Semantic Based Graph Mining Technology for Automatic Service Semantic Link Network Navigation

Anping Zhao
College of Computer and Information Science, Chongqing Normal University, Chongqing, China
Email: zap@swu.edu.cn

Yu Yu
College of Trade and Public Administration, Chongqing University, Chongqing, China
Email: fish@swu.edu.cn

***Abstract*— To reduce the user's cognitive load and make sense of complex interlinked service data for realizing advanced service application, an approach to Service Semantic Link Network (S-SLN) navigation was proposed. Semantic based graph mining technology was employed as an effective solution for automatic service navigation. As one of the most important navigation tools, related service nodes recommendation was mapped as a basic graph mining problem. The closely related service nodes linked by different relations to form a sub-graph of S-SLN, which service nodes and links with different type and weight represent different service scenarios and relationship to the maximum extent such that sub-graph of S-SLN can improve or speed up user's navigation decisions by providing the most probable related Web service and links which are most likely to be of interest. The experimental results and analysis shows that the proposed approach is efficient, feasible and practical.**

***Index Terms*—semantic, graph mining, semantic link network, service navigation**

I. Introduction

With the increasing number of Web service available on the Web, an additional need that arises with the growing number of entries is how to efficiently locate a reasonably small group of service candidates from the enormously services space and to automatically create a navigation list of related services that match a given query. Today's applications which provide access to information typically use two navigation methods, browsing and search. A browser provides a number of links to other items or documents, while a search system shows relevant items given user's query. The efficiency of human navigation in the enormously services space depends on the availability of a powerful and intelligent backend which provides guidance and recommendations [1]. Service Semantic Link Network (S-SLN) is the intelligent semantic model to define semantic structure among Web services by a relationship dependency network which connects services with different types of relationships [2], in which Web services are organized and their functionalities semantically described. Its nodes can be any types of resources. Its edges can be any semantic relations. A S-SLN instance is a directed graph, denoted as *S(ServiceSet, LinkSet)*, where *S* is the name of the S-SLN, *ServiceSet* is a set of services, and *LinkSet* is a set of semantic links in form of $R \xrightarrow{\alpha} R'$, where R, $R' \in$ *ServiceSet*, and α is a semantic factor representing a semantic relation between R and R'. S-SLN is the underlying semantic model for effectively implementing automatic Web service search and composition by a relationship dependency network which connects services with different types of relationships. It allows us to navigate Web service by their semantic relationship. In navigation on S-SLN, one of the most important guiding tools is related service recommendation that is given a set of nodes on an S-SLN, to recommend potentially relevant service nodes. The role of related service recommendation is to reduce cognitive load, provide guidance in navigation and browsing, contextualize, simplify, and make sense of otherwise complex interlinked service data.

As an advanced application of S-SLN, Automatic service recommendation and navigation is a major objective of S-SLN. This paper focuses on how to develop an approach for automatic service navigation generation. We introduce a service navigation method in the S-SLN, which utilizes semantic based graph mining technique. It provides the convenience of having services presented in a way tailored to a particular user, situation or context.

The contribution of this paper is as follows:

(1) We incorporate semantic relationship of S-SLN to reduce the user's cognitive load for facilitating service navigation.

(2) Development of a graph mining technology based approach for automated service navigation of S-SLN from complex interlinked service data.

This work was partially supported by the Scientific Research Foundation for the Doctor Scholar of Chongqing Normal University and the Fundamental Research Funds for the Central University under Grant Project No.CDJRC10010010.
Corresponding author: zap@swu.edu.cn.

doi:10.4304/jsw.7.5.1022-1029

(3) We conduct preliminary experiments and performance studies, which verify the feasibility and effectiveness of our methods.

The rest of this paper is organized as follows: In section II we review related works, Section III describes the approach to S-SLN service navigation based on graph mining technology. We show the effectiveness of the presented approach by experimental result in section IV. Finally, we summarize our conclusions and the future research in section V.

II. RELATED WORKS

Several topics are closely related to graph mining, but have a different focus. Relational learning looks at the graph formed by interlinked relations in a database and attempts to find patterns in it. Studies of rumor or viral propagation in a network look for key properties of the network which determine the susceptibility to epidemics. New graph navigation algorithms try to devise graphs so that local routing decisions can lead to nearly-optimal global routes. The mining of frequent patterns was first introduced in a databases context by Agrawal and Srikant [3], and is possibly one of the most popular data mining techniques. Recently, these ideas have been extended and applied to large graph datasets in order to find the most common patterns or motifs hidden in the graph structure, to compress the graph dataset, and for many other problems. The participants in Milgram's experiment [4] were able to build a chain to an unknown target individual, even though they knew only a few individuals in the full social network. We can navigate to Web sites containing the information we need in spite of the immense size of the Web. Such facts imply that large real-world graphs can be navigated with ease. There is a complete overview about the subject in paper [5], which discussed methods of navigation that can be employed and why real world graphs seem to be so easy to navigate.

Related node recommendation is one of the most important guiding tools for navigation in graph, Spreading activation approach as a technology for recommendation in various kinds of networks belongs to a broader group, which is typically referred to as graph mining based approaches for navigation. In paper [6], author introduces an approach for explaining recommendations in environments that are based on semantic models. Using a constrained Spreading Activation (CSA) technique for recommendation generation, they store and exploit the activation paths leading to recommendations. A mining method for egocentric and polycentric queries in multi-dimensional networks is proposed [7]. The method uses spreading activation technique to combine with so-called ambient navigation for collaborative filtering and community detection based on tag recommendations. Authors describe a method for graph based related item recommendation and present Nepomuk Simple a GUI for managing "piles" of desktop resources [8]. The process of incremental modification of a "pile" might be formally viewed as the process of browsing and exploring data provided by networks, assisted by recommendations resulting from fuzzy polycentric queries. Using the ontology, the search engine is proposed which combines content based search techniques with spread activation techniques applied to a semantic model of a given domain [9]. It is based on a hybrid spread activation algorithm applied to the concept instances graph to find the related concepts in the ontology based on the given query. Social tagging systems present a new challenge to the researchers working on recommender systems. In paper [10], authors argue for the use of spreading activation approach for building tag-aware recommender systems and suggest a specific version of this approach adapted to the multidimensional nature of social tagging networks. Social spaces such as blogs, wikis and online social networking sites are enabling the formation of online communities where people are linked to each other through direct profile connections and also through the content items that they are creating, sharing and tagging. As this space become bigger and more distributed, more intuitive ways of navigating the associated information become necessary. Paper [11] demonstrates the linked semantic data can provide an enhanced view of the activity in a social network, and the Galaxy tool can augment objects from social spaces, by highlighting related people and objects, and suggesting relevant sources of knowledge.

Even though the application area is different in this work, some of the ideas are built on a common ground. As compared with our work, a methodology to automatic navigate service in S-SLN based on the semantic relationship, which can help to reduce the user's cognitive load and make sense of complex interlinked service data for realizing advanced service application, is to be discussed.

III. SEMANTIC BASED GRAPH MINING FOR SERVICE NAVIGATION

A. Problem Description and Basic Idea

In navigation on S-SLN, related service recommendation is one of the most important guiding tools, which is given a set of nodes on the S-SLN, to recommend potentially relevant service nodes. We desire that navigation in our approach is on-the-fly transformation of the underlying service network in such a way that users can focus on one or more services in the S-SLN, and immediately see a conceptual summary of their focus, in the form of transformed reduced service link network, in which unrelated service will be pruned, but not removed completely, and highly relevant services will be brought to the user's attention even if they were not explicitly linked to current user's focus. So users can navigate Web services in a guided yet unconstrained way. Related service recommendation in S-SLN is different from service search, since the goal of recommendation is not to find services with particular properties, because the user himself frequently would not be able to specify what exactly he would like to have as a recommendation, but the search of services with strong cumulative direct and indirect connections to the initial set of services.

Therefore we consider the problem of related service recommendation in S-SLN as a problem of graph mining on multidimensional service semantic link networks. We believe that recommendation of related services for a set of nodes on an S-SLN is a problem which can be addressed by a blend of fundamental graph-mining techniques like Spreading Activation. The aim of service recommendation is not to make choices for the user, but to improve or speed up user's navigation decisions by providing the most probable related Web service and links which are most likely to be of interest.

A successful service recommendation approach for the context of S-SLN should fully employ the semantic link network structure of Web service and use all kinds of service semantic links. We think that the most promising in this context is the spreading activation approach. We employ the spreading activation method as an effective solution for automatic service recommendation. On one hand, one can think that related services are those services which are connected to all of the services in a query by service semantic links of various types which represent the relationships between those services. On the other hand, we can suggest that related services are those nodes on the S-SLN which minimizes the weighted semantic distance to all nodes in the query, and therefore related services are those nodes which might be considered as centroid of semantic neighborhood induced by queries. The notion of Semantic Neighborhood of a node was formally defined, which represents the semantically associated neighborhood of interest for a particular service node. We then arrive at an activation spread model for identifying the members of a node's Semantic Neighborhood. Our goal here is to identify a semantic neighborhood of interest for a given source service node as well as quantify the semantic relationship as the semantic relevance of different service nodes in the neighborhood on the source node. The model uses a semantic based activation function to select nodes of interest with respect to a particular source node, and also to quantify the level of their involvement using commitment values.

B. Preliminary

Spreading activation (also known as spread of activation) is a method for searching associative networks, neural networks or semantic networks. The method is based on the idea of quickly spreading an associative relevancy measure over the network [1]. A formal Spreading Activation Methods Framework can be described as follows.

Output functions $out_v : R \to R$

$$o_v^{(k)} = out_v(a_v^{(k)}) \tag{1}$$

An output function determines the outgoing activation $o_v^{(k)}$ of a node v at time k based on its current activation level.

Input functions $in_v : R^n \to R$

$$i_v^{(k)} = in_v(o^{(k-1)}) \tag{2}$$

An input function aggregates incoming activation of a node v and defines its input $i_v^{(k)}$ at time k.

Activation functions $Act : R \to R$

$$a_v^{(k)} = Act(i_v^{(k)}) \tag{3}$$

An activation function determines the level $a_v^{(k)}$ of activation resulting from the input, and thus in particular whether a node is activated, i.e. whether its activation will be spread in the next iteration.

Activation Constraints

To avoid the activation of all nodes that receive activation at all, a threshold function can be applied.

Data processing

Data processing consists of iterations that contain a set of pulses including checks for termination conditions. In its simplest form, the activation level of a node in spreading activation is determined trough the following formula:

$$l_j = \sum o_i \omega_{ij} \quad (i = 1,\ldots,k) \tag{4}$$

Where

l_j is the activation level of node j , calculated as the total input to node j ;

o_i is the output of unit i , i connected to node j ;

ω_{ij} is a weight associated to the link connecting i and j .

In the following section, we concern ourselves with the following important question. How can we identify the Semantic Neighborhood of interest for a particular source service node, and also develop the activation function based on the quantified service semantic relationship among different nodes, in the S-SLN, with regard to the source service node.

C. Semantic Based Activation function

It is possible to incorporate semantic properties of nodes to encode the activation function. This is of particular importance in intelligent service recommendation applications, where one is interested in identifying subgraphs that match given service nodes.

To obtain eminent semantic neighborhoods when nodes are linked with semantic relationship, we need to consider the semantic relevance of nodes with the source node in S-SLN. Given service S_i and S_j of S-SLN, the semantic relevance between S_i and S_j is a quantitative representation of the degree of correlation based on their semantic relationship, denoted as $R_s(S_i, S_j)$.

We use *TF/IDF* (Term Frequency/Inverse Document Frequency) to compute semantic relevance of between services of S-SLN. TF/IDF is known as one of the most important theory discovery of information retrieval [12]. The TF/IDF weight is a weight often used in information retrieval and text mining. This weight is a statistical measure used to evaluate how important a term is to a document in a collection or corpus. The importance increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus. Variations of the TF/IDF weighting scheme are often used by search engines as a central tool in scoring and ranking a document's relevance given a user query.

The semantic relevance computation problem is to evaluate how important a link between two services is to a service invocation link in a collection of service invocation links. We may solve relevance computation by the way of TF/IDF problem. We map the relevance problem to TF/IDF as follows.

•Each link L_i between two services is mapped to a term t_i in documents.

•Each service invocation link L_i' is mapped to a document d_j in a collection.

•The total number of service invocation links N_S is mapped to the total number of documents N in the corpus.

The service semantic relevance computation is thus formulated as follows:

$$R_s(S_i, S_j) = \frac{f_{i,j} * f_i'}{\sqrt{\sum_1^M (f_{i,j} * f_i')^2}} \tag{5}$$

Where

$f_{i,j}$ is the service link frequency.

$$f_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \tag{6}$$

$n_{i,j}$ is the number of occurrences of the considered link L_i in service invocation link L_i'.

f_i' is the inverse link frequency.

$$f_i' = \log(\frac{N_s}{n_i'}) \tag{7}$$

n_i' is the number of service invocation links L_i' which contains the link L_i.

N_s is the total number of service invocation links L_i' in the collections.

M is the number of links in L_i'.

The activation function is used by the spread algorithm to perform the distribution of different amounts to different nodes depending on topological or semantic features. For a node *v*, the activation function $Act_v()$ is a non-negative function that maps the semantic relevance associated with the neighbors of *v*. Key to the effectiveness of the spread is an activation function Act() that serves as a distribution mechanism.

For further discussion, we formally define the notion of a Semantic Neighborhood of a node, which represents the semantically neighborhood of interest for a particular node.

A Service Semantic Link Network S-SLN can be modeled as a directed graph $G = (V, E, w)$,

Where

$V=\{S_1,S_2,..,S_n\}$ is the set of service;

$E = \{e_1, e_2, \ldots, e_m\}$ is the set of service semantic links;

$w : E \rightarrow R$ is the weight of service semantic links, in particular $w(u,v) = 0$ *if* $(u,v) \notin E$;

$Neighbor(v) = \{u : \{u,v\} \in E\}, v \in V$, is the set of neighbors of node v ;

Activation is spread on a graph G and across incident edges to adjacent nodes and activates these nodes as well.

Definition 1 Let a_v^s represent the level of activation of the node v resulting from the node s , a Semantic Neighborhood of service s $SN = (V, E)$ is a subgraph of $G = (V, E, w)$, which is rooted at s , $V' \in V$, $\forall v \in V'$, $a_v^s > \tau$, τ is a threshold ; $E' \in E$, $\forall (v_i, v_j) \in E', v_i \in V', v_j \in V'$ 。

Definition 2 Let a_v^s represent the level of activation of the node v resulting from the node s , given a *Semantic Neighborhood* of service s , $SN = (V, E)$, $\forall x \in V$,the descendants of *x* are given as $Desc(x) = y : (x, y) \in E$, such that $a_x^s >= a_y^s$.

Let *s* denote the source node. Let the semantic relevance of the descendants of *x* with respect to the source node be

$$R_s(Desc(x)) = \sum_{i \in Desc(x)} R_s(s,i) \tag{8}$$

Given a source node $s \in V$, $\forall x \in V$, the activation function can be defined as:

The amount retained by node *x*:

$$Act(x,s) = \frac{R_s(s,x) \times m_x}{R_s(Desc(x)) + R_s(s,x)} \tag{9}$$

The amount received by each of its siblings and descendants:

$$Act(y,x) = \frac{R_s(s,y) \times m_x}{R_s(Desc(x)) + R_s(s,x)} \tag{10}$$

Where m_x is the total amount of node *x*.

D. Mining Semantic Neighborhood from S-SLN

Our goal here is to identify a semantic neighborhood of interest for a given source service node or a set of source service nodes as well as quantify the relationship or effect of different nodes in the semantic neighborhood on the source service node in the S-SLN. We mining a Semantic Neighborhood *SN* for a given source service node from S-SLN as a recommendation of service.

The initial data for recommendations which allow browsing from one single service to another might be properly treated as a single source recommendation.

Definition 3 Given a source service node *S*, the single source recommendation problem is to find the Semantic

Neighborhood *SN* that represents the sub-graph rooted at node *S* such that each node in *SN* are semantically closely related to *S*.

The activation begins at the source node where we assume a budget *M* is initially available. The source node distributes *M* among its neighbors linked by service semantic links, initiating the activation process. Each node then retains some fraction of the amount for itself and splits the remainder among its descendants. Thus a node is expanded only once. If a node has already been activated, it is not expanded a second time. To handle the inverse-weighting of nodes, we decay the activation as it proceeds farther away from the source node. Each time the activation touches a node, it decays by a factor of the number of the semantic links the service node has to ensure that service nodes closer to the source node are more probable to be chosen. A semantic based activation function $Act_v()$ serves as a distribution mechanism to ensure the effectiveness of the spread. The activation proceeds with the amount constantly decaying until reaching a minimum threshold, at which it is deemed indivisible. At the end of the activation process, the total of the amounts at each service node sums to M. Hence, the fraction of the total amount that each service node has received gives the commitment value for a node in this particular Semantic Neighborhood *SN*.

We also can extend the single source recommendation problem to one where the initial data for recommendations allow browsing from many services to another. And we wish to identify the common shared semantic neighborhood for a set of source service nodes, which is so-called multi-source recommendation. It is important in more complex intelligent applications.

Definition 4 Given a set of *n* source nodes $S=(s_1,s_2,\ldots,s_n)$, the multi-source recommendation problem is to find the Semantic Neighborhood *SN* that represents the intersection of at least *k* of their semantic neighborhoods.

Based on this, semantic based spreading activation on a S-SLN can be used to find semantic neighborhood service sub-graph linked with semantic links as the related service recommendation of interest service node. In such a way we can acquire additional knowledge from the network about the service topic of interest source service.

As stated above, the main steps of mining Semantic Neighbourhood for the given source service nodes from S-SLN can be summarized as follows:

A directed graph S-SLN is populated by Web services $S=(s_1,s_2,\ldots,s_n)$ each having an associated activation value A[i] which is a real number in the range[0.0 ... 1.0]. A semantic link R[i, j] connects source service node s_i with target node s_j. Each link has an associated semantic relevance weight $R_s(S_i,S_j)$ usually a real number in the range [0.0 ... 1.0].

Parameters:

1. Firing threshold *F*, a real number in the range [0.0 ... 1.0].

2. Decay factor *D*, a real number in the range [0.0 ... 1.0].

Steps:

1. Initialize the S-SLN setting all activation values A[i] to zero. Set one or more origin nodes to an initial activation value greater than the firing threshold F. A typical initial value is 1.0.

2. For each unfired service node s_i in the graph having activation value A[i] greater than the node firing threshold *F*:

3. For each semantic link R[i, j] connecting the source service node s_i with target node s_j, adjust $A[j]=A[j]+(A[i]\times R_s(S_i,S_j)\times D)$ where *D* is the decay factor.

4. If a target node receives an adjustment to its activation value so that it would exceed 1.0, then set its new activation value to 1.0. Likewise maintain 0.0 as a lower bound on the target node's activation value should it receive an adjustment to below 0.0.

5. Once a service node has fired it may not fire again, although variations of the basic algorithm permit repeated firings and loops through the graph.

6. Nodes receiving a new activation value that exceeds the firing threshold *F* are marked for firing on the next spreading activation cycle.

7. If activation originates from more than one node, a variation of the algorithm permits marker passing to distinguish the paths by which activation is spread over the S-SLN.

8. The procedure terminates when either there are no more nodes to fire or in the case of marker passing from multiple origins, when a node is reached from more than one path. Variations of the algorithm that permit repeated node firings and activation loops in the S-SLN, terminate after a steady activation state, with respect to some delta, is reached, or when a maximum number of iterations is exceeded.

IV. Experimental Result

In this section, we will present the preliminary experiments to test the effectiveness and feasibility of our proposed service recommendation approaches.

A. Dataset

We conducted experiments using a publicly available test set EEE-05 Web service dataset [13]. The purpose of the EEE-05 Challenge is to establish a venue where researchers can collaborate on implementations in the web service composition domain. The EEE05 test set is a synthetic one that contains artificially created composition scenarios, and the test set and test requests appear to be manually created by human experts. The test set contains just 100 WSDL files. Although the test set is small, the EEE05 test set is still challenging because it is not simple for humans to solve them optimally in a short time. We selected 44 Web services among them and manually created a service semantic link network S-SLN on which activation is spread. The Web service nodes of

S-SLN and relationship information between them are shown as Table I.

TABLE I.
THE WEB SERVICE NODES OF S-SLN AND RELATIONSHIP INFORMATION

*Vertices 44	
1 "checkOutLibraryBook"	2 "findCloseLibrary"
3 "createItinerary"	4 "StateInfo"
5 "CityInfo"	6 "StreetInfo"
7 "airlineName"	8 "CompanyName"
9 "hotelInfo"	10 "ZipCode"
11 "findBanjoInstructor"	12 "PhoneNumber"
13 "findCloseAPT"	14 "findCloseBank"
15 "bankInfo"	16 "CarDealer"
17 "dealerType"	18 "CarRental"
19 "findDoctor"	20 "pickupLocation"
21 "bookISBN"	22 "LocalAttraction"
23 "attractionPrice"	24 "commodityType"
25 "findMostRelevantStock"	26 "getForecast"
27 "stockName"	28 "getStockPrice"
29 "StockResearch"	30 "creditCard"
31 "purchaseALT"	32 "Airport"
33 "ALTPrice"	34 "flightPlan"
35 "purchaseBook"	36 "LocalAttractionTicket"
37 "purchaseStock"	38 "reserveAPTRoom"
39 "hotelName"	40 "reserveDateRestaurant"
41 "scheduleDate"	42 "scheduleShipment"
43 "shipmentCost"	44 "banjoStyleToLearn"

*Edges

1 2 ; 1 4; 1 5; 1 6; 1 21; 2 4; 2 5; 2 6; 2 10; 2 12; 3 4; 3 5 ; 3 6; 3 7; 3 9; 3 15; 3 26; 3 22; 3 30; 3 32; 3 33; 3 38; 3 39; 4 5; 4 6; 4 10; 4 12; 5 6; 5 10; 6 10; 8 4; 8 5; 8 6; 8 10; 8 12; 7 4; 7 5; 7 32; 7 33; 9 4; 9 5; 9 6; 9 10; 9 12; 9 39; 11 4; 11 5; 11 6; 11 10; 11 44; 13 4; 13 5; 13 6; 13 10 ; 13 12; 14 4; 14 5; 14 6; 14 10; 14 12; 14 15; 15 4; 15 5; 15 6 ; 15 10; 15 12; 16 4; 16 5; 16 6; 16 10; 16 12; 16 8; 16 17; 16 18; 17 24; 18 4; 18 5; 18 6; 18 10; 18 12; 18 8; 19 4 ; 19 5; 19 6; 19 10; 19 12; 22 23; 25 27; 25 28; 25 29; 26 4 ; 26 5; 27 29; 27 8; 28 8; 28 25; 28 29; 29 27; 29 28; 30 15 ; 31 4; 31 5; 31 7; 31 12; 31 20; 31 30; 31 33; 31 34; 32 4; 32 5; 32 12; 33 5; 33 7; 34 4; 34 5; 34 7; 34 26; 34 31; 35 20 ; 35 21; 35 30; 36 22; 36 23; 37 25; 37 27; 37 28; 37 29; 37 30; 38 4; 38 5; 38 6; 38 9; 38 12; 38 39; 38 41; 40 5; 40 6; 40 41; 42 4; 42 5; 42 6; 42 12; 42 20; 42 43; 41 34;

B. Results

We set the initial activation value as one to begin at the source node. The source node distributes initial activation value among its neighbours linked by service semantic links, initiating the activation process. And the minimum activation threshold was set as 0.020 by human expert experience.

The experimental results of single source recommendation are shown in Figure. 1. The source node "*FightPlan*" is shown in yellow. The activation path and activation value is shown in Table II. The found Semantic Neighbourhood is consist of blue nodes "*StateInfo*", "*CityInfo*", "*GetForecast*" and "*ALTPrice*", which is semantically closely related to the source service node "*FightPlan*".

TABLE II.
THE ACTIVATION PATH AND ACTIVATION VALUE OF SINGLE SOURCE RECOMMENDATION

No.	Activation Path	Activation Value
1	FightPlan-StateInfo	0.160
2	FightPlan- GetForecast	0.148
3	FightPlan-CityInfo	0.235
4	FightPlan-CityInfo-HotelInfo	0.035
5	FightPlan-CityInfo-HotelInfo-reserveAPTroom	0.023
6	FightPlan-ScheduleDate	0.210
7	FightPlan-airlineName	0.247
8	FightPlan-airlineName-ALTPrice	0.187
9	FightPlan-airlineName-ALTPrice-purchaseALT	0.027

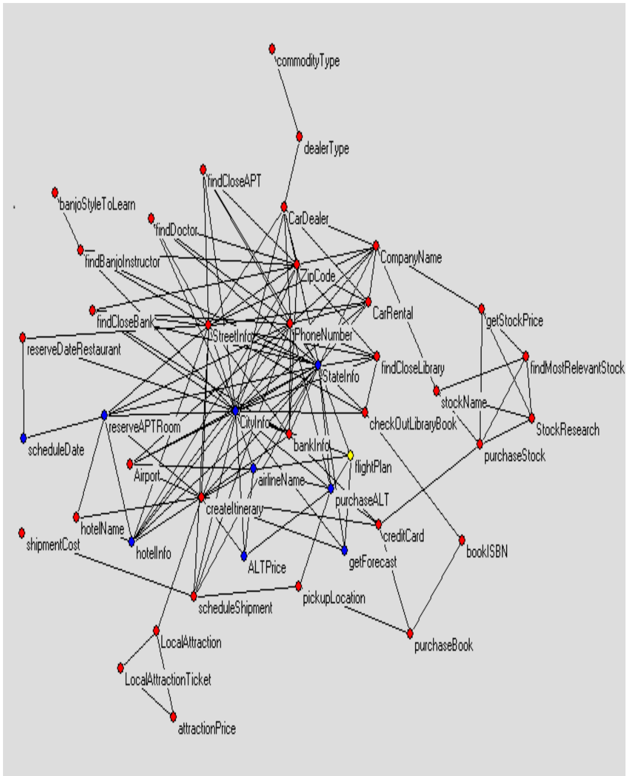

Figure 1. Example of single source recommendation

The experimental results of multi-source recommendation are shown in Figure. 2. Two source nodes "*purchaseALT*" and "*reserveAPTRoom*" in S-SLN are shown in yellow. The mined Semantic Neighbourhood is consist of blue nodes with high activation value, such as "*StateInfo*", "*CityInfo*" and "*creditCard*", which is consist of the common shared semantic neighbourhood and semantically closely related to the source service nodes "*purchaseALT*" and "*reserveAPTRoom*".

The experimental results show that the activated service nodes in S-SLN have higher semantic relevance with the source services, and form the semantically and

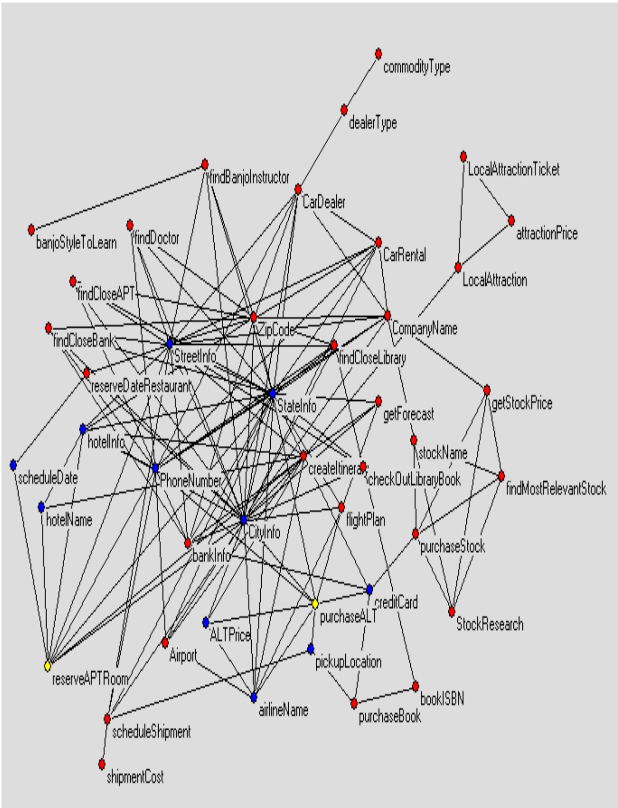

Figure 2. Example of multi-source recommendation

functionally close related Web services sub network of the S-SLN. These experimental results verify the effectiveness of our presented approach. As a result, semantic based graph mining technology is an effective method for S-SLN service automatic navigation.

V. Conclusion

In this paper we proposed and explored a semantic based graph mining methodology for S-SLN service automatic navigation. The main motivation is to improve or speed up user's navigation decisions by providing the most related Web service and links which are most likely to be of interest.

We have presented a simple definition of Semantic Neighborhood of the source service nodes and discussed an approach of service automatic recommendation that are used to reduce user's cognitive load. The methodology revolves around a semantic based activation function technology and map related service nodes recommendation problem as a basic graph mining problem. The work presented here demonstrates the potential of simple graph mining method that automatic navigating S-SLN based on Spreading Activation technology. While because of the complex interrelationships between Web services, the structure of S-SLN will dynamically vary from time to time. Automatic service navigation and recommendation should be adapted to the dynamically S-SLN. This issue is at least worth investigating and will be part of our future work.

Acknowledgment

This work was partially supported by the Scientific Research Foundation for the Doctor Scholar of Chongqing Normal University and the Fundamental Research Funds for the Central University under Grant Project No.CDJRC10010010.

The author thanks all team members of Semantic Grid Research Group of Southwest University for their help and cooperation.

References

[1] Alexander Troussov, Eugene Levner, Cristian Bogdan, John Judge, Dmitri Botvich, "Spread of activation methods", in *Dynamic and Advanced Data Mining for Progressing Technological Development*, IGI Global, 2009, pp.136-167.

[2] Anping Zhao, Zhixing Huang, Yuhui Qiu, " Service semantic link network discovery based on markov structure," in *Proc. 6th International Conference on Semantics, Knowledge and Grid*, 2010, pp.9-16.

[3] Agrawal, R., Srikant, R., "Fast algorithms for mining association rules", *In International Conference on Very Large Data Bases*. Morgan Kaufmann, San Francisco, CA, 1994.

[4] Travers, J. and Milgram, S. , "An experimental study of the small world problem", *Sociometry* ,vol32(4), pp.425–443, 1969.

[5] D. Chakrabarti and C. Faloutsos , "Graphs mining: laws, generators, and algorithms", *ACM Computing Surveys*, Vol. 38, pp. 1-69, March 2006.

[6] Tim Hussein and Sebastian Neuhaus, "Explanation of spreading activation based recommendations," *SEMAIS 2010*, 2010, pp.1-5.

[7] Alexander Troussov, Mikhail Sogrin, John Judge, Dmitri Botvich, "Mining socio-semantic networks using spreading activation technique," in *Proc. International Workshop on Knowledge Acquisition from the Social Web KASW08*, 2008, pp. 405-412.

[8] Troussov, A., et al. "Navigating networked data using polycentric fuzzy queries and the pile UI metaphor navigation", *Proceedings of the International SoNet Workshop*, 2008, pp. 5-12.

[9] R. Renuga et al, "Semantic search using content based search technique with spread activation technique", *International Journal of Recent Trends in Engineering*, Vol. 1, No. 2, pp.226-229, May 2009.

[10] Alexander Troussov, Denis Parra, Peter Brusilovsky, "Spreading activation approach to tag-aware recommenders: modeling similarity on multidimensional networks", *Proceedings of Workshop on Recommender Systems and the Social Web at the 2009 ACM conference on Recommender systems, RecSys '09*, New York, NY, October 25, 2009, pp.1-6.

[11] Sheila Kinsella et al, Navigating and annotating semantically-enabled networks of people and associated objects, In *Why Context Matters: Applications of Social Network Analysis* (T. Friemel, ed.), VS Verlag, 2008, pp. 79-96.

[12] Salton.G and Buckley.C, "Term-weighting approaches in automatic text retrieval," *Information processing & Managament*, vol.24(5), pp.513-523, 1988.

[13] EEE-05 Challenge Samples Repository (2005), http://cssun.georgetown.edu/~blakeb/EEE05/WSDL_Repos.zip

[14] Michael R. Berthold et al, "Pure spreading activation is pointless", *CIKM'09*, November 2–6, Hong Kong, China, 2009, pp.1915-1919.
[15] Sitaram Asur, "A viewpoint-based approach for interaction graph analysis", *KDD'09*, June 28–July 1, Paris, France, 2009, pp. 79-88.

Anping Zhao graduated from Southwest University, China and received his Doctor of Artificial Intelligence from the Department of Computer and Information Science in 2011. Now, he is working in the College of Computer and Information Science, Chongqing Normal University, China. His main research interests include Artificial Intelligence, Semantic grid, Cloud computing and Software Engineering.

Yu Yu received her Doctor of Logics from the Southwest University in 2010. She is a lecturer in the College of Trade and Public Administration, Chongqing University, China. Her main research interests include Artificial Intelligence, Semantics and Language logic.