2.

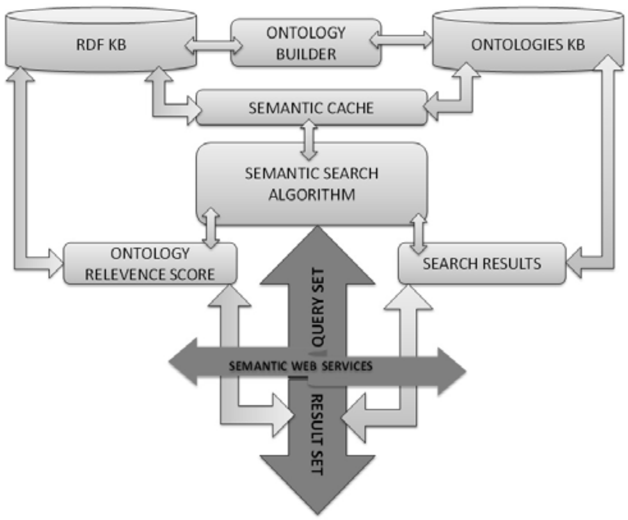

Figure 2: System Architecture of Semantic Search Providers

Each semantic search providers possess the RDF KB. The RDF KB could be represented as $R_{KB} = \{ rkb_1, rkb_2, rkb_3, rkb_4, \dots \dots \dots . rkb_r \}$

Where rkb_r is the r^{th} RDF data record available with the semantic search provider $Pr_n \in PR$.

The RDF data is usually embedded in XML documents as they support segment declaration using tags. The RDF data usually is enclosed within tags represented as $< rdf{:}RDF >$ and $</rdf{:}RDF >$ [31]. The RDF records are said to consist consists of triplets [32]. Based on the concept of triplets the record rkb_r could be represented as $rkb_r = < trkb_{r_{sub}}, trkb_{r_{prd}}, trkb_{r_{obj}} >$

Where $trkb_{r_{sub}}$ is the subject triplet, $trkb_{r_{prd}}$ is the predicate triple and $trkb_{r_{obj}}$ represents the object triplet.

The concepts extracted from the RDF KB include some complex relations that cannot be represented in RDF alone hence the SSFSWS presented here adopts representation of the ontologies through $SROIQ$ Axioms due to its benefits [9] as compared to $SHOIN$. A brief summary of the syntax and semantics of the $SROIQ$ is as shown in Table 1.

The KB of the semantic search providers which constitute of both the RDF KB and Ontology KB are humongous in nature and size. A search executed on huge databases would affect the response times due to numerous disk read and disk write operations involved in the semantic search. To compress the knowledge base and create cache the SSFSWS utilizes a hierarchical data ordering algorithm.

The ontology KB of r RDF data records is defined as $O_{KB} = \{ okb_1, okb_2, okb_3, okb_4, \dots \dots \dots . okb_r \}$

Name	Syntax	Semantics
Concepts		
Atomic Concept	A	A^I (given)
Nominal	$\{a\}$	$\{a^I\}$
Top Concept	T	Δ^I
Negation	$\neg C$	Δ^I / C^I
Conjunction	$C \sqcap D$	$C^I \cap D^I$
Existential Restriction	$\exists R.C$	$\{x \mid R^I(x, C^I) \neq \emptyset\}$
Min Cardinality	$\geq nS.C$	$\{x \mid \parallel S^I(x, C^I) \parallel \geq n\}$
Exists Self	$\exists S.Self$	$\{x \mid < x, x > \in S^I\}$
Axioms		
Complex Role Inclusion	$\rho \sqsubseteq R$	$\rho^I \subseteq R^I$
Disjoint Roles	$Disj(S_1, S_2$	$S_1^I \cap S_2^I = \emptyset$
Concept Inclusion	$C \sqsubseteq D$	$C^I \subseteq D^I$
Concept Assertion	$C(a)$	$a^I \in C^I$
Role Assertion	$R(a,b)$	$\langle a, b \rangle \in R^I$

Table1: SROIQ Rules Semantic Equivalent

Let the cache of a concept s_{c_a} which represents the a^{th} semantic concept be represented as $Cache_{s_{c_a}} = < s_{c_a}, r_{c_a}, e_{c_a} >$ where r_{c_a} is the number of relations of the concept and e_{c_a} represents the number of edge concepts. It is evident that greater the number of concepts and greater the relations that exist larger is the KB size increasing the number of disk operation for a semantic search. The number of occurrences of a concept in an ontology is directly proportional or equivalent to the number of relations r_{c_a} of a concept. Also it can be stated that for a constant m is equivalent to a function of the number of relations (f_{num_r}) r_{c_a} of a concept s_{c_a} and a function of the edge depth (f_{edg_dpth}) of a concept s_{c_a}.

$$r_{c_a} \approx \sum_{x=1}^{x=m} m/x \approx m \int_{x=1}^{x=m} 1/x \, dx = m \ln m$$

$$f_{num\_r}(r_{c_a}) \times f_{edg\_dpth}(r_{c_a}) \approx m$$

Also

From the above equation it is clear that even if the number of relations r_{c_a} of a concept s_{c_a} increase the cache size does not increase by a great extend. Generally the concepts require $2S_{Unit}$ storage space per concept (s_{c_a}). The space utilized in storing the cache defined above is given by

$$\sum_{r_{c_a}} (2 + f_{num\_r}(r_{c_a})) \approx S_{Util}(2 + \ln S_{Util})$$

Where S_{Util} is the space required to store the same concept s_{c_a}. It is considered that only one entry of a s_{c_a} concept is allowed in the cache. In order to compare the normal caching strategy with the caching strategy used in SSFSWS the comparison ratio is defined as

ISSN 2229-5518

$$\frac{2S_{Util}\left(1+{}^{\ln S_{Util}}/_{2}\right)}{2S_{Util}\ln S_{Util}} = 1/\ln S_{Util} + 1/2$$

Hence the proposed caching strategy improves the storage space utilization by approximately 50% . The access cost for

$$ACost_{Cache} = \sum_{S_{Util}} f_{num_r}(r_{c_a}) \approx \int_{S_{Util}/t} \frac{S_{Util}\{r_{c_a}: f_{num_r}(r_{c_a}) \le t\}}{f_{edg\_dpth}} df_{edg\_dpth} = S_{Util} \ln t$$

the caching strategy is defined as

Where $f_{num\_r}(r_{c_a}) \le t \leftrightarrow f_{edg\_dpth}(r_{c_a}) \ge S_{Util}/t$

The probability of $AProb$ finding the concept s_{c_a} in the knowledge base is defines as $AProb_{Cache} = \ln t/\ln S_{Util}$

The access time of the cache to search for a concept s_{c_a} within the knowledge base with a probability $AProb_{Cache}$ is defined as $ATime_{Cache} = AProb_{Cache} \log_b ACost_{Cache}$

$$+ (1 - AProb_{Cache})(\log_b ACost_{Cache} + 1)$$
$$= \log_b ACost_{Cache} + (1 - AProb_{Cache})$$

Where b represents the branching factor of the ontology tree. The cache created based on the RDF KB and Ontology KB is encoded in a binary format for faster access.

The ontology relevance score is a ration between the query concept and the response concept based on the ontologies constructed. The ontology relevance score is used by the Semantic Search Application in ranking the search responses received by the n semantic service provider considered in the SSFSWS. The semantic query SS_Q could be defined as a set of concepts and relational operators. The semantic search web service offered supports queries containing Boolean operators like $AND, OR, NOT, +, -, ""$ commonly available with the major search providers. $ss_Q = < s_{C_{ss_Q}}, R_{c_{ss_Q}} >$

The semantic query ss_Q could be represented as a $p \times q$ matrix where p represents the number of concepts queried for and q represents the number of relations, logical operators and special characters defined for querying amongst the p concepts. The semantic response ss_R is a set of responses and the corresponding ontology relevance score defined as $ss_R = < s_{R_{ssR}}, ORS_{R_{ssR}} >$

The semantic response ss_R could also be represented as a $r \times r$ matrix where r is the number of responses obtained for the semantic search query ss_Q. The ontology relevance score is defined as

$$ors_{R_{ssR}}(ss_Q, ss_R) = \frac{\sum_r s_{r_{ssR}} \cdot ss_Q}{\| ss_Q \| \| s_r \|}$$

To represent the ontology relevance score to a scale of $0\ to\ 1$ .Normalization is considered in the SSFSWS hence the ontology relevance score could be defined as

$$ors_{R_{ssR}}(ss_Q, ss_R) = s_{r_{ssR}}{}', ss_Q$$

Where $s_{r_{ssR}}{}' = \frac{s_{r_{ssR}}}{s_r}$

The semantic search providers could be considered as the core of the SSFSWS architecture. The providers discussed in this section not only rely on the RDF KB to provide effective search queries but also rely heavily on the Ontologies KB to provide effective and accurate search responses. The semantic search provider's not only incorporate effective hierarchical caching strategies enhancing query response time but also provide relevant query responses. In addition to the query responses the search providers also provide ontology relevance scores associated with each query responses enabling effective ranking when multiple semantic search responses are composed.

3.2 Semantic Search Application

The semantic search application is a user interface which accepts user search queries represented by SS_Q. The SSFSWS accepts logical, conditional and simple term based search queries. The response of the search is represented as SS_R. The semantic search application provides the search responses SS_R by using semantic web service composition techniques. The depth first search based semantic web service composition algorithm is discussed in the next section of this paper. The semantic search response not only consists of search responses but additionally provides the ontology relevance score used in ranking the search responses i.e. higher the ontology relevance score greater is the rank of the search response. The semantic search application also provides the ontologies constructed after consuming the semantic services provided by the n semantic service providers. The provided ontologies are constructed by the possible concepts and axioms obtained post the semantic web service composition. This enables the SSFSWS to provide better semantic search results and overcome the drawback currently prevalent in the semantic web search sphere (discussed in the previous section of this paper).

Let us consider semantic concept set S_C and two concepts $s_{c_x} \in S_C$ and $s_{c_y} \in S_C$. There exists 4 possible relations amongst concepts s_{c_x} and s_{c_y}. The possible relations could be defined by using the subsume represented by Sb_{sum} and defined as $Sb_{sum} : (S_C \times S_C) \mapsto \{T, F\}$

Where T represents the conditional true relation and F represents a conditionally false relation. Using the above definition we could define the first possible relation between the concepts s_{c_x} and s_{c_y} as

$Sb_{sum}(s_{c_x}, s_{c_y}) = T$ Holds if and only if the semantic concept s_{c_x} is a generalization of the semantic concept s_{c_y}. Also then it could be stated that the semantic concept s_{c_y} is a specialization of the semantic concept s_{c_x}.

$Sb_{sum}(s_{c_y}, s_{c_x}) = T$ Holds if and only if the semantic concept s_{c_x} is a specialization of the semantic concept s_{c_y}. Also then it could be stated that the semantic concept s_{c_y} is a gener-

alization of the semantic concept s_{c_x} .
If the semantic concepts s_{c_x} and the semantic concept s_{c_y} are not related then $Sb_{sum}\left(s_{c_x}, s_{c_y}\right) = F$ and $Sb_{sum}\left(s_{c_y}, s_{c_x}\right) = F$.
If the semantic concepts s_{c_x} and the semantic concept s_{c_y} are equal then $Sb_{sum}\left(s_{c_x}, s_{c_y}\right) = T$ and $Sb_{sum}\left(s_{c_y}, s_{c_x}\right) = T$.
The generalization, specialization and the subsume Sb_{sum} relations are transitive. Let us consider a parameter p_x of the semantic service provider Pr_x and a parameter p_y of the semantic service provider Pr_y. If the parameters $p_x = p_y$ then the semantic web service could be called if only $Sb_{sum}(Pr_x, Pr_y) = T$. It could also be stated that the parameter p_x requires less or equal data than the parameter p_y.For the semantic web service composition required there is no requirement for a demarcation amongst the concepts and the semantic concepts. Let's define a set of semantic web services available with the semantic search application as follows
$Sws_{Pr} = \{ sws_{Pr_1}, sws_{Pr_2}, \dots\dots\dots, sws_{Pr_n} \}$

where sws_{Pr_n} represents the n^{th} semantic web service offered by semantic service provider Pr_n
Each semantic web service offered by semantic service provider Pr_n required a set of inputs denoted as $ss_{Q_{Pr_n}}$ and if the set of inputs is provided in an orderly fashion the semantic web service provides a set of output concepts denoted by $ss_{R_{Pr_n}}$ and $ss_{R_{Pr_n}} \in S_C$. The depth first search semantic web service composition algorithm discovers the semantic web services available Sws_{Pr} . On successful execution of the semantic web service execution algorithm the next semantic web service i.e. sws_{Pr_2} could be processed only if the execution of the previous sws_{Pr_1} (provided with the input parameters $ss_{Q_{Pr_1}}$ and the output concepts $ss_{R_{Pr_1}}$ are obtained in response) is processed successfully. Let the depth first search based semantic web service composition be represented as $Comp_{Sws}\,(Sws_{Pr})$ then the semantic web service composition is said to successfully process all the requests if

$$Comp_{Sws}\,(Sws_{Pr}) \leftrightarrow \forall\, \mathcal{X} \in ss_{Q_{Pr_1}} \exists\, \mathcal{Y} \in SS_{Q_{Pr}} : Sb_{sum}(\mathcal{X}, \mathcal{Y}) \wedge \forall\, \mathcal{X} \in ss_{Q_{Pr_z}}, z \in \{2, 3, \dots, n\}$$
$$\exists\, \mathcal{Y} \in SS_{Q_{Pr}} \cup ss_{R_{Pr_{z-1}}} \cup \dots\dots \cup ss_{R_{Pr_1}} : Sb_{sum}(\mathcal{X}, \mathcal{Y}) \wedge \forall\, \mathcal{X} \in ss_{R_{Pr_1}} \cup \dots\dots \cup ss_{R_{Pr_n}} \cup SS_{Q_{Pr}} : Sb_{sum}(\mathcal{X}, \mathcal{Y})$$

Let f_{SSWS} represent a semantic service provider search function based on a concept S_C which provides all the set of semantic web services available defined as
$f_{SPr}(S_C) = Sws_{Pr}$
Also it could be stated that
$\forall\, s_{c_a} \in f_{SPr}(S_C)\ \exists\, ss_{R_{Pr_a}} \in SS_{R_{Pr}} : Sb_{sum}\left(S_C, ss_{R_{Pr_a}} \right)$

The semantic search application is an interface which provides the search criteria to the composed services the results obtained are then there by provided to the user. On receiving the user's semantic search query SS_Q the application of the SSFSWS performs the semantic web services search function f_{SSWS}. The web service offerings amongst the varied semantic service providers are obtained by the advertisement process invoked by the f_{SSWS}. Based on the semantic web services offered and the user query appropriate web services are selected. The selected web service offerings Sws_{Pr} are composed using the semantic web service composition function $Comp_{Sws}\,(Sws_{Pr})$. On completing the composition the semantic web services are invoked by parsing the required user parameters SS_Q . The results obtained are aggregated and ranked based on the ontology relevance score. Higher is the ontology relevance score higher is the rank. The ranking could be easily achieved using any sorting algorithm.
Let the semantic web search response set be defined as
$SS_R = \{ ss_{R_{Pr_1}}, ss_{R_{Pr_2}}, ss_{R_{Pr_3}}, \dots, ss_{R_{Pr_n}} \}$

Where $ss_{R_{Pr_n}}$ represents the semantic search response received from the n^{th} semantic service provider for a given query set SS_Q . As stated earlier the semantic search algorithm available at the semantic service provider's end, provide the result page info, the ontologies behind the search, the ontology relevance score. Based on this argument $ss_{R_{Pr_n}}$ could be defined as
$ss_{R_{Pr_n}} = \{ r_1 ss_{R_{Pr_n}}, r_2 ss_{R_{Pr_n}}, r_3 ss_{R_{Pr_n}} \dots, r_m ss_{R_{Pr_n}} \}$

Where $r_m ss_{R_{Pr_n}}$ represents the m^{th} search result received from the n^{th} semantic service provider for a given query set SS_Q .
The semantic web service composition is an important entity of the semantic web search application. The next section of this paper discusses the depth first search algorithm utilized in composing the semantic web services Sws_{Pr} offered by the n semantic service providers.

3.2.1 Semantic Web Service Composition Using Depth First Search Algorithm

The semantic search framework SSFSWS introduced in this paper utilizes the depth first search algorithm for semantic web service composition. The dept first search algorithm is selected for the sole purpose of quicker responses it offers and it is computationally lighter when compared to other semantic web service composition algorithms. The web service composition function introduced in the earlier section of this paper $Comp_{Sws}\,(Sws_{Pr})$ receives the set of semantic web services Sws_{Pr} over which the composition has to be performed. The semantic web services composition is performed using the depth first search algorithm. Let us define a function f_{SWS-DS} which performs the depth first algorithm is defined as
$f_{SWS-DS}\left(ss_Q, ss_R, Sws_{tmp}, d_c \right) = sws$

ISSN 2229-5518

Where ss_Q represents the input query set, ss_R is the desired response, Sws_{tmp} represents the current temporary semantic web services identified, d_c represents the current depth and sws represents the resultant semantic web service identified.

The f_{SWS-DS} is solved by the following algorithm

Step 01: START
Step 02: For Each $Var_1 \in ss_R$
Step 03: For Each $sws_{Pr} \in f_{SPr}(S_C) = Sws_{Pr}$
Step 04: Initialization $ss_{R_{tmp}} = ss_R$
Step 05: **For Each** $Var_2 \in ss_{R_{tmp}}$
Step 06: IF $\exists\, Var_3 \in ss_{R_{Pr}} :\ Sb_{sum}(Var_2, Var_3)$
Step 07: $ss_{R_{tmp}} = ss_{R_{tmp}} / Var_2$
Step 08: **End IF**
Step 09: **End For Each**
Step 10: **For Each** $Var_4 \in ss_{Q_{tmp}}$
Step 11: IF $\exists\, Var_5 \in ss_{Q_{Pr}} :\ Sb_{sum}(Var_4, Var_5)$
Step 12: $ss_{R_{tmp}} = ss_{R_{tmp}} \cup Var_4$
Step 13: **End If**
Step 14: **End For Each**
Step 15: $sws_{tmp} = sws_{Pr} \oplus Sws_{tmp}$
Step 16: **IF** $ss_{R_{tmp}} = \{\}$
Step 17: **Return** sws_{tmp}
Step 18: **End IF**
Step 19: **ELSE**
Step 20: **IF** $d_c < d_{max}$
Step 21:
$$ss_{R_{tmp}} = f_{SWS-DS}\left(ss_Q, ss_{R_{tmp}}, sws_{tmp}, d_c + 1 \right)$$
Step 22: **End IF**
Step 23: **IF** $sws_{tmp} \neq \{\}$
Step 24: **Return** sws_{tmp}
Step 25: **End IF**
Step 26: **End ELSE**
Step 27: End For Each
Step 28: End For Each
Step 29: Return $\{\}$
Step 30: END

Where $Var_1, Var_2, Var_3, Var_4, Var_5$ represent temporary processing variables and d_{max} represents the maximum depth.

The semantic web service composition function denoted by $Comp_{Sws}(Sws_{Pr})$ is realized using the following algorithm

Step 01: START
Step 02: Initialization $d_{max} = 2$
Step 03: **DO**
Step 04: $Sws_{tmp} = f_{SWS-DS}\left(ss_{Q_{Pr_n}}, ss_{R_{Pr_n}}, \{\}, 1 \right)$
Step 05: $d_{max} = d_{max} + 1$
Step 06: While $Sws_{tmp} \neq \{\}$
Step 07: END

The $Comp_{Sws}$ composes the semantic web services from the back. In each iteration of the $DO-WHILE$ loop the algorithm checks if a service within the Sws_{Pr} could provide the service requested for. In the proposed framework SSFSWS we use the composition algorithm to identify the semantic service providers offering the semantic search web services to support the semantic search application.

The SOA architecture considered for the SSFSWS is described in this section of the paper. The SSFSWS is designed to provide appropriate search responses. The SSFSWS relies on the RDF KB and the Ontologies KB housed as the KB component of the semantic web service providers for provisioning of the search responses. The semantic search web services offered by the service providers are composed using the depth first search algorithm. The next section of the paper discusses the prototype implementation adopted to realize the SSFSWS.

4. CASE STUDY AND EXPERIMENTAL EVALUATION

This section of the paper discusses a small case study undertaken to prove the functional feasibility of the SSFSWS. The case study considered a popular semantic corpus Edubase2 [33][34][35]. The Edubase consists of the records of the educational establishments in England and Whales. This corpus is maintained by the Education Department of the United Kingdom Government [36]. The Edubase2 corpus is huge and in order to construct an SOA architecture the Edubase2 data was split into three categories namely establishments offering primary education , establishments offering secondary education and educational establishments offering higher education and collegiate education. Each corpus resulting from the classification discussed is considered as the RDF KB of the semantic search providers. The SOA adopted to demonstrate the functionality of the SSFSWS is as shown in Figure 3.

ISSN 2229-5518

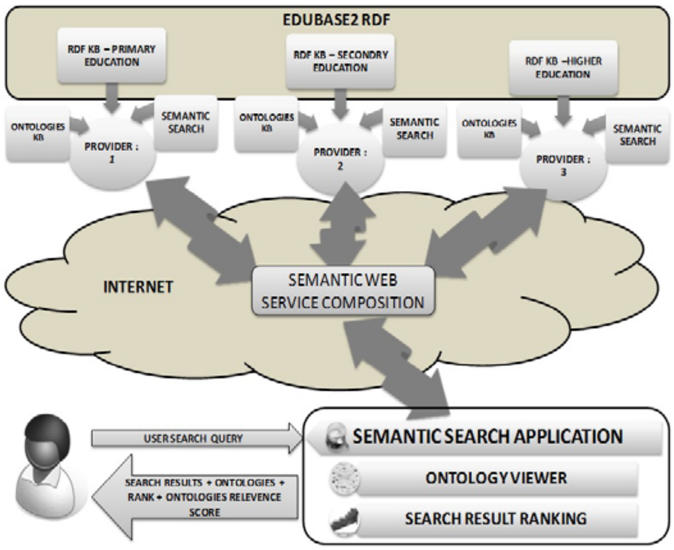

Figure 3: Case Study Architecture Based on EDUBASE

Microsoft .Net platform 4.0 is considered for the development of the case study presented here. The three semantic service providers offering the semantic search web services were developed on the ASP.Net platform. The web services offered were composed using the depth first search algorithm. The semantic search application was developed on a Windows Presentation Platform using C#.net and Extensible Application Markup Language (XAML) as the implementation language. The ontology viewer which basically is an interface application to display the ontology graphs resulting from the search query was also developed using C#.net and XML. Most of the research work done in the area of semantic search preferred a Java based platform for implementation this is a major motivation to develop the SSFSWS on the .Net Platform.

The RDF KB obtained from the Edubase2 corpus consists of a total number of 66655 records of various educational establishments in England and Whales. A comprehensive data of nearly 218 concepts per school were provided. The corpus on splitting in the terms of the type of education levels provided consists of 24167 establishments offering primary education , 5099 providing secondary education and 37389 establishments offering higher education. The RDF KB housed in the three semantic search providers was considered to build ontologies and represent the relations using $SROIQ - DL$. The number of relations extracted based on the RDF KB housed with each semantic search service providers is as shown in Table 2. The number of RDF Records available with each of the semantic search providers is shown in the Figure 4. The average number of relations extracted was found to be around 121. Figure 5 shows the number of relations extracted based on the RDF KB available with the semantic service providers.

Education Type Offered	Number of RDF Records in Corpus	Number of Relations Extracted
HIGHER EDUCATION	37389	3663441
SECONDARY EDUCATION	5099	764476
PRIMARY EDUCATION	24167	3653042

Table 2: RDF KB and Ontologies KB Statistics

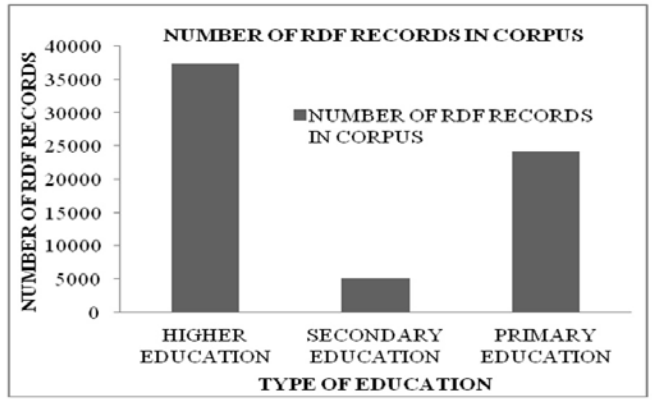

Figure 4: Size of the RDF KB Available with the Semantic Search Providers

The semantic search application supports ontologies built using the $SROIQ - DL$ language. A sample ontology obtained after performing the search is shown in Figure 6. The presented case study is compared with the existing semantic search offered using Edubase/Edubase2 corpus maintained by the Education Department of the UK Government. [37][38]. The existing semantic search available online provides no support for relational queries and supports no logical or Boolean operators in the search query. The SSFSWS case study developed over comes these drawbacks and supports logical queries, Boolean operator based queries and queries using special characters. In addition the proposed SSFSWS also provides the ontology view of the response obtained for a query. Ranking the search results based on the ontology relevance score provides more legitimate search responses to be provided to the users.

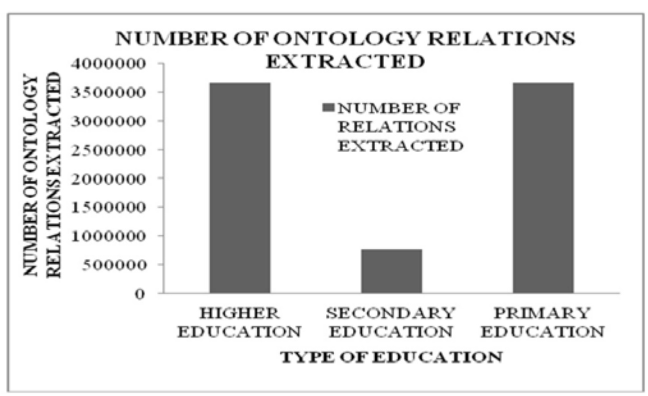

Figure 5: Number of Ontology Relations Extracted to form the Ontology KB

ISSN 2229-5518

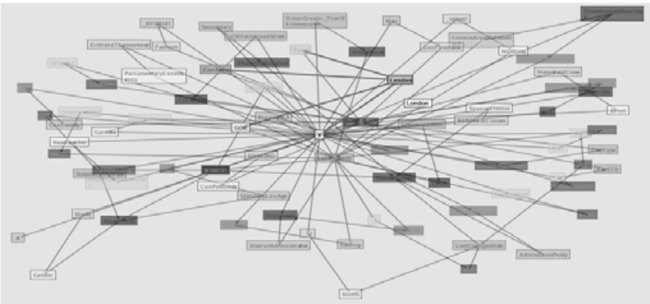

Figure 6: Ontology View Generated at the Semantic Search Application

5. CONCLUSION AND FUTURE WORK

The designed framework named SSFSWS is based on SOA semantic search. It also offers semantic search web services and a semantic search application. The depth first search algoithm is used for composing semantic search web services. RDK KB and ontologies KB gives the bases for the semantic search responses in the SSFSWS framework. It also constructs using the SROIQ-DL web ontology language. The response time of the search result can be improvised by using hierarchical caching mechanisms. Ranking based on the ontology relevance score is also supported by SSFSWS. The semantic web service composition was posted by the construction of ontologies. The constructed ontologies also provide semantic search response which enables better graphical representation and analysis for the search query provided for users. The image based search can be supported by providing SSFSWS and also semantic web services composition algorithm can be optimized in future.

REFERENCES

[1]T. Berners-Lee and M. Fischetti, “Weaving the Web,” Chapter Machines and the Web, pp. 177-198, 1999.

[2]W3C 1999 specification."Resource Description Framework (RDF) Model and Syntax Specification",W3C Recommendation 22 February 1999

[3] Web Ontology Language, http://www.w3.org/2004/OWL/

[4] OWL 2 Web Ontology Language, http://www.w3.org/TR/owl2-overview/

[5]T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic-Web, " Scientific American Magazine, vol. 284, no. 5, 2001, pp. 34–43

[6]Sheila A. McIlraith, Tran Cao Son, and Honglei Zeng."Semantic Web Services",IEEE INTELLIGENT SYSTEMS,MARCH/APRIL 2001 pp 46-53

[7]M. Burstein, C. Bussler, M. Zaremba, T. Finin, M. N. Huhns, M. Paolucci, A. P. Sheth, and S. Williams, "A Semantic Web Services Architecture" IEEE Internet Comput. , vol. 9, no. 5, 2005, pp. 72–81.

[8]Yufei Li, Yuan Wang, and Xiaotao Huang,"A Relation-Based Search Engine in Semantic Web",IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 19, NO. 2, FEBRUARY 2007. pp 273-282

[9]Horrocks I, Kutz O, Sattler U. The even more irresistible SROIQ. 2006

[10]Moldovan, D.I. and Mihalcea, R,"Using wordnet and lexical operators to improve internet searches". IEEE Internet Computing 4 (2000) pp 34–43

[11]D. Tümer, M. A. Shah, and Y. Bitirim, An Empirical Evaluation on Semantic Search Performance of Keyword-Based and Semantic Search Engines: Google, Yahoo, Msn and Hakia, 2009 4th International Conference on Internet Monitoring and Protection 2009.

[12]Manola, F. and McBride, B. (2004) ‘RDF primer, W3C recommendation 10 February 2004’, Technical Report,W3C, http://www.w3.org/TR/rdf-primer/ (2004-07-15).

[13]Davies, J. and Weeks, R ‘QuizRDF: search technology for the semantic web’, 37th Hawaii International Conference on System Sciences (HICSS-37 2004), CD-ROM/AbstractsProceedings, 5–8 January, Big Island, HI, USA, IEEE Computer Society, ISBN 0-7695-2056-1.

[14] SEWISE: www.georges.gardarin.free.fr/Articles /Sewise_NLDB2003.pdf

[15]Heflin, J. and Hendler, J. ‘Searching the web with SHOE’,Artificial Intelligence for Web Search. Papers from the AAAI Workshop, WS-2000-01,pp.35–40

[16]A. Maedche, B. Motik, L. Stojanovic, R. Studer, and R. Volz, "An infrastructure for searching, reusing and evolving distributed ontologies," in Proceedings of WWW '03Budapest, 2003, pp. 439-448

[17]Rocha, C., Schwabe, D. and de Aragao, M.P. ‘A hybrid approach for searching in the semantic web’, WWW ‘04: Proceedings of the Thirteenth International Conference on World Wide Web, pp.374–383

[18]Burton-Jones, A., Storey, V.C., Sugumaran, V. and Purao, S. ‘A heuristic-based methodology for semantic augmentation of user queries on the web’, Conceptual Modeling – ER 2003, 22nd International Conference on Conceptual Modeling, October 13–16, Proceedings, pp.476–489

[19]Amaral, C., Laurent, D., Martins, A., Mendes, A. and Pinto, C “Design and implementation of a semantic search engine for Portuguese”, Proceedings of 4th International Conference on Language Resources and Evaluation, LREC 2004, Vol. I, pp.247–250

[20]Christoph Mangold. "A survey and classification of semantic search approaches",Int. J. Metadata, Semantics and Ontology, Vol. 2, No. 1, 2007. pp 23-34

[21]Eetu Makel. "Survey of Semantic Search Research",www.seco.tkk.fi/publications/2005/makela-semantic-search-2005.pdf

[22]F. F. Ramos, H. Unger, V. Larios (Eds.): LNCS 3061, pp. 145–157, Springer-Verlag Berlin Heidelberg 2004

[23]Cohen, S. Mamou, J. Kanza, Y. Sagiv, Y “XSEarch: A Se-

mantic Search Engine for XML" proceedings of the international conference on very large databases, pp 45-56, 2003.
[24]E. Kandogan, R. Krishnamurthy, S. Raghavan, S. Vaithyanathan, and H. Zhu, "Avatar semantic search: a database approach to information retrieval," in Proceedings of SIGMOD '06 Chicago, 2006, pp. 790-792
[25]D. Bhagwat and N. Polyzotis, "Searching a file system using inferred semantic links," in Proceedings of HYPERTEXT '05 Salzburg, 2005, pp. 85-87
[26]Yuzhong Qu and Gong Cheng, "Falcons Concept Search: A Practical Search Engine for Web Ontologies",IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART A: SYSTEMS AND HUMANS, VOL. 41, NO. 4, JULY 2011,pp 810-816
[27]Xiaoou Tang, Ke Liu, Jingyu Cui, Fang Wen, and Xiaogang Wang."IntentSearch: Capturing User Intention for One-Click Internet Image Search", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 34, NO. 7, JULY 2012,pp 1342-1353
[28]Xinmei Tian, Yijuan Lu, and Linjun Yang," Query Difficulty Prediction for Web Image Search", IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 14, NO. 4, AUGUST 2012, pp951-962
[29]David Booth and Canyang Kevin Liu. Web Services Description Language (WSDL) Version 2.0 Part 0: Primer. World Wide Web Consortium (W3C), June 26, 2007. W3C Recommendation. Online http://www.w3.org/TR/2007/REC-wsdl20-primer-20070626
[30]David Martin, Mark Burstein,Grit Denker,Daniel Elenius,Joseph Giampapa,Drew McDermott,Deborah McGuinness,Sheila McIlraith,Massimo Paolucci ,Bijan Parsia ,Terry Payne ,Evren Sirin,Naveen Srinivasan and Katia Sycara. "OWL-S 1.2 Release", OWL-based Web Service Ontology.Web-Ontology Working Group at the World Wide Web Consortium, 2008-12.Online : http://www.ai.sri.com/daml/services/owl-s/1.2/
[31] RDF/XML Syntax Specification ,http://www.w3.org /TR/rdf-syntax-grammar.
[32] N-Triples,http://www.w3.org/2001/sw/RDFCore/ntriples.
[33] Edubase RDF, http://source.data.gov.uk/data/education /edubase/2009-08-14/edubase-rdf-r2rc3.tgz
[34]R.Guha, R.McCool and E.Miller "Semantic Search",WWW2003, May 20-24, 2003, Budapest, Hungary.ACM 1.58113-680-3/03/0005 pp 700-709
[35]Jonathan A. Abourbih, Luke Blaney, Alan Bundy, and Fiona McNeill, "A Single-Significant-Digit Calculus for Semi-Automated Guesstimation"Springer-Verlag Berlin Heidelberg ,IJCAR 2010, LNAI 6173, pp. 354–368, 2010
[36]Department of Education, www.education.gov.uk
[37] Department of Education,http://services.data.gov.uk /education/search Accessed May 26th 2012
[38]Department of Education, http://www.education.gov.uk/schools/performance/index.html Accessed September 1st 2012

AUTHORS

P. Nandhakumar (nandhap@gmail.com) completed M.C.A., M. Phil. and currently pursuing Ph.D in computer science at Karpagam University, Coimbatore under the guidance of Dr.M.Hemalatha, Professor and Head, Dept. of Software System, Karpagam University, Coimbatore. He is working as Senior Software Engineer in Easy Design Systems Private Limited, Coimbatore.

Dr. M. Hemalatha (csresearchhema@gamil.com) completed M.Sc., M.C.A., M. Phil., Ph.D (Ph.D, Mother Terasa women's University, Kodaikanal). She is Professor & Head and guiding Ph.D Scholars in Department of Computer Science at Karpagam University, Coimbatore. Twelve years of experience in teaching and published more than hundred papers in International Journals and also presented more than eighty papers in various national and international conferences. She received best researcher award in the year 2012 from Karpagam University. Her research areas include Data Mining, Image Processing, Computer Networks, Cloud Computing, Software Engineering, Bioinformatics and Neural Network. She is a reviewer in several National and International Journals.