Clustering of Semantic Web to Develop Mashup of Web Services
Shruti Gupta¹ Salunke Avinash Nimba² Jaishankar N.³
¹²³ M.Tech ³Professor
1,2,³Department of Computer Science & Engineering
1,2,³VIT University Vellore, Tamil Nadu, India

Abstract—Semantic web refer to the web of data which construct meaningful data from unstructured and semi-structure data architecture. It allows data to be share and reuse across different platforms, frameworks and countries. So the clustering of the semantic web of data will provide effective solution for web of data access. And in today’s web, there are different web services which is develop across different frameworks and protocols. So this paper combines all the web services which are develop across different framework and protocols in one service. This represents the mashup of web services, it has used the content which is from more than one source and creates into single new service displayed in single GUI. So doing the clustering of semantic web of data will enhance the effectiveness of mashup of web services.

Key words: Semantic Web, Clustering, Mashup, Software Development Life Cycle Model

I. INTRODUCTION

As in today’s scenario web is moving from web 2.0 to web 3.0 with respect to the need of enterprise solutions. So there is drastic need of the technology which can use the previously defined methods and web services to provide the same functionality and information across the internet. This leads to combine the different-different web services in one web service with respect to the same type of information or data types, which can be reduce the overload of unwanted data information and forms the semantic nature of the combine web services. To develop this type of technology, basically it requires identification of the web of data and the services with the same contents which leads to clustering of this data on the basic of similarity. Clustering[6] refers to the grouping of similar data set into the one group selected from the different and large data sets. Hence the first part of this technology is to cluster the meaningful data which can be provided by the previously defined web services methods and after clustering all data will be efficiently utilized by the proposed web services mashup[6]. Here the cluster data is of the type semantic web. Clustering follows some similarity approaches based on taxonomy, relation and attribute similarity[6]. Here the cluster data is of the type semantic web. Clustering follows some similarity approaches based on taxonomy, relation and attribute similarity[6]. Clustering is used due to its-Simplifications, Pattern detection, Useful in data concept construction etc. Clustering is used in various areas such as Data mining, Information retrieval, text mining ,Web analysis ,marketing ,medical diagnostic and many other areas.

Major Existing clustering methods are-Distance-based, Hierarchical , Partitioning, Probabilistic. In distance-based clustering case we easily identify the clusters into which the data can be divided; the similarity criterion is distance: two or more objects belong to the same cluster if they are “close” according to a given distance. Hierarchical clustering is of two types-Agglomerative (i.e. bottom up) and Divisive (i.e. top down). In Agglomerative, it starts with a single step then recursively add two or more appropriate clusters and finally Stop when k number of clusters is achieved. In Divisive start with a big cluster than recursively divide into smaller clusters and finally stop when k number of clusters is achieved. Partitioning clustering begin with division of data into proper subset than recursively go through each subset and relocate points between clusters (opposite to visit-once approach in Hierarchical approach).This recursive relocation generates a higher quality cluster .In Probabilistic clustering data is picked from mixture of probability distribution than use the mean, variance of each distribution as parameters for cluster.

As user may search for any key on the internet and for replying the query which encounter by the user there must the same scenario which fulfill this. But in reality along with the result of query, web of data provides some large number of irrelevant return services which may not be useful for user so this simply leads to overloading the traffic and leads to bad services formation for users. The semantic web [7] provides the meaningful data which semantically and optimally perfect for the given query. Mashup is actually refer for audio music in which different songs are combine in one genre as one new song from existing songs, this song is refer to as mashup [1]. Hence mashup can be say as something which can be combined into one for creating something new. Here the different web services providing from enterprise are combined into one to create new web services which can be displayed in one GUI [1]. This mashup is implies to fast integration, easy implementation, frequently using open API and semantic web of data to enriched the web of services [13]. Combination of data, visualization of data and aggregation of involved web services are main characteristic of mashup services [6]. It makes the existing data more accurate, more useful and meaningful with respect to user for the personal or professional use. Mashup allows to web application to integrate data and functions in one API instead of building all are in separate web service methods. Mashup technology hardly require the programming skill cause it supports GUI based API’s which are very simple for understanding and using for end user [13]. Mashup having following types as

- Business Mashup[12]: It is also refer to as enterprise mashup which combines the resources, data of applications and other web services methods. This mashup is perfectly suitable for the Agile Development project, this gives the collaboration between customer specially a product manager and the developers for satisfying and invoking business requirements. This type of mashup is very secure and Rich Web Application which contains all external and internal information which fulfill the requirements.
- Consumer Mashup [12]: This mashup implies for the combination of multiple data from the public sources web sites and gives well defined user interface through browsers.
- Data mashup [12]: This is just opposite to above one where similar information and utility media are combine into single representation from multiple sources. By this type of mashup a new web service is created which is distinct from the originally existed.

II. RELATED WORKS

- Instance Based Clustering of semantic web of data [7]: This paper explain the challenges and issues surrounding to the application of a clustering algorithm to the semantic web of data. For extracting the instance it adopted RDF conversion from databases, another is based on the Ontology of Semantically Rich Web Enable Application. As in traditionally clustering uses two methods instance extraction and distance measure, also it requires the number of cluster of the data set. The cluster the semantic data which can be important tool for the autonomous semantic agent. Discuss challenges like how to extract the instance representation from RDF graph and after extraction these instance how the computer can distance the instance of two such instances. This paper uses algorithm as Hierarchical Agglomerative Clustering (HAC) algorithm.
- Clustering Ontology-based Metadata in the Semantic Web [6]: This paper proposed approach Ontology Based Clustering of Metadata for mining the semantic web of data. The ontology implies a formal, clear and conceptualization of the describe domain interest. It defines the ontology structure into 6 tuple set which forms the relation with other data set of same domain. Also it explained the structure of metadata which follows the given ontology. It made computational similarity approaches.
- Mashup Of Web services [1]: This defines the overall procedure for developing the mashup. Explains the what are the main advantage and important features of it. Paper demonstrates simple mashup which can make the web services more effective.
- Smart Travel Planner AMashup of travel-related web service [2]: This paper proposed the mashup of traveling related information through the several API’s. This showed that in one service all the information related with traveling such as route map, hotels, food and shopping details in single API.
- Using Logically Hierarchical Meta Web Services to Support Accountability in Mashup Services [3]: Mashup integrate different API into single one and form new GUI for users which provides flexibility for users. This paper proposed the new model to resolve the problem relating with accountability. It uses PKI and hierarchical meta web services for supporting traceability and self identification of web mashup.
- An aggregation search engine based On RESTful Web services and Mashup [4]: As this paper proposed the RESTful service to form Mashup by introducing Aggregation serach engine (ASE). This also showed the Google AJAX and Microsoft Bing and focus on scalability, reusability and flexibility.
- Social-Based Web Service Discovery and Composition for Step-by-step Mashup Completion [5]: This paper describe the services composition for mashup environment by explaining the RDF graphs for it. It explain the detail methods for discovering the semantically related web service methods and form mashup of it.

III. PROPOSED ARCHITECTURE

The proposed architecture consists of three phases

A. Clustering

For developing this service first it need to cluster the semantic web of data into similar cluster so that irrelevant data will be avoided while replying the query. We here proposed Clustering Probabilistic Semantic Approach which is efficient for finding Web Services whose content is not compatible with the fired query by the user. The irrelevant web services should be eliminated within the collection of web services.

![Fig. 1: Clustering Method](image)

To finding the irrelevant web service we group the same kind of web service along with its query. Let q be the query and s be the service then similarity computation can be done as:

$$\text{Sim}(q,s) = \frac{|q \cap s|}{q + s - |q \cap s|}$$

From the above similarity computation we can acquire an initial set of samples through selecting a predefined threshold and the web service clustering can be managed along with given fig.

As here consider that the initial service may content some irrelevant services hence to remove this follow the procedure:

Given w returned services

$$S = \{s_1, s_2, ..., s_n\}$$ with respect to a query, cluster S to k groups

$$C = \{c_1, c_2, ..., c_k\}$$ and remove service $$s_i$$ such that

$$|S_i \cap C_j| \geq E,$$

$$K < w$$

$$J \in (1, 2, ..., k)$$

All rights reserved by www.ijsrd.com 612
Where $\varepsilon$ is predefined threshold and $C_j$ is the center of cluster.

Hence by implementing the above procedure we can find out the irrelevant services which can help in forming the Mashup of web services.

The mashup of web services is occurred. This follows the architecture as form semantic web of data, clustering this semantic web of data and then forming the mashup into one service. The architecture combines all the web services methods into one. As in real time according to the need of enterprise the web services are developed by using SOAP, REST, XML, JSON, RSS, ATOM etc [4]. The overall architecture is given in figure 1 where the globe is a SOA including various web services protocols which make the semantically rich web services.

![Fig. 1: Mashup of web services](image)

These are all different web protocols which can be used according the model followed by enterprise.

Three layer architecture of mashup

1. Presentation/user interaction [12]: This layer defines the user interface for mashup, it uses various techniques such as HTML, CSS, Javascript, AJAX.

2. Web Services [12]: To achieve the product main functionality it provides API of services. By using API the newly form web service can be accessible through the use of technologies like XML, HTTPRequest, XML-RPC, JSON-RPC, SOAP, REST.

3. Data [12]: As the data is main source and challenge for developing web services so this layer describe how data is carried along the operation. It uses the technologies such as XML, JSON etc.

Architecturally mashup is divided into two styles as for server side mashup and client side mashup [13]. These are main type of mashup from which server side mashup provides mashup of data across databases from different vendors and also combines software’s [9]. On server side all the above protocols are integrated to one with the help of programming language and deployed on browser (AJAX Application) [10]. And in other hand client side mashup access the service of mashup using Ajax Client Mashup Browser [9]. This client side data mashup implies to taking the remote information from web services and combines it with data from another source, this can be generate new information which may doesn’t exist before the result is generated [10].

Client Side Software mashup implies the code is directly integrated from various methods in the browser to form new result to show different capabilities of existed code. This provides great flexibility to use old method’s code in one integrated code to form new code without modifying the existing code. Presentation Mashup implies the presentation of the combined web of data into some well organized form by using AJAX service.

IV. METHODOLOGY

- Semantic web of data collection [12]: The semantic web refers to build new architecture for www web for enhancing the formal semantically content [6]. For collecting and identifying the semantic data for web it must be consider by using the define ontologies. Ontology elaborates the engineering data structure, similar data set, relation between data sets which refers to the clear conceptualization of the defined data set of common domain interest [6]. Based on the given relations of ontology it is very easy task for indentifying the semantic web of data which will be perfectly suitable for the given user queries. For gathering semantic web of data, it uses taxonomy similarity of ontology which computing the similarity on the basic of concepts and position of two instances. Relation similarity determines by the relations of corresponding objects. Attribute similarity evaluated by the respective attributes and their values.

- Clustering of semantic web of data [6]: As the collected semantic web of data must be cluster according to the some specific groups based on the above discuss similarity criteria’s. For clustering there are various algorithms are present such as k-means algo, hierarchical clustering algo, self-organizing map algo, expectation maximization clustering algo and based on the performance analysis here instance based clustering is used which cluster the semantic web of data with respect to instances. It follows the instance extraction and distance measure which gives the RDF graph specification. For instance extraction it uses same similarity computation of above semantic and for distance measure it follows Vector Based Distance Measures. There are also various approaches for distance measures such as Graph Based Distance Measures, Ontologically Based Distance Measures. Also other approaches for clustering are as PAM and SOM. PAM is partition around medoids which finds the sequence of objects called medoids which are centrally located in clusters. The main goal of this PAM algorithm is minimize the average dissimilarity of objects to their closest selected objects. Hence clustering of semantic data is becomes very important requirement for replaying the web services along with the correct meaningful data by avoiding the irrelevant data sets.

- Mashup Development [13]: For developing the mashup of web services it should be follow the following steps:

1. Identify the requirement for your mashup.
2. Collect and cluster the semantically web of data into similar groups and make data collaborations.
(3) Choose the development platform
(4) Design and select good mashup editor, as this editor will provide user interface.
(5) Sign up for the services which can be accessible from API.
(6) Integrate the selected web services using API into one and this will be the final step for mashup.

V. RESULT DISCUSSION

As Google created the Map Mashup which are using for the determine desired the locations and mashup with GPS and other highlighted feature or the daily life events such as rating for coffee shops, book shops etc.

![Google Mashup](image)

**Fig. 2: Google Mashup**

It also provides pop up details about the specific locations and this Google map can be again mashup into desired website as user can get view on one page about the navigation. Other example of mashup is combining SMS services in website portals which gives the current process state by alerting him through the SMS. Also weather forecasting API gives the best mashup which can be embedded in site for various purposes. As mashup, gives better user interface with different-different application in one interface as shown in figure 3.

![Google Gadgets](image)

**Fig. 3: Google Gadgets**

As doing this, various gadgets into one page can achieved the better reuse capability of the web services. It increases the faster time for marketing in one cloud environment. As there are so many mashup tools are present such as Google Mashup Editor, Microsoft Popfly, Yahoo Pipes.

VI. CONCLUSION

By mashup technology reuse of various web services methods is possible under the single web services user interface. The technology provides API which can be used for developing the new mashup in desired ways hence instead of coding same functionality which already exits, mashup embed these functionality in single source. It generates new information which may not be existed as it combines the semantic web of data gather from large set of data domains.

REFERENCES

[7] Gunnar AAsstrandGrimnes, Peter Edwards, AlunPreece, “Instance Based Clustering of Semantic Web Resources”, Kaiserslautern, Germany,Kaiserslautern, Germany
[12] [http://semanticweb.org/wiki/Main_Page](http://semanticweb.org/wiki/Main_Page) and [http://www.w3.org/standards/semanticweb/](http://www.w3.org/standards/semanticweb/)