Aspect based Enhanced Web Service Discovery

Vandan Tewari[*], Nirmal Dagdee[**], Aruna Tiwari[***]

[*] C.T.A. Dept., Shri G.S.Institute of Tech. & Sc. Indore
[**]Computer Engg. Dept, S.D. Bansal College of Tech.,Indore
[***]Comp.Sc. & Engg.,IIT,Indore

Abstract- Web services [1] are the latest paradigm of distributed systems which are implemented as typically application programming interfaces (API) or web APIs that are accessed via HTTP and executed on a remote system hosting the requested services. The Web services are used to interact between components, programs, and applications. In this paper, we have proposed an approach to allow for support of web service descriptions and queries using registries that conforms to UDDI V3.0.2 specification. Specifically, our main aim is to discover user requested services based on not only on functional parameters but also non functional parameters such as Quality of Service (QoS) parameters. Web Service discovery can be augmented using query expansion as well as appended with QoS in requirement. Since the metadata associated with web services is very less as compared to HTML pages, to offer a web service appropriate to a client we have used two techniques from Information Retrieval: Query Expansion and Query Rewriting. The proposed approach does not require any modification to the existing UDDI data structures. The add-on modules for searching reside at the service requestor's end who wishes to take advantage of better web service discovery. This approach is completely backward compatible and can integrate seamlessly into existing infrastructure. An e-learning technique has been used to demonstrate the effectiveness of proposed approach.

Index Terms- Web Services, Discovery, Query Expansion, Quality of Service (QoS)

I. INTRODUCTION

The service oriented architecture SOA is the latest and most popular realization for web services. Web Services are building blocks of SOA and makes the e-business processes interoperable and loosely coupled. These characteristics makes web services very popular. The basic SOA contains three entities, Service Providers which provide or deploy the services to be used by Service Requestors and Service Brokers which acts as directory services for discovering web services. The Service Broker can be realized as a UDDI Registry. The following is an explanation of underlying technology for deploying , discovering and invoking web services. An E-Learning domain has been chosen to presented for the proposal and therefore an elementary discussion on E-Learning System has also been presented here.

Universal Description, Discovery, and Integration **(UDDI)**[2] The Universal Description, Discovery, and Integration UDDI is a realization of service broker which provides a standard method for publishing and discovering web services. The UDDI is an industry standard that has been developed as a platform-independent, open framework for describing services, discovering businesses, and integrating business services. UDDI is a de facto standard for the process of discovery in the service-oriented architecture. The basic information model and interaction framework of UDDI registries consists of the following data structures[2]:

businessEntity: Contains information about the business or organization that is offering the published service.

businessService: Contains information about the individual service offered by a business.

BindingTemplate: Has information about technical details of as ervice and a reference to the service's programmatic interface or API.

tModel : Additional information such as categorization metadata, taxonomies and policies, are stored in tModels

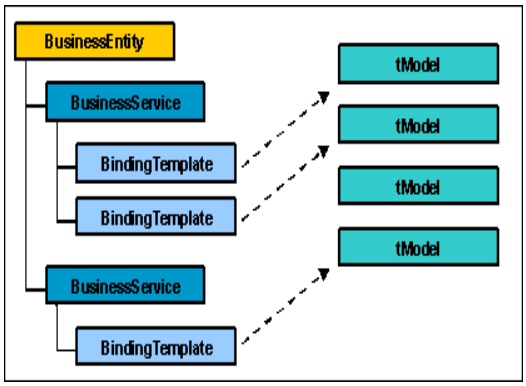

Fig 1 Basic Data structure of UDDI

Extensible Markup Language (XML)

XML[3] is widely known paradigm for exchanging information over the web and follows the semi structured model for a rapidly growing number of software development activities. XML is intended as a markup-language for arbitrary document structure, as opposed to HTML, which is a markup language for a specific kind of hypertext documents.

E-Learning Systems: E-Learning [4] is the use of Technology to enable people to learn ubiquitously. E- Learning can include training, the delivery of learning content management and information and guidance from experts.

Publishing Services for an E-Learning system

A Service Provider starts by modeling the real-world concepts associated with the service and creates a WSDL[5] Documentto define the Service's syntax by specifying the structure of the input and output messages, along with aspects of the service's runtime bindings.It then mapselements of WSDL Messages to

Semantic Concepts. The next step is expressing the meaning of the WSDL by mapping the operations and the schemas of the WSDL input and output messages to the semantic model.It then publishes the Model, WSDL, and Mappings in a UDDI Registry. The concepts in the WSDL, the semantic model, and the mappings between them are defined using tModels in the UDDI registry.

Searching an E-Learning system

Identify the Required Service using requested string.Query UDDI, Locate Service, & Retrieve all Interface Descriptions.Access the Web Service.

Using QoS Context for web services discovery

Current keyword based searches for web services in the UDDI Registries do not provide effective response given the small amount of information about services in UDDI registries. Due to inherent ambiguity and incompleteness of web service requests, it is difficult to maintain the relevancy in results. Interestingly this problem can be mitigated if we consider the context of a request while searching for services. This aspect based search refers to automatically augmenting the user request with information extracted from search context. This context can be generated if we add additional information to the query entered by the user. Context can also be defined as the quality of service parameters for a user.

II. RELATED WORK

Discovering web services using current search techniques offered by existing UDDI APIs may not give results which fulfills the need of a service requestor's. When discovering web services, clients search for the services that meet their functional needs as well as quality of service (QoS)[6]. Authors here stresses upon that it is very important to consider QoS parameters while discovering a service. Extending search using existing search API [7] only exploits keyword based search techniques which may not be suitable for web services when we want to select services between services which have similar functionalities. Differentiating between web services that share similar functionalities can be effectively done by observing non functional web service attributes such as throughput,response time,cost , availability etc.

Authors propose to use a third party agent to certify the QoS parameters proposed by the service provider in [8]. E. El- Masri [9] in his paper stresses upon the need of discovering best services for a client based on non functional parameters of a service .Xu et al. [10] in his research work proposes to use a web service discovery model that contains an extended UDDI to accommodate the QoS information, a reputation management system to build and maintain service reputations and a discovery agent to facilitate service discovery. Authors have also developed service matching, ranking and selection algorithm, but have not not provided any process of verification process for QoS in that proposed discovery model. R.Kraft et. al. [11] have proposed three different algorithms for implementing contextual search of web pages. However the algorithm of query rewriting may be used while discovering appropriate web services.

III. PROPOSED APPROACH

We have developed a web service discovery framework that aims to provide best E-learning services based on non functional Quality of Service parameters so that the services can be searched not only on the basis of user defined keywords but also non functional quality of service parameters so as to get better results. Also we have developed multiple E-Learning web services so to as to enable the search of services from multiple domains classified based on their properties. To achieve these objectives we have created a module for:

- Deploying e-learning services.
- Publishing of services by the service provider.
- Enabled aspect based discovery of web services by the client which includes Query expansion on keywords as well as QoS Parameters.

Architecture design has been shown in Fig. 2 given below which shows the system not only uses functional but also non functional parameters to search for appropriate service for a client.

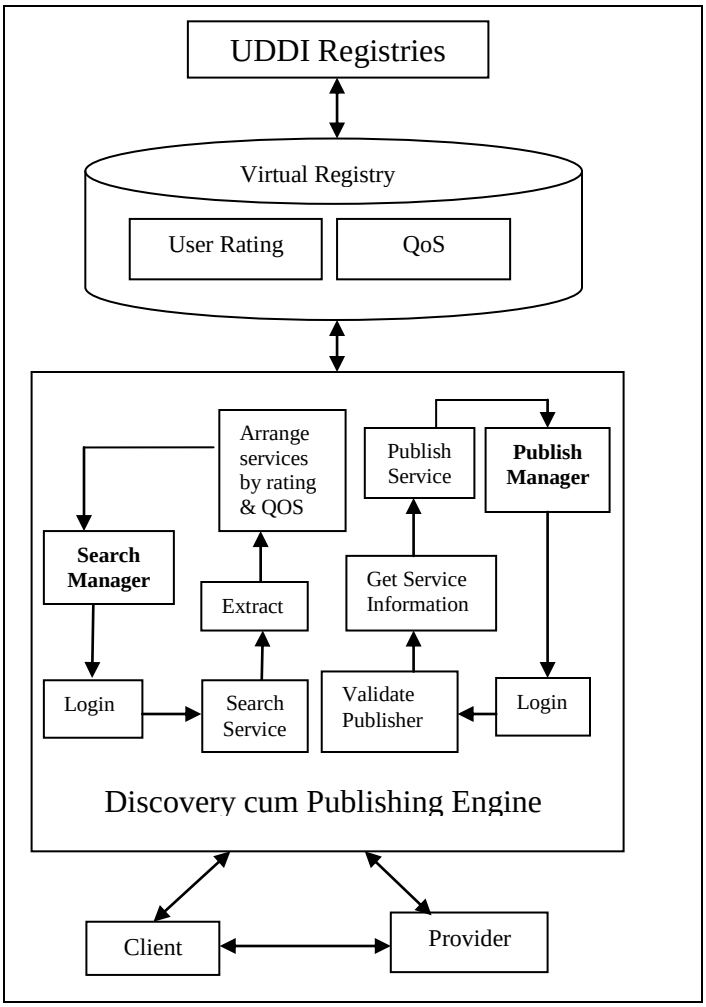

Fig. 2 Architectural Design of Proposed Technique

Proposed Architecture of the System:

The proposed architecture uses anenhanced web service crawler engine [14] specially designed for crawling web services and

implements an enhanced keyword based discovery as well as incorporated searching of web services on non functional parameters. The discussion related to QoS publishing and discovery has been discussed by us in [8]. The Virtual Registry is a repository and has various functions as:

- It stores and indexes the web services fetched from multiple UBRs under a keyword based categorisation or domain specific classification scheme and hence is the single point for web service discovery.
- It stores the addresses of all UBRs or seed points of all web services sources for crawling.
- It is also used as a **Web Service Cache** for search results of similar queries which have been fired previously. The policy of caching is based on how popular the query has been for the past three days of usage.
- It stores the service rating data and associates it with the service metadata for ranking of services based on user rating as well as popularity.
- It gets populated by the crawler which runs as a background process periodically using focused crawling.

To implement the various modules the techniques used are discussed in subsequent sections.

A. Query Expansion

The keyword based search in current UDDI Registries does not result in appropriate services for a client. This search can be improved if the keyword based query is augmented with additional information. A way to improving web service discovery is using methods of information retrieval used by search engines. Reiner Craft et al. [11] have suggested some techniques for context for contextual search. They have discussed a technique known as Query Rewriting which refers to augmenting the query with appropriate terms from search context such as QoS and then using a search engine for improving the query. In another proposal [13] P. A. Chirita et al. have suggested to devise a personal information repository for a client and using it to expand a query. However we suggest another method for query expansion for keyword based web services discovery, looking to the minimal information available in web services descriptions. In the proposed approach at first the query is run by the crawler [14] but these results are not presented to the user but are used to expand the query. The query is expanded using keywords extracted from top ten of the results.

An Example Scenario: A user is trying to locate an e- learning service through proposed architecture. He enters a query : 'java network programming '

1. The engine first extracts keywords from user's query(Parsing +stemming+ removing stop words to produce a vector of keywords e.g in this case [java, network, programming]) on WSDLs on Internet. Say it fetches Six services either containing one two or all three keywords from query.
2. Now from all the fetched services ,Keywords are extracted and a matrix is devised as:

Services	S1	S2	S3	S4	S5	S6	P(Keyword)
Keyword1 (Java)	1	0	0	1	1	1	4/6=0.66
Keyword2 (network)	1	1	0	1	0	0	3/6=0.50
Keyword3 (Book)	0	1	0	1	0	0	2/6= .33
Keyword4 (Programming)	0	1	0	0	0	0	1/6=0.16
Keyword5 (learn)	0	1	1	0	0	0	2/6= .33
Keyword6 (ebook)	1	0	1	0	1	1	4/6= .66
Keyword7 (socket)	1	0	0	0	0	0	1/6= .16

Here Keywords are all the keywords collected from those six services. A '1' signifies the presence of a keyword in a WS Description.

The probability of topmost Keywords are Keyword1 (Java=0.66) and Keyword6 (ebook=0.66), Keyword2(network=0.50), Keyword3(Book=0.33) and Keyword5(learn=0.33) ; after deleting the keywords in original query considering top two of rest ,the query Vector is now expanded as : [java, network, programming, ebook, learn]

The expanded query is now forwarded to the Discovery engine which results in a list of services which is better suited to a user. The results have demonstrated the effectiveness of such an approach.

Enhanced Keyword based Service Discovery Algorithm using Query Expansion

FOR each user query Do{
1.Parse Query.
2. Run query and collect top ten results
2.1. For the top ten results do {Parse the service descriptions to collect keywords}
2.2. Calculate the Probability of occurrence for all keywords.
2.3. Choose the Top ten keywords based on highest probability of occurrence to form a term expansion set.
2.4. Remove keywords of user's query from the term expansion set .
2.5. Select the five (at most) top ranked keywords to expand the query.
3. Rerun the query using the expanded query
4. Return the results.

B. Query Rewriting

Using the additional information entered by the user say {Cost C, Response time R, Accuracy A, security S} where each term has been normalized between (0,1),the query vector can be thought of as :

Q = [q , CV] where q is the query entered by the user whereas CV represents the Context vector which is defined as below: CV={a*w_1,b*w_2,c*w_3,d*w_4} where w_1, w_2 represent the weights

associated with various attributes such that ∑w= 1 and the context vector is written as per increasing weights as entered by the user . Let's say CV={A,B,C,D}

Theoretically Q is either of the values {q , q+A , q+A+B, q+A+B+C,…} The higher terms of this expression obviously suffers from low recall as more specialization takes place.

For example Say user entered the following values (cost 0.4 with weight 0.5, response time 0.4 with weight 0.3, Accuracy 0.1 with weight 0.2 and security 0.1 with weight 0.1) Then using the normalized values of all attribute

Required Overall value of QoS Context = 0.2+0.12+0.02+.01=0.35 so the topmost results fulfilling the functional requirement should have at least QoS Context value >= 0.35.

The publication of QoS for web services and calculation of overall QoS using associated weights has been discussed by us in [8] & [12].

IV. TEST CASES

1. QoS based Search:If User wants to search based on context then user has to select QoS search option, Search results based on QoS have been shown following snapshot.

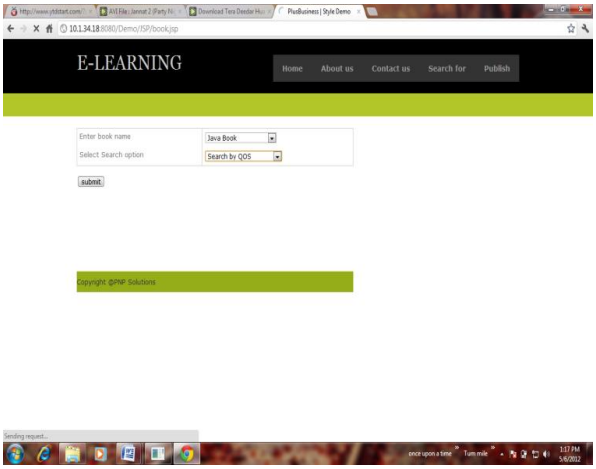

Fig 3 This search facilities the user to search by QOS

2. These are the results as per the given Query for a java e book in step1.

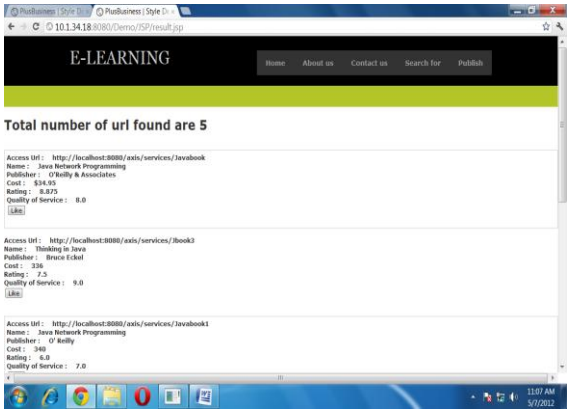

Fig 4Results for QoS based search

2.User rating based simple search:

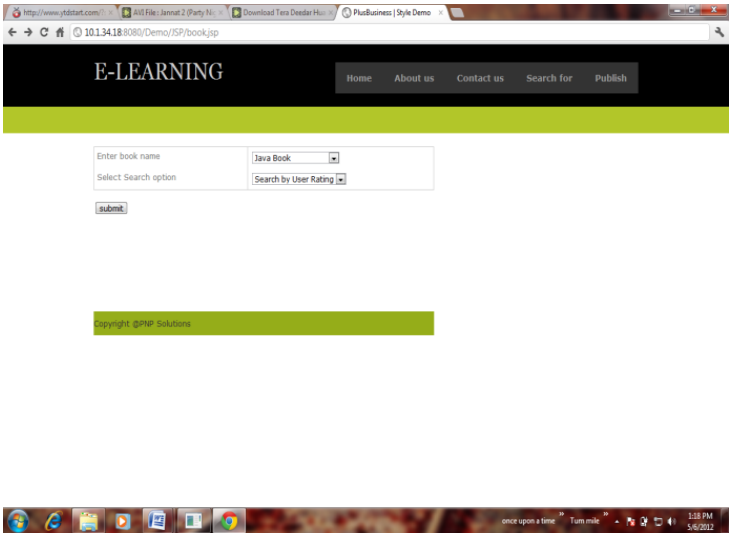

Fig 5 Simple search based on user rating

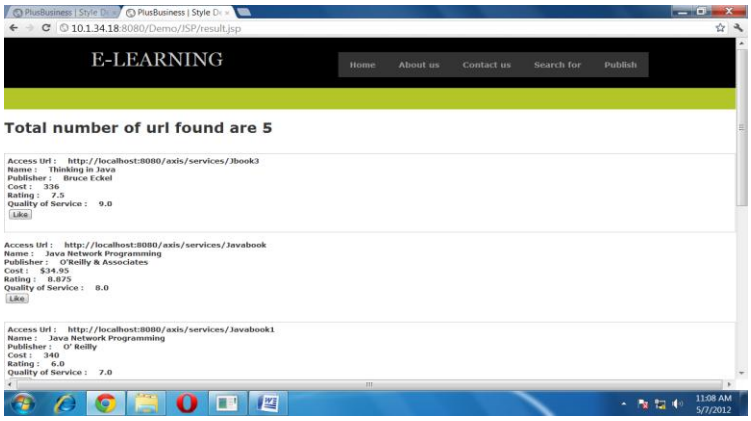

Fig 6Results of a simple search with User rating

There is a clear distinction between the search results obtained from a QoS context based search and the simple search based on User rating. In addition the improvement in precision can be seen from the Figure 6 when the keyword enhanced algorithm was used. Precision represents the

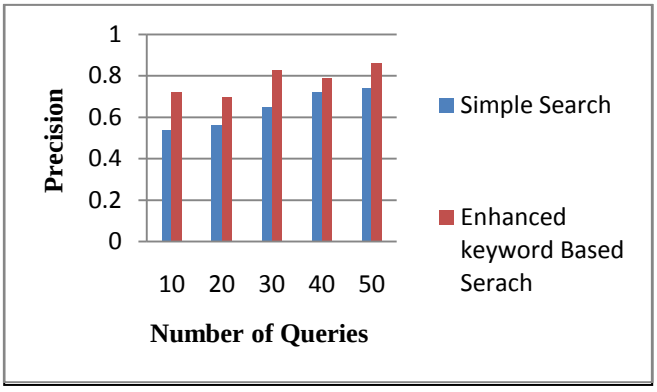

Fig. 7 Precision Curve with Enhanced Search

V. DISCUSSION AND FUTURE ENHANCEMENT

The web service discovery has been improved by the augmentation of intelligence based search algorithm which incorporates both functional as well as non functional parameters. In addition an enhanced keyword based algorithm has been proposed which improves the service discovery on functional parameters. The current system can be extended to search all types of web-services rather than E-learning services which have been taken as a test domain. E-learning module can also be enhanced via implementing through Semantic Web as their practical implementation.Other criteria for future implementation of learning can be feedback facilities, data mining functionalities and credit discussion.

VI. CONCLUSION

In this proposal concept of **"Aspect based searching in E-Learning"** domainhas been presented for the purpose of effective and appropriate discovery of web services within acceptable performance time. The presented searching scheme incorporates intelligence mechanism which hires Client Centric approaches like user rating anduser defined quality of service. Both the mentioned parameters are taken into the consideration and an average resultant is generated on which the sorting of the services is performed. As a result, the optimal service as per the desired QoS parameter is fetched from the registry. In addition an enhanced keyword based algorithm using Query expansion has been used to increase the appropriateness of search results Thus the system extracts the best services as per the user's interest. Further it has been established from the experiments that the search results usually gets refined when user chooses to use QoS for web service discovery in addition to simple search on functional parameters.

REFERENCES

[1] Web Services
http://www.w3schools.com/webservices
http://en.wikipedia.org/wiki/Web_service
http://www.w3schools.com/webservices/ws_intro.asp

[2] Universal Description Discovery and Integration (UDDI) Version 3.0.2 UDDI SpecTechnical Committee Draft, Dated 20041019, Available at:http://www.uddi.org/pubs/uddi_v3.htm

[3] XML,http://www.w3schools.com/xml/default.asp

[4] E-Learning, *en.wikipedia.org/wiki/**E-learning***

[5] Web Services Description Language (WSDL) version 2.0 Part 1,Core Language, W3C working draft 26 March 2000 www.w3.org/TR/2004/WD-wsdl20-20040326/wsdl20.pdf

[6] Eyhab Al-Masri, Qusay H. Mahmoud. "QoS-based discovery and Ranking of Web Services" IEEE2007.http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4317873&tag=1

[7] Nirmal Dagdee, Urjita Thakar, Vandan Tewari, Appropriate Service Discovery from UDDI through application of Data Mining Techniques, International Conference on Emerging trends in Computing (ICETiC'09), January 08-10, 2009.

[8] Vandan Tiwari, N. Dagdee, A. Tiwari, and D. Dixit, "Extended SOA to Enable Web Service Discovery on Non Functional Parameters," IJECCE, vol. 3, no. 2, pp. 97–100, 2012

[9] Eyhab AI-Masri and Qusay H. Mahmoud, "Discovering the best Web Service", WWW 2007, Canada, 2007

[10] Ziqiang Xu, Patrick Martin, Wendy Powley and FarhanaZ.2007,"Reputation Enhanced QoS-based Web services Discovery", IEEE International Conference on Web Services (ICWS 2007), IEEE 2007

[11] Reiner Kraft, Chi Chao Chang, Farzin Maghoul, and Ravi Kumar. 2006. Searching with context. In *Proceedings of the 15th international conference on World Wide Web* (WWW '06). ACM, New York, NY, USA, 477-486. DOI=10.1145/1135777.1135847 http://doi.acm.org/10.1145/1135777.1135847

[12] Vandan Tewari, Nirmal Dagdee and Aruna Tiwari. Article: User Oriented Web Services Discovery based on QoS Parameters in Multiple Registries. *International Journal of Computer Applications* 46(24):8-12, May 2012. Published by Foundation of Computer Science, New York, USA

[13] [13] Paul - Alexandru Chirita, Claudiu S. Firan, and Wolfgang Nejdl. 2007. Personalized query expansion for the web. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (SIGIR '07). ACM, New York, NY, USA, 7-14. DOI=10.1145/1277741.1277746 http://doi.acm.org/10.1145/1277741.1277746

[14] Tewari, Vandan; Dagdee, Nirmal; Singh, I.; Garg, N.; Soni, P., "An Improved Discovery Engine for Efficient and Intelligent Discovery of Web Service with Publication Facility," In proceedings of Services - II, 2009. World Conference on SERVICES-2 '09., vol., no., pp.63,70, 21-25 Sept. 2009,doi:10.1109/SERVICES-2.2009.21

[15] Semantic Discovery for Web Services | SOA World Magazine *soa.sys-con.com/node/39718*

AUTHORS

- **Vandan Tewari** is currently pursuing Ph.D. program in Comuter engineering in RGPV University, M.P., India.She is presently working as Asst. Prof. at Dept. of C.T.A at S.G.S.I.T.S.,Indore E-mail: vandantewari@gmail.com
- **Dr. Nirmal Dagdee** is working as Director and Professor in Dept. of Comp, Engg. at, S. D. Bansal College of Technology, Indore. E-mail: nirmal_dagdee@rediffmail.com
- **Dr. Aruna Tiwari** is working as Assistant professor at Dept. Of Computer Sc. & Engg. at IIT, Indore
  E-mail: artiwari@iiti.ac.in

Correspondence Author – Vandan Tewari,vandantewari@gmail.com, vtiwari@sgsits.ac.in, 09425108291