An Effective Web Service Selection based on Hybrid Collaborative Filtering and QoS-Trust Evaluation

S. Urmela[1], Dr. K. Suresh Joseph[2]
[1]M.Tech Scholar, [2]Assistant Professor, Department of Computer Science, Pondicherry University, Puducherry, India

***Abstract*— Web Service mining has become one of the predominant areas of Service Oriented Architecture. Web service discovery methods include syntactic based system and semantic based system. In the proposed work, both syntactic and semantic based approach is followed. The most widely used recommender technique is collaborative filtering. In this paper, we proposed architecture for Web Service Selection based on Hybrid Collaborative Filtering (Memory and Model based Collaborative Filtering) and QoS-Trust evaluation. Memory based Collaborative Filtering is achieved by matrix formulation and Model based Collaborative Filtering is achieved by Cluster Based Hierarchical Probabilistic Latent Semantic Analysis – Feature extractor model. The proposed work addresses various collaborative filtering problems namely, cold-start problem, gray sheep problem, synonym problem, ramp-up problem, shillings attack, data sparsity and scalability.**

***Keywords*—Web Service, recommendation system, Collaborative Filtering, Cluster Based Hierarchical Probabilistic Latent Semantic Analysis – Feature extractor.**

I. Introduction

The amount of information found in Internet is growing more and more every year, making it necessary to develop new Information Retrieval techniques. Personalized Information Retrieval system is emerging more nowadays especially when not limited to just search information but also recommend product or service to customer based on certain parameters like QoS and trust metrics thereby increasing level of users' satisfaction.

Hence, the so-called recommender system plays a key role for their ability to recommend a product or service to customer increasing customers' satisfaction. At present recommender system, proves to be effective in recommending music, financial services, twitter followers, in particular for search queries also.

In their simplest form recommender systems provide a personalized and ranked lists of items by predicting what the most suitable items users' need are, based on the users' history, preferences and constraints.

Typically, a recommender system compares a user profile to some reference characteristics, and seeks to predict the 'rating' or 'preference' that a user would give to an item. These ratings or preference can be collected either actively or passively. Active user profile collection includes: asking a user to rate an item or product after usage, presenting two different items or products and asking user to rate them on a scale of 10.

Passive user profile collection includes: Recording users' history, analyzing his/her products purchased, analyzing social network profiles and discovering his/her likes and dislikes, etc [1].

Since multiple Web Services provide same functionality, another parameter must be introduced to be set as a deciding factor. QoS is the suitable deciding factor, set of non-functional requirements like response time, accessibility, throughput, availability, etc. Current Universal description, discovery and Integration (UDDI) provide support of Web Service retrieval by functional-requirement only. Web Service mining based on Collaborative Filtering and QoS is gaining importance [2].

The organization of the paper is as follows: Section 2 describes an overview of web Service concepts, Hybrid Collaborative Filtering and Probabilistic Latent Semantic Analysis. Section 3 discusses the related work on Web Service mining by Collaborative Filtering and QoS. Section 4 depicts a novel architecture for Web Service selection based on Hybrid Collaborative Filtering and QoS-Trust evaluation. Section 5 presents experimental results and analysis of them. Section 6 summaries the paper.

II. Concepts Overview

This section gives on overview of various concepts applied in the proposed architecture for an effective Web Service selection based on Hybrid Collaborative Filtering and QoS-Trust evaluation.

A. *Web Service*

Web Services are server (service provider) and client (service requester) applications that communicates over the World Wide Web (WWW) HyperText Transfer Protocol (HTTP). Web Services are wide spreading by their interoperability, loose coupling, reusability and extensibility with the help of its components namely UDDI (Universal Description, Discovery and Integration), WSDL (Web Service Description Language), XML (eXtended Markup Language), and SOAP (Simple Object Access Protocol). UDDI is a registry where service provider registers their services; WSDL is used for describing the services; SOAP is used to transfer the data, enables communication between service provider and service requester; XML uses custom defined tags to describe the data in a structured manner [3].

B. *Collaborative Filtering*

Recommender systems are systems that recommend products or services based on users' past behaviour or consumption patterns. Recommender system is broadly classified as Content-based, Knowledge-based, Collaborative Filtering and demographic-based [4] [5]. Collaborative filtering technique aims to recommend a product or service to targeted user based on other users' ratings towards the product or service. In general

ISSN: 2278 – 1323

collaborative filtering can be classified into 2 broader areas namely; Memory based Collaborative Filtering and Model based Collaborative Filtering. Table 2.1 depicts various Collaborative Filtering techniques.

TABLE 2.1 COLLABORATIVE FILTERING TECHNIQUES

Approach	Description	Cons
Memory-based Collaborative Filtering [6][5]	Recommends product or service based on user-user similarity or item-item similarity	Data Sparsity Scalability Doesn't generalize data Overfit Gray sheep problem Cold-start problem Ramp-up problem
Model-based Collaborative Filtering [7][5]	Recommends product or service based on Bayesian network model, clustering and assumptions	Inflexible Prediction Quality Synonyms Problem Cold-start problem Ramp-up problem
Hybrid Collaborative Filtering[5][8][9] [10]	Recommends product or service by Memory and Model-based Collaborative Filtering	Different ratings evaluated-malicious attack

C. QoS Profile

We explored various research papers to study about various QoS parameters and their weightage for discovering an effective Web Service based on user query.

The various QoS parameters which we had considered essential for web service mining are listed here,

- *Availability:* Availability was measured by the mean ratio of the whole of times that the users can access the service successfully divided by the whole of time that the users use to request for the service.
- *Best practices:* Compatibility with WS-I Basic Profile.
- *Compliance:* Compatibility with WSDL specification.
- *Documentation:* Measure of documentation (description tags) in WSDL.
- *Latency:* Latency is value obtained from subtracting response time from request time of the web service invocation.
- *Reliability:* Reliability is measured by ratio of all of the times that the users request for the service successfully divided by all of the times that the users request for the service in specific time.
- *Response time:* Response time is an important quality aspect of Web service, measured from the time since the users send their requests to service sever until it was responded.
- *Successability:* Successability refers to the extent to which web service provider yields successful results over service requesters' request messages [11][12].
- *Throughput:* Throughput refers to the maximum number of services that a platform providing web services can process for a unit time [11][12].

D. Probabilistic Latent Semantic Analysis

Probabilistic Latent Semantic Analysis (PLSA) topic model is applied for retrieving Web Services by Model based Collaborative Filtering. PLSA clusters Web Services into finite number of semantic matched groups. For each domain, latent words are assumed and their probability values are fixed based on analyzing WSDL document for each domain. Each cluster is formed based on certain probability range. The cluster with maximum probability value will be the group with the most apt Web Service corresponding to Web Service requester query.

III. LITERATURE SURVEY

In general web service mining can be categorized into non-functional requirement (syntactic) based system and functional requirement (semantic) based system. With large number of web services available, retrieval based on keyword or tags alone proved to be an ineffective technique. Firstly, large number of web services might be obtained by keyword based search. Secondly, identical web services results in poor precision. As a result it leads to unusable discovered web services in complex business environment. In recent times, solutions to overcome this problem have been motivated [2].

A. By Mohamad Mehdi et al, **―Probabilistic approach-Trustworthy web service selection based on QoS**[13][20], involves a probabilistic approach for predicting the quality of a Web service based on the evaluation of past experiences (ratings) of each of its consumers[24].

QoS ratings of services are represented using different statistical distributions, namely multinomial Dirichlet (MDD), multinomial generalized Dirichlet (MGDD), and multinomial Beta-Liouville (MBLD). Bayesian inference method is employed to estimate the parameters of the mentioned distributions, which presents a trustworthy web services to service consumer

Experimental evaluation involves 3 classifier namely: classifier 1- Bayesian approach with the Beta-Liouville distribution, classifier 2- Bayesian approach with a Dirichlet and classifier 3- compare them to the state of the art naive Bayes (NB) classifier. Inflexible, quality of predictions, synonyms problem are the cons of this work.

B. By Lin, S-Y et al, **―Web service discovery-Trustworthy QoS-based collaborative filtering approach**[14], deals with a trustworthy two phase web service discovery mechanism based on collaborative filtering and QoS. In the first phase, the observer agents will collect records of user behavior, including querying and invoking web services and monitor actual QoS, and then store the profile information in the public cloud database.

This phase involves 3 sub-phases namely establishing query and web services matrices, finding query similarity and calculating the relevance between query and web services.

ISSN: 2278 – 1323

This phase mainly establishes item based (memory based) collaborative filtering. The result of phase 1 discovered services may satisfy users' functional requirements and have correct QoS information.

In the second phase, the QoS scores of the selected web services are derived from the QoS information stored in database. This phase involves 3 sub-phases namely establishing a matrix of QoS and web services, normalizing the QoS value, and calculating the QoS score. A high QoS score indicates that the web service meets the requirements of a user.

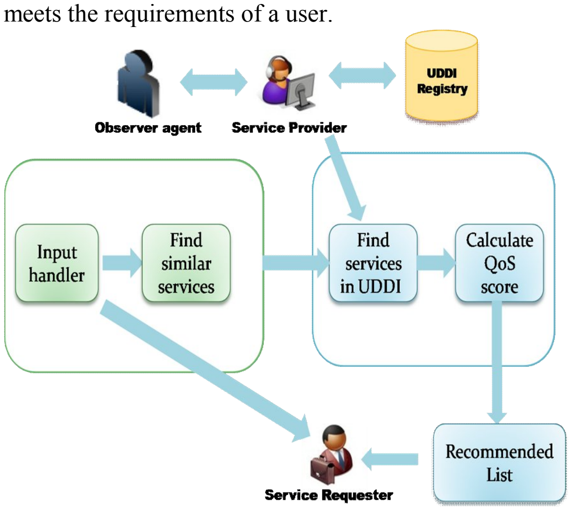

FIG 3.1 TRUSTWORTHY QOS-BASED COLLABORATIVE FILTERING APPROACH-WEB SERVICE SELECTION

Limitations:

i) Presence of observer agent, between service provider and requester which collects user's behavior is maintained in a public cloud engine, which is prone to malicious attacks

ii) In phase 1, only the similarity between users (by user profile) is considered

iii) In phase 2, ranks are calculated for web service by analyzing each QoS attribute and assigning weights to individual attribute based on user query, which is prone to frequent modification of weights of QoS attributes

iv) Data sparsity, scalability, doesn't generalize data, overfit, gray sheep problem, ramp-up problem, cold-start problem, gray sheep problem, synonym problem, shillings attack

v) Involves only 4 QoS attributes

C. By Chen et al, —**Similarity-Aware Slope One Collaborative Filtering- QoS Prediction for Web Services**[15], employs similarity-aware slope one algorithm for QoS ratings prediction. The proposed work combines both Pearson similarity and slope one measurement for QoS ratings prediction. Weight adjustment and SPC (Statistical Process Control) based smoothing is also utilized for abnormal data smoothing.

The proposed work shows better precision result compared with slope-one and famous WSRec system. The work has the capacity to reduce noise in QoS ratings data. Time-consuming, cold start problem and gray sheep problem are the cons of this work.

D. By Qi Yu et al, —**Collaborative QoS evaluation QoS-aware service selection**[16][21], proposes a service selection scheme that provides automation for assessment of QoS of an unknown service providers thereby providing a reliable web service that matches service requester's query. Relational Clustering based Model (RCM), which effectively addresses the data scarcity issue.

Experimental results of RCM model on both real and synthetic datasets demonstrates that the proposed automation model can more accurately and reliably predict the QoS parameters of an unknown web service, matching service requester's query. Data sparsity, scalability, doesn't generalize data, overfit, gray sheep problem and ramp-up problem are the cons of this work.

E. By Zheng et al, —**Collaborative Filtering QoS Aware Web Service Recommendation**[17][19], proposes a Collaborative Filtering recommendation method for QoS prediction of web services, making advantage of past usage experience of service requester. Initially, a user-collaborative mechanism for collecting past Web service QoS information from different service requester is done. Finally, based on the QoS data collected, a collaborative filtering recommendation is designed. A prototype model named, WSRec is implemented and experimental results show that proposed model achieves better prediction accuracy than traditional approaches. Data sparsity, scalability, malicious user ratings, doesn't generalize data and gray sheep problem are the cons of this work.

F. By Yechun Jiang et al, —**Personalized Collaborative Filtering- Effective Web Service Recommendation**[18][21], describes a personalized method for web service recommendation to service requester based on Collaborative Filtering. Personalized influence of service requester is taken into consideration for this approach.

An effective Personalized Hybrid Collaborative Filtering based on Personalized technique[25] is developed by integrating personalized user-based (memory-based) algorithm and personalized item-based (memory-based) algorithm. Data sparsity, scalability are the cons of this work[22].

IV. PROPOSED WEB SERVICE SELECTION ARCHITECTURE

In the proposed work, Web Service selection involves two phases i) Web Service selection by Memory-based Collaborative Filtering along with QoS-Trust evaluation ii)Web Service selection by Model-based Collaborative Filtering along with QoS-Trust evaluation. Meta-level collaborative Filtering is applied where model earned by Model-based Collaborative Filtering is applied over Memory-based Collaborative Filtering.

ISSN: 2278 – 1323

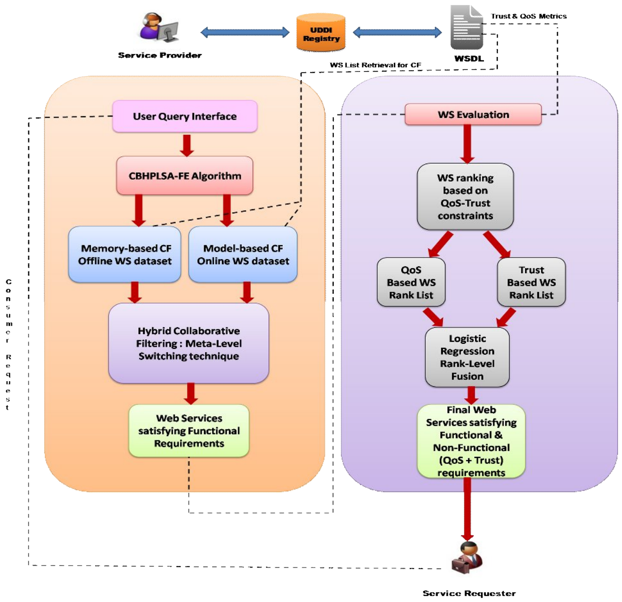

FIG 4.1 ARCHITECTURE - WEB SERVICE SELECTION BASED ON HYBRID COLLABORATIVE FILTERING AND QOS-TRUST EVALUATION

Before working on two-phases, let's take a look into the proposed Model, named Cluster Based Hierarchical Probabilistic Latent Semantic Analysis-Feature Extractor model.

Cluster Based Hierarchical Probabilistic Latent Semantic Analysis- Feature Extractor Model

By CBHPLSA-FE model (Cluster Based Hierarchical Probabilistic Latent Semantic Analysis-Feature Extractor*)*, in first level latent words for each domain are formulated and their probabilistic values are fixed by analyzing each WSDL document.

Clusters are formed with certain probability range values. First cluster is formed with having higher probability to certain range (say 1-0.8) valued Web Services, second cluster with (say 0.79-0.6), and so on.
Probability of occurrence of keyword of user query in WSDL document will be calculated simultaneously.

Given, **keywords retrieved (KW), WSDL (di), Latent Variable (Lf)**

P(di, KWj) = P(di) * P(KWj|di)
where P(di) = 1.0
P(KWj|di) = P(KWj ∩ di) / P(di)
P(KWj |di) = (f = 1 to n) P(Lf|di) * P(KWj| Lf)
where P(KWj| Lf) = P(KWj ∩ Lf) / P(Lf)
P(Lf) = 1.0

In Second level, QoS parameters of each retrieved Web Services from Web service cluster are analyzed. QoS score for each Web Service is calculated depending on domain (weather forecast, currency converter) by assigning weightage for each QoS metric. For Trust evaluation, operations of each retrieved Web Services from Web service cluster are analyzed. Ranks are assigned for each Web Service based on higher rank value (in this case 1, higher rank value). Similarly for Trust metric too on analyzing operations port as https value of 1 is given and value 0 for http port. Higher rank is given for value of trust metric with value as 1. Then both the QoS and Trust ranks are combined by Logistic Rank-Level fusion by assigning weightage to QoS and Trust ranks. Finally the weightage value of QoS and Trust metric is multiplied with their respective ranks and divided by 2(number of parameters taken for calculating ranks). One with higher rank value is the suitable Web service with promising QoS-Trust metrics.

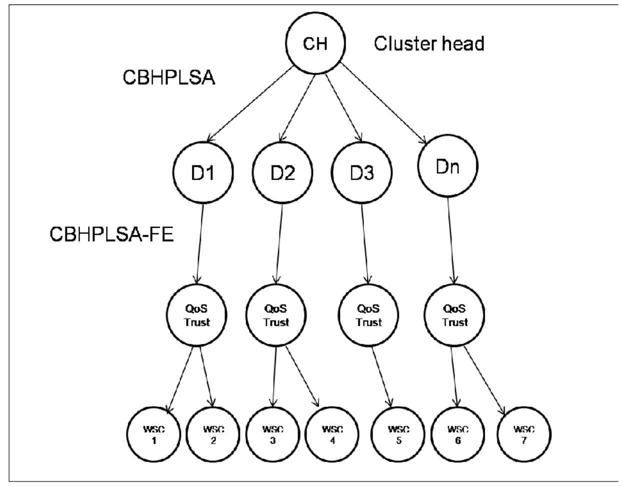

FIG 4.2 CLUSTER FORMATION

Algorithm 1: Cluster Based Hierarchical Probabilistic Latent Semantic Analysis-Feature Extractor Model (CBHPLSA-FE)

```
Input: User Query Q (keywords retrieved KW), WSDL (di), Latent Variable (Lf)
Output: Cluster of Web Services WSCn
CBHPLSA-FE(KW,di,Lf)
{

//Retrieve WS and form cluster by domain
If KW (user query) = KW (WSDL)

//Retrieve WS by Latent Variable and form new Cluster
//Retrieve WS with higher P(di, KWj)
For i = 1 to n
For j = 1 to n
//For each Latent words compute P(di, KWj) and P(KWj |di)
Compute P(di, KWj) = P(di)* P(KWj|di)      where P(di) = 1.0 , P(KWj|di) = P(KWj ∩  di) / P(di)
Compute P(KWj |di) = (For each latent word) (f = 1 to n) P(Lf|di)* P(KWj| Lf)
           where P(KWj| Lf) = P(KWj ∩  Lf) / P(Lf) , P(Lf) = 1.0
End for
End for

//Parse WSDL to identify operations (WS) which need QoS-Trust Evaluation
//Parse WSDL and apply Feature Extractor (FE) to extract WS operations
//Cluster WS by QoS-Trust by WS operations on applying FE

For each WS retrieved check operations address port
For each WS retrieved check QoS parameters
Combine QoS and Trust rank by rank-level fusion
Include it in appropriate QoS-Trust Cluster
End for
End for
}
```

Algorithm 1 depicts Cluster Based Hierarchical Probabilistic Latent Semantic Analysis-Feature Extractor Model (CBHPLSA-FE)

A. Phase 1: Web Service selection by Memory-based Collaborative Filtering along with QoS-Trust evaluation

Web Services which are stored offline are retrieved by the user query by Memory-based Collaborative Filtering. For each user query, results are obtained from the matrix formulated. It involves:

i) Cluster formation
ii) Matrix formulation
iii) Query similarity

ISSN: 2278 – 1323

i) Cluster formation by Cluster Based Hierarchical Probabilistic Latent Semantic Analysis-Feature Extractor

As said earlier, clusters are formed by CBHPLSA-FE model in-prior for the offline datasets stored.

ii) Matrix formulation

Matrix is formulated with possible user queries and Web Service Cluster. A frequency count I_{ij} indicating number of times Web Service cluster is invoked.

	WSC_1	WSC_2	WSC_3	WSC_4	WSC_5	WSC_{n-1}	WSC_n
Q_1							
Q_2				I_{ij}			
Q_3							
Q_4							
Q_5							
Q_{n-1}							
Q_n							

FIG 4.3 USER QUERY-WEB SERVICE CLUSTER MATRIX

iii) Query similarity by Jaccard Coefficient Similarity

The corresponding Web Service cluster of user query with maximum Jaccard Coefficient Similarity will be retrieved.

Jaccard Coefficient Similarity,

$$J(p,q) = Tpq / Tp + Tq + Tpq$$

Algorithm 2: Memory-Based Collaborative Filtering

Outline Algorithm - Offline Dataset: Memory-based CF

Input: Initial set of web services (Offline WS), user query records-Matrix (Previous user query-Web Service Cluster), user query
Output: WS satisfying Functional Requirements (Based on QoS-Trust)

1. Initial set of Web Services
2. Apply **CBHPLSA-FE Algorithm** to Offline Dataset
3. Get user Query
4. Calculate **Jaccard Coefficient Similarity** to user Query and Previous user Query
5. Retrieve Matched Cluster to user Query
6. If Result (Matched Cluster) = matched query then Goto step 9
7. Else Goto step 4
8. End If
9. End

B. Phase 2: Web Service selection by Model-based Collaborative Filtering along with QoS-Trust evaluation

Web Services which are stored online are retrieved by the user query by Model-based Collaborative Filtering. CBHPLSA-FE model is applied over the retrieved online Web Services.

Algorithm 3: Model-based Collaborative Filtering

Outline Algorithm - Online Dataset: Model-based CF

Input: Initial set of web services (Online WS), user query
Output: WS satisfying Functional Requirements (Based on QoS-Trust)

1. Initial set of Web Services
2. Apply **CBHPLSA-FE Algorithm** to Online Dataset
3. Get user Query
4. Retrieve Matched Cluster to user Query by semantic Matching
5. If Result (Matched Cluster) = matched query then Goto step 8
6. Else Goto step 4
7. End If
8. End

C. Removing redundant Web Services

Finally from retrieved Web Service cluster by Memory and Model-based Collaborative Filtering, redundant Web Services are checked by their names and WSDL address and they are removed. The final output being Web Services corresponding to user queries with promising QoS-Trust ranks and being non-redundant.

Algorithm 4: Memory and Model-based Collaborative Filtering

Algorithm - Offline Dataset: Memory-based CF

Input: Initial set of web services (Offline WS), user query records-Matrix (Previous user query-Web Service Cluster), user query
Output: WS satisfying Functional Requirements (Based on QoS-Trust)

```
OfflineDataset (User Query Q, Record-Matrix Mij, p – Recorded user query)
{
For previous user Query compute CBHPLSA-FE algorithm
//Apply CBHPLSA-FE Algorithm to Offline Dataset
For each Q compute Jaccard Coefficient Similarity, J(p,q) = Tpq / Tp + Tq + Tpq
            where Tpq – No. of terms common in p and q
                  Tp – No.of terms common in p
                  Tq – No.of terms common in q
End for
Retrieve Matched Cluster WSC corresponding to matched user Query from Mij
}
```

Algorithm - Online Dataset: Model-based CF

Input: Initial set of web services (Online WS), user query
Output: WS satisfying Functional Requirements (QoS-Trust)

```
OnlineDataset (User Query Q)
{
Apply CHPLSA-FE Algorithm to Online Dataset
Retrieve Matched Cluster corresponding to user query Q
}
```

The various QoS parameters considered for CBHPLSA-FE Web Service discovery model is listed here,

i) Response time
ii) Availability
iii) Throughput
iv) Successability
v) Compliance
vi) Reliability
vii) Best Practices
viii) Latency
ix) Documentation

The architecture model has promising,

i. retrieval time,
ii. better precision-recall value
iii. accuracy

In addition it overcomes problem like cold-start problem, gray sheep problem, synonym problem, ramp-up problem, shillings attack, data sparsity and scalability.

V. RESULTS AND DISCUSSIONS

Our goal is to design novel architecture for Web Service selection based on Hybrid Collaborative Filtering and QoS-Trust evaluation.

Our approach result will be Web Services retrieved by Hybrid Collaborative Filtering (Memory and Model based Collaborative Filtering) and promising QoS-Trust metrics.

A. Datasets Description

i) QWS :(Quality of Web Service) QWS is considered for Memory-based Collaborative Filtering. QWS includes 11 parameters along with values for each parameters calculated

ISSN: 2278 – 1323

in prior. QWS dataset includes more than 2500 real world Web Services.

TABLE 5.1 QUALITY OF WEB SERVICE[23]

Parameters	Description
Response Time (ms)[23]	Time for sending and receiving Web Service request[23]
Availability (%)[23]	Count of relevant service[23]
Throughput (ips)[23]	Number of requests served per unit time[23]
Successability (%)[23]	Number of responses/number of request messages[23]
Compliance (%)[23]	Compatibility with WSDL specification[23]
Reliability (%)[23]	Stable performance of retrieved Web Service[23]
Best Practices (%)[23]	Compatibility with WS-I Basic Profile[23]
Latency (ms)[23]	Processing request delay time[23]
Documentation (%)[23]	Measure of documentation (description tags) in WSDL[23]
Service Name	Web Service name[23]
WSDL Address	Corresponding WSDL address of Web Service[23]

ii) Titan Web Service search Engine:

Titan Web Service search engine is considered for Model-based Collaborative Filtering. Titan Web Service search engine includes more than 20 Web Services for each domain.

B. *Latent words and Probability values*

Latent words are analyzed using Probabilistic Latent Semantic Analysis. WSDL documents are analyzed for fixing the probability values for each domain.

TABLE 5.2 LATENT WORDS AND PROBABILITY VALUES

Domain	KW1	KW2	Latent Words	Probability
weather forecast	weather	forecast	location	0.2344
weather forecast	weather	forecast	ZIP code	0.2423
weather forecast	weather	forecast	climate	0.4172
weather forecast	weather	forecast	condition	0.2723
weather forecast	weather	forecast	wind	0.3573
weather forecast	weather	forecast	temperature	0.3732
email validator	email	validator	mail server	0.5432
email validator	email	validator	address	0.3452
email validator	email	validator	email verification	0.4523
email validator	email	validator	telecommunication	0.3456
email validator	email	validator	geolocation	0.2376
currency converter	currency	converter	rate	0.3231
currency converter	currency	converter	conversion	0.3456
currency converter	currency	converter	value manipulation	0.5674
currency converter	currency	converter	onsale	0.0934
currency converter	currency	converter	commercial	0.0912

C. *Experimental Results*

For experimental result analysis, five service requester queries was considered namely, weather forecast, domain name validator, temperature converter, currency converter and email validator. Ranking algorithm evaluation metric namely, Precision, Recall, F-Measure, Accuracy and Mean Average Precision(MAP) was evaluated for CBHPLSA-FE model of Web Service discovery. The experimental results of Hybrid Collaborative Filtering along with QoS and Trust – based on CBHPLSA-FE model was compared with 3 Web Service discovery approaches namely, keyword-based along with QoS, Memory-based Collaborative Filtering along with QoS and Trust – based on CBHPLSA-FE, Model-based Collaborative Filtering along with QoS and Trust – based on CBHPLSA-FE. The ranking metrics considered for evaluation,

i) **Precision** is a measure that the retrieved Web Service is relevant to service requester query search.

$$\text{Precision (P)} = \frac{\text{True positive}}{\text{True positive} + \text{False positive}}$$

ii) **Recall** is a measure that a relevant Web Service is retrieved in the service requester query search.

$$\text{Recall (R)} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}}$$

iii) **F-measure** is the harmonic mean of precision and recall.

$$\text{F-Measure (F)} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

iv) **Mean Average Precision** is the mean of average precision values of each service requester query.

$$\text{Mean Average Precision (MAP)} = \frac{\sum (q = 1 \text{ to } N)\ P(q)}{N}$$

v) **Accuracy** measures how well ranking algorithm correctly identifies or excludes Web Service based on service requester query.

$$\text{Accuracy (AC)} = \frac{\text{True positive} + \text{True negative}}{\text{Total population}}$$

Experimental results were observed for precision, recall, F-measure, Mean Average Precision and Accuracy on all the four Web Service discovery approaches mentioned earlier.

TABLE 5.3 PRECISION, RECALL AND F-MEASURE VALUES

Service Requester Query	WS Discovery Approach	Precision	Recall	F-measure

ISSN: 2278 – 1323

Weather forecast	**A**	0.80	0.40	0.53
	B	0.75	0.30	0.53
	C	0.80	0.40	0.53
	D	1	0.64	0.78
domain name validator	**A**	0.88	0.71	0.79
	B	0.87	0.69	0.76
	C	0.75	0.71	0.73
	D	1	0.92	0.96
temperature converter	**A**	0.75	0.80	0.77
	B	0.67	0.40	0.50
	C	0.75	0.60	0.67
	D	1	0.87	0.99
currency converter	**A**	0.66	0.71	0.69
	B	0.50	0.72	0.60
	C	0.80	0.73	0.68
	D	0.88	0.78	0.83
email validator	**A**	0.85	0.67	0.75
	B	0.84	0.72	0.78
	C	0.67	0.72	0.88
	D	1	0.88	0.94

***A – Keyword-based + QoS**
B- Memory-based CF (CBHPLSA-FE) + QoS +Trust
C- Model-based CF (CBHPLSA-FE) + QoS + Trust
D- Hybrid CF (CBHPLSA-FE) + QoS + Trust

Table 5.3 depicts the observed precision, recall, F-measure, accuracy and Mean Average Precision on the proposed Web Service discovery approach-Hybrid Collaborative Filtering (CBHPLSA-FE) along with QoS and Trust, keyword-based along with QoS, Memory-based Collaborative Filtering along with QoS and Trust, Model-based Collaborative Filtering along with QoS and Trust.

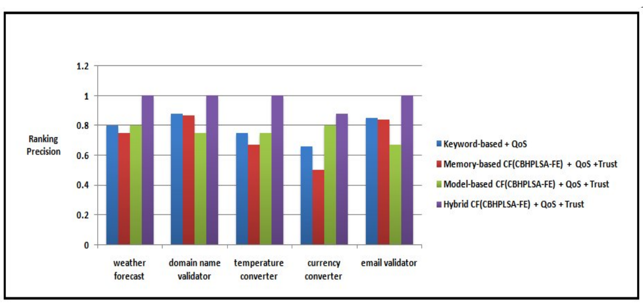

FIG 5.1 RANKING PRECISION VS WEB SERVICE DISCOVERY APPROACH

Fig 5.1 depicts average Precision value of all 5 service requester queries for Hybrid Collaborative Filtering (CBHPLSA-FE) + QoS + Trust is 0.97 which is seen to be an increased value compared to keyword-based + QoS with 0.79, Memory-based Collaborative Filtering (CBHPLSA-FE) + QoS + Trust with 0.73 and Model-based Collaborative Filtering (CBHPLSA-FE) + QoS + Trust with 0.75.

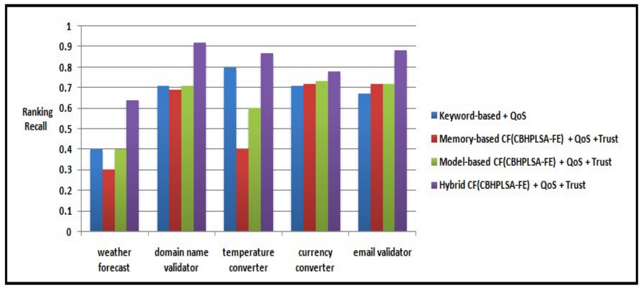

FIG 5.2 RECALL VS WEB SERVICE DISCOVERY APPROACH

Fig 5.2 depicts average Recall value of all 5 service requester queries for Hybrid Collaborative Filtering (CBHPLSA-FE) + QoS + Trust is 0.82 which is seen to be an increased value compared to keyword-based + QoS with 0.66, Memory-based Collaborative Filtering (CBHPLSA-FE) + QoS + Trust with 0.57 and Model-based Collaborative Filtering (CBHPLSA-FE) + QoS + Trust with 0.63.

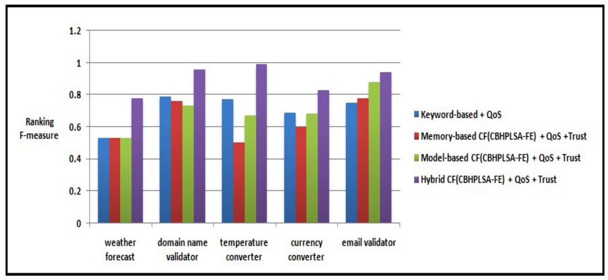

FIG 5.3 F-MEASURE VS WEB SERVICE DISCOVERY APPROACH

Both the values of Precision and Recall ranking metrics for Hybrid Collaborative Filtering seen to shown an improved value compared to other traditional Web service discovery approaches discussed.

Fig 5.3 depicts average F-measure value of all 5 service requester queries for Hybrid Collaborative Filtering (CBHPLSA-FE) + QoS + Trust is 0.90 which is seen to be an increased value compared to keyword-based + QoS with 0.71, Memory-based Collaborative Filtering (CBHPLSA-FE) + QoS + Trust with 0.63 and Model-based Collaborative Filtering (CBHPLSA-FE) + QoS + Trust with 0.70.

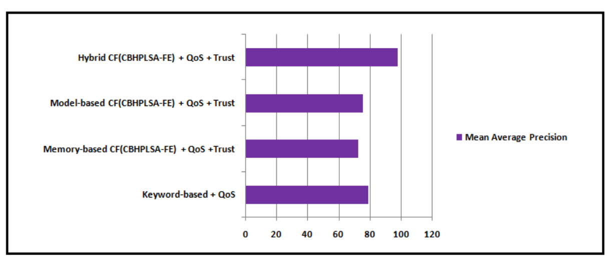

FIG 5.4 MEAN AVERAGE PRECISION VS WEB SERVICE DISCOVERY APPROACH

Similarly fig 5.4 and 5.5 Mean Average Precision(MAP) and accuracy for Hybrid Collaborative Filtering (CBHPLSA-FE) + QoS + Trust is 97.6 and 85.71 which is seen to be an improvised value compared to keyword-based + QoS, Memory-based Collaborative Filtering (CBHPLSA-FE)

ISSN: 2278 – 1323

+ QoS + Trust and Model-based Collaborative Filtering (CBHPLSA-FE) + QoS + Trust.

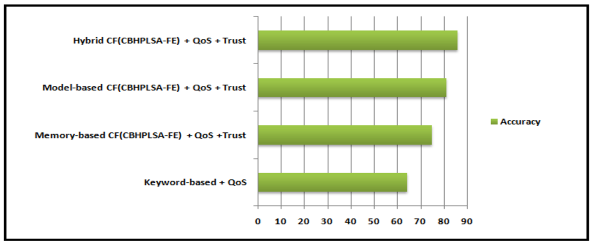

FIG 5.5 ACCURACY VS WEB SERVICE DISCOVERY APPROACH

VI. CONCLUSION AND FUTURE WORK

Web Service selection based on Hybrid Collaborative Filtering and QoS-Trust evaluation addresses various Collaborative Filtering problems namely cold-start problem, gray sheep problem, synonym problem, ramp-up problem, shillings attack, data sparsity and scalability. In addition it achieves improved precision-recall value, result accuracy and promising retrieval time.

Experimental analysis indicates that the proposed model shows an improvised result. There are many areas that can be worked out in future namely implementing genetic algorithm along with Collaborative Filtering for Web Service discovery, including more QoS parameters namely cost, accessibility, accuracy, scalability and Trust parameters namely competence, usability, efficiency, payment satisfaction, Input-Output satisfaction, Delivery satisfaction and credibility.

REFERENCES

1. P.W. Foltz, S.T. Dumais, ―Personalized information delivery: an analysis of information filtering methods, Communications of the ACM, Special issue on information filtering, Volume 35, Issue 12, December 1992, pp: 51 – 60.
2. Pradnya Khutade, Rashmi Phalnikar, ―QoS based web service discovery using oo concepts, International Journal of Advanced Technology & Engineering Research (IJATER), ISSN No: 2250-3536, Volume 2, Issue 6, November 2012, pp. 81-86.
3. Introduction to Web Services - NetBeans IDE: https://netbeans.org/kb/docs/websvc/intro-ws.html
4. Atisha Sachan, ―A Survey on Recommender Systems based on Collaborative Filtering Technique, International journal of Innovations in Engineering and technology (IJIET), ISSN. 2319-1058, Volume 2 Issue 2, April 2013.
5. A survey of Collaborative Filtering Techniques: http://www.hindawi.com/journals/aai/2009/421425/
6. Xiaoyuan Su, Taghi M. Khoshgoftaar, ―A Survey of Collaborative Filtering Techniques, Hindawi Publishing Corporation, Advances in Artificial Intelligence, Article ID 421425, 19 pages, http://dx.doi.org/10.1155/2009/421425.
7. Y. Koren, ―Tutorial on recent progress in collaborative filtering,in Proceedings of the the 2nd ACM Conference on Recommender Systems, 2008.
8. Jun Wang, ―Unifying user-based and item-based collaborative filtering approaches by similarity fusion, Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, New York, USA, ISBN:1-59593-369-7, pp. 501-508, DOI. 10.1145/1148170.1148257.
9. David M. Pennock, Eric Horvitz, Steve Lawrence, C. Lee Giles, Collaborative Filtering by Personality Diagnosis: A Hybrid Memory- and Model-Based Approach, Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI-2000), pp. 473-480, Morgan Kaufmann, San Francisco, 2000.
10. http://link.springer.com/article/10.1007/s11280-012-0186-0
11. Understanding quality of service for Web services: http://www.ibm.com/developerworks/library/wsquality/index
12. Summary of Quality Model for Web Services: https://www.oasisopen.org/committees/download.php/15444/Comparsion%20of%20WSQM%20& %20QoS%20FE.doc
13. Mohamad Mehdi, Nizar Bouguila, Jamal Bentahar, ―Probabilistic approach for QoS-aware recommender system for trustworthy web service selection, Springer Journal of Applied Intelligence, Volume 41, Issue 2 , pp 503-524, Online ISSN. 1573-7497, DOI. 10.1007/s10489-014-0537-x.
14. Szu-Yin Lin, ―A trustworthy QoS-based collaborative filtering approach for web service discovery, Journal of Systems and Software, Volume 93, July 2014, pp. 217–228, DOI: 10.1016/j.jss.2014.01.036.
15. Chengying Mao, Jifu Chen, ― QoS Prediction for Web Services Based on Similarity-Aware Slope One Collaborative Filtering, Journal of Informatica 37 (2013), pp. 139–148.
16. Qi Yu ,QoS-aware service selection via collaborative QoS evaluation, Springer Journal of World Wide Web, Volume 17, Issue 1 , pp 33-57, Online ISSN. 1573-1413, DOI. 10.1007/s11280-012-0186-0.
17. Zibin Zheng , Hao Ma, Lyu, M.R., King, I., ― QoS-Aware Web Service Recommendation by Collaborative Filtering, IEEE Transactions on service Computing, Volume. 4, Issue. 2, ISSN. 1939-1374, pp. 140-152, April-June 2011, DOI. 10.1109/TSC.2010.52.
18. Yechun Jiang , Xiangtan, China, Jianxun Liu, Mingdong Tang, Xiaoqing Liu, ―An Effective Web Service Recommendation IEEE International Conference on Web Services, Washington, DC, 4-9 July 2011, pp. 211-218, ISBN. 978-0-7695-4463-2
19. WEB SERVICE: http://kaashivinfotech.com/web-service.php
20. Probabilistic approach for QoS-aware recommender system for trustworthy web service selection : http://link.springer.com/article/10.1007/s10489-014-0537-x
21. http://link.springer.com/article/10.1007/s11280-012-0186-0
22. http://ieeexplore.ieee.org/xpl/abs_all.jsp?arnumber=6009391
23. http://www.uoguelph.ca/~qmahmoud/qws/index.html
24. http://link.springer.com/article/10.1007/s10489-014-0537-x
25. https://www.scribd.com/doc/248602141/Reduction-of-Data-Sparsity-in-Collaborative-Filtering-based-on-Fuzzy Inference-Rules Methods_based_on_Personalized_Collaborative_Filtering,2011.

ISSN: 2278 – 1323