The survey: Integrated Mashup Service Using Data Mining Logic

[1] Ms. Priti V. Rane [2] Prof. Umesh Kulkarni
[1] PG Scholar,Alamuri Ratnamala Institute of Engineering and Technology,Shahapur
[2] Assistant Professor, Department of Computer Engineering,
Vidyalankar Institute of Technology, Mumbai

Abstract: -- Web mashup is an activity to integrate various types of E-commerce services. With the advent of Web 2.0there has been a substantial increase in development and usage of E-commerce websites.The web pages that are open do not have an efficient data mining technology for searching, robust & reliable infrastructure.This paper presents efficient technique to cluster data from web repository based on user behavior analysis using levenshtein distance algorithm which gives the structural result.

*Keywords:-*Clustering, Data Mining, Web Mashup

I. INTRODUCTION

The World Wide Web popularity leads to a revolution towards electronic data. An integrated mashup service is a web application that takes information from one or more sources. It is necessary to forecast mandatory data before executing service integration process. No of web mining algorithm are available predict required data like page rank, weighted page rank, HITS etc.

Knowledge Discovery in Databases (KDD) is initial idea from which mashup candidates are identified by addressing a repository of open services. Within this approach, there was a customized improvement life cycle that software engineers can use to generate new applications based on service integration techniques. KDS also define service integration qualification by discovering different aspects of web service specifications[1]. Integrating Information, Simply states that majority of work addressed the tools and environment that used to accomplish mashups techniques. This tool supports the visualization and presentation of mashupresults [2][3].

For modernizing systems using the information resulting from successful mashups various policies was investigated. These policies also protect data in mashup environment [4].

In web mining framework four phases was developed' web structure mining analysis, Web Content Mining analysis, decision analysis and security analysis. In web structure mining analysis page rank algorithm and trust rank algorithm was used. In Web Content Mining analysis Hierarchical agglomerative clustering and k-means cluster analysis was used. In decision analysis trust calculation of web site and statistical techniques ware used to analyses the result of the evaluation[5]. This paper is organized in the sections. Section 2 describes theproposed work for Integrated Mashup Service, section 3 describes research methodology. The result is presented in section 4 followed by references.

II. INTEGRATED MASHUP SERVICE USING DATA MINING LOGIC

The emerging phenomenon of web 2.0 describes the new characteristics of web. It demonstrates that the end user have keen interest in developing services through different static services available on web. The customizable web portals are easy to use but they do not support advanced application because the software services and data repositories cannot be combined. This process is time consuming and becomes honeypots for intruders. In proposed system, data is extracted from web repositories based on user requirement. User can search data with security features. User need not to fill their personal information. User can just type combination of queries and based on user behavior analysis data will be predict. Using levenshtein

distance algorithm data will predict andclusters are created using clustering algorithm, this exact data will display to user.

III. RESEARCH METHODOLOGY

User interface is composed of main query and sub queries. As per query, proposed work searches in web service repositories for viable web service candidate by extraction of parts and other descriptive service-based information. This can be a service name, operation name, type name and descriptive fields from the service specification. Clustering is done on the basis of user behavior. Once equivalences are identified between input query and specific parts of web services, web services are linked together by input and output messages. Fig.1 represents System architecture of proposed system.
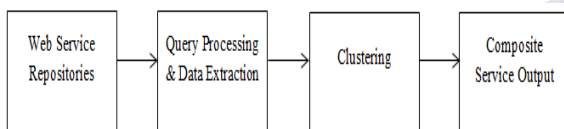

Fig.1: System architecture of proposed system

a. Web Service Repositories

Web service repositories Universal Description Discovery and Integration (UDDI) contain viable candidate of web services from which data is searched.

b. Query processing & Data Extraction

Query processer contains combination of main queries and sub query based on which data is extracted from web service repositories.Levenshtein distance algorithm is use to process user query. Data is refined based on user input.

c. Clustering

Using Levenshtein distance algorithm the closest values of an instance are determined. Categorization is done by "Categorization On Pairing" (COP). This approach clusters services into categories based on the similarity of their specifications. Each service is compared to every other service in the repository once. The algorithm starts at the pair of services with the greatest similarity and makes a category for these two services.

d. Composite Service Output

Filtering is done by "Prediction Algorithm". This approach takes data from clustering. A prediction algorithm result has added value to particular user. At the end composite service is display to the user.

IV.RESULT

Equivalence processing module is implemented by using Levenshtein distance algorithm. Input Query and Service Specifications of web services are matched. It creates cluster for further processing, which is continuously refine to give structured data. Using web 2.0 user can perform various activities through e-commerce. The proposed system integrates various web services. It helps user to predict exact data from web repository and to preserve security in E-commerce website.

REFERENCES

[1] M. Brian Blake, Michael F. Nowlan, "Knowledge Discovery in Services (KDS): Aggregating Software Services to Discover Enterprise Mashups," Knowledge and Data Engineering, IEEE Transactions on, vol.23, no.6, pp.889-901, June 2011.

[2] M.B. Blake, "Knowledge Discovery in Services," IEEE Internet Computing, vol. 13, no. 2, pp. 88-91, Mar. 2009.

[3] M.B. Blake and M.F. Nowlan, "Taming Web Services from the Wild," IEEE Internet Computing, vol. 12, no. 5, pp. 62-69, Sept./Oct.2008.

[4] Cristina Ofelia Stanci "K- NearestNeighboralgorithm for instance based learning", 2012.

[5] R.Manjusha, Dr.R.Ramachandran, "Web Mining Framework for Security in E-commerce", IEEE-ICRTIT 2011.

[6] Ajaya K. Akasapu1, P. SrinivasaRao, L. K. Sharma and S. K. Satpathy, "Density Based k-Nearest Neighbors Clustering Algorithm for Trajectory Data", International Journal of Advanced Science and Technology Vol. 31, June 2011.

[7] "Data Mining Introductory and Advanced Topics" by Margaret H. Dunham, Pearson Education.

[8] Mrs.PallaviChandratre, Prof. UmeshKulkarni, "Security Issues Related to Web Services In E-commerce" 978-1-4799-7319-4/14/©2014 IEEE.

[9] PraveenkumarArjun Patel, Prof. UmeshLaxmanKulkarni "Different Syntactic Methods and Clustering For Web Service Integration"International Journal of Innovative Research in Computer and Communication Engineering Vol. 3,pp. 357-362 January 2015.