CUSTOMIZING COMMONALITIES GROUNDED INTERNET SERVICE RECOMMENDER SYSTEM USING COLLABORATIVE FILTERING

K. SELVARANI¹, Dr. A. ANNADHASON²

¹Research Scholar, Department of Computer Applications, Alagappa University, Karaikudi, Sivagangai (District), Tamilnadu, India
selvaranik647@gmail.com

²Head & Assistant Professor & HOD, Department of Computer Science, Govt., Arts & Science College, Kadaladi, Ramnad (District), Tamilnadu, India

ABSTRACT

In the realm of the Internet, Recommender Systems (RS) play a pivotal role in enhancing data retrieval techniques, thereby optimizing the utilization of online data. These systems offer tailored recommendations for items or services to end-users, facilitating well-informed decisions. This paper explores methodologies, specifically the Pearson Correlation-Coefficient (PCC) method in conjunction with Collaborative Filtering, and the Novel Recovery-Collaborative Filtering (NRCF) method, to identify suitable web services. These methods incorporate algorithms for similarity measurement and computation within the domain of Web service recommendation systems. In contrast to existing approaches, this study introduces a novel composite clustering technique that bolsters the accuracy of similarity measurement through precise prediction. The primary goal is to enhance the efficiency of web service recommendations. The evaluation involves subjecting the PCC method to this innovative composite clustering technique, with results presented through comparative analysis.

Keywords: Customization, Commonalities, Internet Service, Recommender System, Collaborative Filtering

1. INTRODUCTION

Web services recommendation systems are the function of automatically finding the needed service and also recommending the web services over the requested user [1]. Nowadays numeric growth of information systems and the rapid increase in the number of users and data have become challenging problems in filtering the needed information. An efficient web service recommendation is to fulfill the both basic functional and non-functional requirements of the end users. Basic Functional-requirements based upon what a service does and basic non-functional requirements foundation over the Service Quality i.e.) Quality of Service (QoS) like response time, throughput, and Round-Trip Time (RTT), etc., the major role in QoS followed those same web services could be ranked and also selected by users [2]. Collaborative filtering is one on one technique that is pre-owned to portend the items that an end user may like on the grounds of ratings provided by other end-users who have the same taste in these items[3][4][5]. When identifying the similarity computation within the end-users and items, most Top N-rated recommendation algorithm is pre-owned to refer the Top N-ranked web service to end users.

1.1 Contributions

The novel contributions of this paper are:
1. We propose a novel composite clustering technique to enhance the accuracy of similarity measurement in Internet service recommender systems.
2. We introduce a tailored approach to web service recommendations through customized commonalities, ensuring more relevant suggestions.
3. We apply Pearson Correlation-Coefficient (PCC) method coupled with Collaborative Filtering to refine the identification...
of suitable web services.

4. We demonstrate the efficacy of the Novel Recovery-Collaborative Filtering (NRCF) method in augmenting the accuracy of web service recommendations.

5. We conduct comparative analysis to evaluate the impact of the composite clustering technique on the efficiency of Internet service recommendations based on the PCC method.

2. LITERATURE REVIEW

Shah et al., [16] (2022) used a recommendation system from shopping on Amazon to watching a movie on Netflix. A recommender system bases its predictions, like many machine learning algorithms, on past user behavior. They specifically forecast user preference for a group of items based on prior usage. The two most well-liked methods for developing recommender systems are collaborative filtering and content-based filtering. They used the traditional methods, named content-based filtering (CB) and collaborative-based filtering (CF), which are lacking behind because of some issues or problems like a cold start and scalability. The approach of this paper is to overcome the problems of CF as well as CB. They built an advanced recommendation system that is built with neural collaborative filtering which uses implicit feedback and finds the accuracy with the help of hit ratio which will be more accurate and efficient than the traditional recommendation system.

Sowmya et al., [17] (2022) introduced a multi-domain item reordering system based on the best explanation for an item, which are the best ranked paths extracted from a Linked Open Data knowledge graph connecting recommended and interacted items. To order paths, the algorithm assigns a value to the node attributes connecting two items by calculating the popularity of the property between interacted items that are rare in the full set of items. Results from two datasets of the movie and music domains comparing the developed reordering system with six baselines of different collaborative filtering families showed that their easy-to-explain approach improved diversity and/or accuracy metrics.

SIMANJUNTAK et al., [18] (2023) developed model-based hybrid recommendation systems that typically required extensive feature engineering to construct a user profile. Name Entity Recognition (NER) provides a straightforward way to identifies one item from a set of other items that have similar attributes of the related objects. Due to the large scale of the data used in real world recommendation systems, little research exists on applying NER models to hybrid recommendation systems in job vacancy environment. They adapted the name entity recognition approaches to construct a real hybrid job recommendation system. Furthermore, in order to satisfy a common requirement in recommendation systems the approach of accuracy, precision, recall and F-measure are using in this recommendation system in a principled way. The experimental results demonstrated the efficiency of their approach as well as its improved performance on recommendation precision.

Ghori et al., [19] (2022) explored the users’ perception and understanding of the recommender system in an empirical study using a grounded theory methodology. They draw on the cognitive theory of mind to propose a comprehensive theoretical framework that explains the users’ interpretation of the recommender system’s knowledge, reasoning, motivation, beliefs and attitudes. Their findings, based on individual in-depth interviews, suggested that users possess a sophisticated understanding of the recommender system’s behavior. Identifying the user’s understanding is a necessary step in evaluating their impact and improving recommender systems accordingly.

Lin et al., [20] (2022) developed Content-aware KG-enhanced Meta-preference Networks as a way to enhance collaborative filtering recommendation based on both structured information from KG as well as unstructured content features based on Transformer-empowered content-based filtering. They employed a novel training scheme, Cross-System Contrastive Learning, to address the inconsistency of the two very different systems and propose a powerful collaborative filtering model and a variant of the well-known NRMS system within their modeling framework. They also contributed to public domain resources through the creation of a large-scale movie-knowledge-graph dataset and an extension of the already public Amazon-Book dataset through incorporation of text descriptions crawled from external sources. We present experimental results showing that enhancing collaborative filtering with Transformer-based features derived from content-based filtering outperforms strong baseline systems, improving the ability of
knowledge-graph-based collaborative filtering systems to exploit item content information.

Rostami et al., [21] (2022) developed a hybrid food recommender-system to overcome the shortcomings of previous systems, such as ignoring food ingredients, time factor, cold start users, cold start food items and community aspects. Their method involved two phases: food content-based recommendation and user-based recommendation. Graph clustering is used in the first phase, and a deep-learning based approach is used in the second phase to cluster both users and food items. Besides a holistic-like approach is employed to account for time and user-community related issues in a way that improved the quality of the recommendation provided to the user. They compared our model with a set of state-of-the-art recommender-systems using five distinct performance metrics: Precision, Recall, F1, AUC and NDCG. Experiments using dataset extracted from “Allrecipes.com” demonstrated that the developed food recommender-system performed best.

Ibrahim et al., [22] (2023) proposed the Hybrid Neural Collaborative Filtering (HNCF) model, leveraging deep learning and interaction modeling for recommender systems using a rating matrix. To address the cold start issue, they aggregated Deep Multivariate Rating (DMR) from various external data sources, utilizing overall ratings to overcome discrepancies across sites. Their model, comprising HUAPA-DCF+NSC+DMR modules, effectively embeds user preferences and product characteristics through hierarchical attention mechanisms. The deep collaborative filtering module captures user-product interactions, while the Neural Sentiment Classifier (NSC) extracts user semantic preferences. Incorporating explicit information from DMR, their approach demonstrated superior performance over state-of-the-art models on IMDb, Yelp2013, and Yelp2014 datasets. The HNCF model significantly increased accuracy, confidence, and trust in recommendation services.

Honka et al., [23] (2022) introduced a theory-based framework called the virtual individual (VI) model to enhance personalization of digital well-being interventions. They developed an HRS prototype, With-Me HRS, utilizing knowledge-based filtering for recommending behavior change objectives and activities across lifestyle domains. The HRS incorporated essential VI model features related to well-being, lifestyle, and behavioral intention. In a real-life health-coaching program involving 50 participants, recommendations from the HRS were either visible or hidden for decision-making. Notably, 73% of participants (85% with visible vs. 62% with hidden recommendations) integrated at least one recommended activity into their coaching plans. The HRS reduced coaches' effort in identifying coaching tasks (effect size: Vargha-Delaney AD = 0.71), and participants perceived improved coaching quality (effect size: AD = 0.71). These findings established a foundation for evaluating the impact of additional user model features on the validity of recommendations from knowledge-based multi-domain HRSs.

Patro et al., [24] (2022) delivered an orderly explanation for hybrid RS along with a novel method with slight modification of the contemporary techniques, such as collaborative filtering. It also described their evolution, progression, and fruitfulness and also identifies various future implementation areas selected for future, present, and past importance.

Nachiappan et al., [25] (2022) developed an ontology focused semantic intelligence approach for the book recommendation is proposed using Jaccard similarity, NPMI, and Kullback Leibler divergence with the optimization done by the Krill Herd algorithm. The developed system OSIBR reaches an overall Precision and Accuracy of 94.16% and 95.55% respectively and with the highest Normalized Discounted Cumulative Gain (nDCG) of 0.96.

Thirupurasundari et al., [26] (2022) used a filtering method that dubbed the “hybrid approach.” The Proposed OTPS Cluster technique is used to determine TPS (Time, Place, and Service). Users’ interests and TPA recommendations are taken into account. Users can forecast the location of the temple based on the temple’s history. Collaborative Filtering and Material-Based Filtering were used to propose sites based on comparable users and content, respectively. Testing showed that the algorithm solved difficulties in standard tour routing and providing a temple visited route that was tailored to each individual’s preferences. Their study made use of data from the South Indian city of Temple in the form of temples.
<table>
<thead>
<tr>
<th>Ref No.</th>
<th>Authors</th>
<th>Method</th>
<th>Findings</th>
<th>Research Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>Shah et al., [16]</td>
<td>Neural Collaborative Filtering</td>
<td>Advanced recommendation system utilizing neural collaborative filtering for improved accuracy.</td>
<td>Addressing limitations of traditional recommendation methods (CF and CB) to enhance accuracy.</td>
</tr>
<tr>
<td>17</td>
<td>Sowmya et al., [17]</td>
<td>Multi-Domain Item Reordering</td>
<td>Introduction of a reordering system based on Linked Open Data knowledge graph, enhancing diversity and accuracy metrics.</td>
<td>Integrating Linked Open Data in multi-domain reordering systems to improve accuracy and diversity.</td>
</tr>
<tr>
<td>18</td>
<td>SIMANJUNTAK et al., [18]</td>
<td>Name Entity Recognition (NER) for Hybrid RS</td>
<td>Utilizing NER for hybrid job recommendation system, focusing on accuracy metrics.</td>
<td>Applying NER models to hybrid recommendation systems in job vacancy context, improving recommendation accuracy.</td>
</tr>
<tr>
<td>21</td>
<td>Rostami et al., [21]</td>
<td>Hybrid Food Recommender System</td>
<td>Introducing a multi-phase recommender system for food, considering ingredients, time, and user community.</td>
<td>Addressing limitations of previous food recommender systems through graph clustering and deep learning.</td>
</tr>
<tr>
<td>22</td>
<td>Ibrahim et al., [22]</td>
<td>Hybrid Neural Collaborative Filtering</td>
<td>Proposing a hybrid neural model for recommender systems using DMR and hierarchical attention mechanisms.</td>
<td>Overcoming cold start and scalability issues in traditional CF and CB methods through advanced neural model.</td>
</tr>
</tbody>
</table>
3. PROBLEM STATEMENT

The existing recommendation systems, such as collaborative-based filtering (CF) and content-based filtering (CB), encounter challenges like the cold start problem and scalability issues. These limitations hinder their ability to provide accurate and efficient recommendations, impacting user satisfaction and trust. To address these issues and enhance recommendation quality, this study aims to develop advanced models, such as neural collaborative filtering and hybrid systems, incorporating hierarchical attention mechanisms, deep learning, and semantic intelligence. The research strives to bridge the gap between traditional methods and emerging technologies, ultimately improving the accuracy and effectiveness of recommendation services across various domains, such as e-commerce, well-being interventions, job vacancies, and personalized tour routing.

3.1 Research Gap

Despite the advancements in recommendation systems, there remains a research gap in effectively combining traditional collaborative and content-based filtering techniques with innovative approaches like neural networks and semantic intelligence. Additionally, while existing studies have explored specific domains such as e-commerce and well-being interventions, there is room for further investigation into novel hybrid models that address specific challenges like the cold start problem and enhance recommendation accuracy and personalization across diverse domains such as job vacancies and personalized tour routing. This research gap highlights the need for comprehensive and adaptable recommendation approaches that harness the strengths of both established and emerging methods to provide more accurate and user-centric recommendations.

4. PROPOSED METHODOLOGY

4.1 Recommendation System

Web service Recommendation System is one kind of the strongest tools for rising profits and preserving buyers. Recommendation Systems are mostly used in various kinds of fields like E-learning, E-Commerce, Social media and so on. The recommendation system predicts the interest and preferences of users.

A simple framework of the recommendation system is as follows:

![Figure 1: Framework of Web service Recommendation system](image)

In Fig 1 for Framework of Web service Recommendation system, Each component and the arrows in the described diagram are:

1. User Profile and Items to Recommendation System:
   a. User Profile (Rectangle): This represents the collection of attributes and information about the user, including their preferences, behaviors, history, etc.
   b. Items (Rectangle): This represents the items, products, or content that the user interacts with or is interested in.

2. Arrow from User Profile to Recommendation System: This arrow indicates the flow of user profile information from the "User Profile" rectangle to the "Recommendation System" for processing.

3. Arrow from Items to Recommendation System: This arrow indicates the flow of item-related data from the "Items" rectangle to the "Recommendation System."

4. Recommendation System (Square): This represents the component that processes user profile information and item data to generate
5. Recommendation System to Recommendations:
   a. Arrow from Recommendation System to Recommendations: This arrow signifies that the output or recommendations generated by the "Recommendation System" are directed to the "Recommendations" component.
   b. Recommendation System (Square): This represents the component that processes the input data and generates personalized recommendations.
   c. Recommendations (Oval): This represents the output of the recommendation system, which includes the personalized recommendations for the user.

The Fig illustrates the flow of information and actions in a recommendation system process:
1. User profile information and item data are collected.
2. This data is fed into the recommendation system for processing.
3. The recommendation system processes the data and generates personalized recommendations.
4. The recommendations are then directed to the "Recommendations" component, which could involve presentation to the user, further filtering, or other actions.
5. Arrows visually show the connections and data flow between these components, helping to visualize how user information leads to the creation of personalized recommendations.

Web service Recommendation system is broadly classified into various types according to the Knowledge principles they use to make recommendations.

1) Collaborative Filtering Recommendation
   One of the most used recommendation system is collaborative filtering. Collaborative filtering determines the features of item or web services for a particular user to the same in order to same set of web services given by many other users.

2) Content-based Recommendation
   Content-based recommendation system is recommended items towards the user similar to the ones she and he favored in the past. It is adapted to design to recommend for the text-based items.

3) Context-aware Recommendation
   Context-aware recommendations are the multilateral idea that has been learned beyond various research zone disciplines, computer science, well-known linguistics, philosophy, organizational sciences and psychology.

4) Knowledge-based Recommendation
   Knowledge-based recommendation applies the knowledge around end users, items and also which type of products meet user's requirements. An important advantage of Knowledge-based recommendations is the non-existence of cold-start problems.

5) Demographic-based Recommendation
   Demographic-based Recommendations work on the basis of user demographic data. The main aim of Demographic-based Recommendation approach is to categorize the end user over the base of features and user's demographic information stored in their profiles, for suggesting the item.

6) Hybrid Recommendation
   Hybrid Recommendation approach combines aspect of multiple kinds of filtering approaches by computing choice of every feature there by attaining better performance. Hybrid Recommendations merges all the techniques to rise better system optimization.

4.1.1 Collaborative Filtering
   The process of identifying same end users and same web services and also recommending what similar users like is called collaborative filtering. The collaborative-filtering suggested the web services into the user, on the base of the past behavior of web service. The end user can hardly invoke all services, meaning that the Quality of Service (QoS) (round-trip time i.e. RTT) ethics of web services that the end user has not invoked are unknown. Hence, providing accurate Web service recommendation QoS prediction is very important for service users. Table 2 shows a simple example. The numerical ethics in the given table correspond to the response time into the users to invoke the indicated service. Question Mark (i.e.?) means that the end user has not invoked this service.

<table>
<thead>
<tr>
<th>Service 1</th>
<th>Service 2</th>
<th>Service 3</th>
<th>Service 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use r</td>
<td>Locati on</td>
<td>Servic e 1</td>
<td>Servic e 2</td>
</tr>
</tbody>
</table>

Table 2: User-Item Matrix with Round-trip time (RTT) A
4.1.2 Similarity Computation Measures

There are two types of similarity measures, i.e., the functional similarity measure and the nonfunctional similarity measure. Input/output/operation names are usually employed to measure the functional similarity computation within two web services. In our paper, rather than the functional similarity, focus is on the nonfunctional similarity (QoS similarity).

Given a data set consisting of M web service users and N Web services, the supplication records within users and services can be defined by an M × N matrix, which is called a user–service matrix. An entry in this matrix m, n represents a record of invocation (QoS values, e.g., latency and availability), as demonstrated in Table 1. Similarity mainly split into two type user-based and Item-based similarity measures.

1) PCC (Pearson correlation-coefficient) [7] was introduced in amount of recommender systems for similarity computation, then it may be smoothly implemented and can attain high accuracy.

A) User-based Similarity Measures:

In user-based similarity measures, the PCC can be engaged to assess the similarity computation within two user x, y by

$$\text{sim}(x, y) = \frac{\sum_{i \in I} (r_{x,i} - \bar{r}_x)(r_{y,i} - \bar{r}_y)}{\sqrt{\sum_{i \in I} (r_{x,i} - \bar{r}_x)^2} \sqrt{\sum_{i \in I} (r_{y,i} - \bar{r}_y)^2}}.$$  

(1)

Where I is the group of end users who invoked both web services i,j, and $r_{x,i}$ and $r_{y,i}$ are the QoS evaluation of the web service i observed b user x and y, respectively.

B) Item-based Similarity Measures:

The PCC could be engaged to scale the similarity computation within two web services x, y by

$$\text{sim}(i, j) = \frac{\sum_{x \in U} (r_{x,i} - \bar{r}_x)(r_{x,j} - \bar{r}_x)}{\sqrt{\sum_{x \in U} (r_{x,i} - \bar{r}_x)^2} \sqrt{\sum_{x \in U} (r_{x,j} - \bar{r}_x)^2}}.$$  

(2)

Where $U = U_i \cap U_j$ is the group of end users who invoked both web services i,j, and $r_{x,i}$ is the QoS estimate of the web-service j observed b user and $\bar{r}_x$ defines the moderated rate of the web-service j observed through users in the periods of -1 and 1.

2) NRCF

NRCF has been introduced for computing similarity within end users or items (web services) followed by the user-item value shown in Table which is usually called user-item matrix and prediction of unknown QoS value followed by the similarity computation of users and items.

A) User-based Similarity Measures:

In terms of user-based similarity measures, the PCC can be employed to measure the similarity between two users x and y by

$$\text{Sim}(x, y) = \frac{1}{|I|} \sqrt{\text{Sim}(x, y)}.$$  

(3)

Where $I = I_x \cap I_y$ is the group of web services enforced through users x and y, I is also the number of I value, $r_{x,i}$ and $r_{y,i}$ are the QoS value of item i observed by user x and y, respectively, and $r_{min}$ and $r_{max}$ defines the lowest and highest QoS value from user x in Respectively, and $r_{min}$ and $r_{max}$ defines the lowest and highest values from user x in.
y in P respectively. In (3), $Sim(x, y) \in [0, 1]$, where $Sim(x, y) = 0$ represents that two end users are dissimilar and $Sim(x, y) = 1$ indicates that two end users are the same, while the rate of the PCC is in the interval of -1 and 1.

**B) Item-based Similarity Measures:**
To calculate the correlation between two web services, similarly, based on normalizing the items QoS values, propose system guides the standard user-item matrix P into the column-normal user-item matrix $P_{ni}$, where each column is in the range of $[0,1]$. The method of NRCF is to measure the similarity computation between the two items $i,j$ by

$$Sim(i, j) = 1 - \frac{\sum_{u \in U} (r_{ui} - r_{min} - r_{ui} - r_{min})^2}{\sqrt{|U|}}$$

(4)

Where $U = U_i \cap U_j$ is the bunch of end users who invoked both items $i,j$ and $|U|$ is numeric of U, $r_{ui}$ is the scale of item i from user u in the way of original matrix $P$, $r_{min}, r_{max}, r_{min}$, and $r_{max}$ define the lowest value of item i, the highest rate as to of the item i, the lowest rate of the item j, and the foremost value of item j in the initial standard matrix Respectively.

**4.2 Clustering Algorithm**
Clustering Algorithm choose a set of similar item or set of similar users for target users. The Process of selecting similar item for similar users is crucial for the accuracy of prediction because the prediction of the unknown value depend on the corresponding values of similar users. Propose system uses clustering methodology. Clustering is of two-types: user-based clustering and item-based clustering.

**A) User-based clustering and item-based clustering**
User-based clustering is process of finding similar neighbor and makes cluster of it. Each user having its own user cluster and this cluster are used for finding unknown QoS values for that respective user. Our proposed system uses the composite clustering algorithm [7] is used for user-based clustering. User-based clustering algorithms use user-based similarity values and also Item-based clustering algorithm is used for making a cluster of the same web services. This kind of algorithm is mainly used in NRCF similarity measurement. Algorithm 1 shows used-based clustering algorithm.
Algorithm 1: User-based clustering algorithm

**Input:**
- \(x\): Target User \(T(x)\): set of another user
- \(I\): Similarity threshold
- \(K\): Number of similar neighbor to be selected

**Output:** \(S(x)\): Cluster of items for User \(x\).

1. int \(N_{sim} = 0\)
2. for all \(x_i\) such that \(x_i \in T(x)\) do
3. \(N_{sim}++\)
4. end for
5. If \(N_{sim} \geq K\) then
6. \(S(x) \leftarrow Top - K\) user
7. else if \(0 < N_{sim} < K\) then
8. \(S(x) \leftarrow Top N_{sim}\) items
9. else if \(N_{sim} = 0\) then
10. \(S(x) \leftarrow 0\)
11. end if

**B) Composite clustering**

Identifying the QoS values of a user who are all have very lowest similarity to the target user are useless or very harmful user. So, traditional Top-K algorithm is not suit for this problem. Item-based clustering [8] of NRCF similarity computation measurement is not suitable for this kind of problem. So, with the aim of address this problem, our system proposes composite clustering algorithm. Composite clustering-algorithm is the system proposed technique, which is required for both PCC item similarity computation along with PCC user similarity computation measurement values and gives the output \(S(x)\) which means that group of items for user \(x\). \(K\) is the top selected items from the whole group of items, selection is made by choosing of maximum similarity computation between items.

Algorithm 2 shows composite clustering algorithm having three threshold values is used. The first value is \(I\), which is used as item similarity computation also which is the lower limit of the qualified similar item. The second value is used as the user similarity computation measurement threshold, which is the lower limit of qualified users. Finally, the third threshold is limiting of the number of users who accessed web services \(i\).
Algorithm 2: Composite clustering algorithm

Input:
x : Target user \( W(u) \): set of web services used by target user u.
\( W(j) \): set of other users web services
I : PCC item similarity threshold U: PCC user similarity threshold
UI: Number of users those used web service i threshold
K: Number of similar items selected
Output: \( S(x) \): PCC cluster of items for user x.

1. int \( N_{sim} = 0 \)
2. for all i such that \( i \in W(i) \)do
   3. count users those access service i into C
4. for all j such that \( j \in W(j) \) and \( sim(i,j) \geq I \) and \( C \geq UI \) do
   5. \( N_{sim++} \)
6. end for
7. end for
8. if \( N_{sim} \geq K \) then
9. for all y such that \( y \in U(y) \) and \( sim(x,y) \geq U \) do
10. \( S(x) \leftarrow \) web services of y
11. end for
12. else if \( 0 < N_{sim} < K \) then
13. \( S(x) \leftarrow \) TopN \( _{sim} \) items
14. else if \( N_{sim} = 0 \) then
15. \( S(x) \leftarrow 0 \)
16. end if

5. RESULTS AND DISCUSSION

5.1 Experiment
This section describes the experimental evaluation of our prospective composite clustering techniques pre-owned within the system.

A) Web-service Recommendation QoS-Dataset:
Many kinds of standard datasets are available for experimental uses. They are Planet Lab, WS-Dream. The Proposed system uses WS-

Dream dataset of testing.

B) Evaluation Metric:
To appraise the QoS value forecast accuracy, we use the well-known mean absolute-error (MAE) metric. The MAE is the standard absolute derivation of predictions to the ground truth values.
The MAE is defined as:

$$\text{MAE} = \frac{\sum_{i,j} |r_{x,i} - P(r_{x,i})|}{N}$$

C) Performance difference of similarity computation measures:

To show the strength of our PCC similarity measure, we correlate it with another similarity computation measures like NRCF similarity Measures. Table 3 shows the forecast certainty of distinct similarity computation measures. From table, we could see that the best MAE performance of proposed PCC is superior to existing NRCF. Compared with existing NRCF, proposed PCC approach significantly improves the forecast certainty.

<table>
<thead>
<tr>
<th></th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>0.9</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCC</td>
<td>0.1208</td>
<td>0.0110</td>
<td>0.1036</td>
<td>0.0951</td>
<td>0.0866</td>
<td>0.075</td>
<td>0.695</td>
<td>0.060</td>
<td>0.0524</td>
<td>0.0439</td>
</tr>
<tr>
<td>NRCF</td>
<td>0.1357</td>
<td>0.1254</td>
<td>0.1151</td>
<td>0.1048</td>
<td>0.0946</td>
<td>0.0843</td>
<td>0.074</td>
<td>0.0637</td>
<td>0.0535</td>
<td>0.0432</td>
</tr>
</tbody>
</table>

5.2 Discussion

Here is a detailed comparison of the proposed work with other relevant studies presented in the literature:

(1) Difference from Other Studies:

A. Innovative Clustering Technique: Unlike previous works that often focus on traditional collaborative or content-based filtering, our study introduces a novel composite clustering technique tailored for web service QoS prediction. This approach is distinct from the conventional methods employed in most literature.

B. QoS Emphasis: While some prior studies explore recommendation systems in general, our work specifically addresses the challenge of web service recommendation by placing a strong emphasis on the prediction of Quality of Service (QoS) attributes, which is essential for enhancing user satisfaction.

C. Holistic Approach: Our proposed technique integrates Pearson correlation-coefficient (PCC) and Normal Recovery Collaborative Filtering methods with a composite clustering technique. This holistic approach capitalizes on the strengths of both traditional and innovative methods, resulting in improved accuracy and performance.

D. User Similarity Enhancement: The novel composite clustering technique enhances user similarity identification by considering the characteristics of web service QoS values. This is a marked departure from standard clustering techniques used in previous studies, which often do not incorporate domain-specific attributes.

E. Overcoming Cold Start and Scalability: Our proposed method strives to address limitations like the cold start problem and scalability issues commonly faced in web service recommendation. The composite clustering technique's focus on enhancing similarity measurement contributes to tackling these challenges.

(2) Significance of Difference:

A. Enhanced QoS Prediction: The integration of a composite clustering technique leads to improved prediction accuracy of QoS attributes. This advancement is particularly significant for users and service providers seeking reliable and accurate recommendations.

B. Domain Adaptability: The proposed approach's domain-specific clustering technique contributes to more tailored recommendations, catering to the unique characteristics of web services and their associated QoS attributes.

C. User Engagement: By effectively addressing the cold start problem and scalability issues, our method ensures better user engagement by providing relevant recommendations even for new users or less-explored services.

D. Improved Trust and Confidence: Enhanced recommendation accuracy and efficiency foster increased trust and confidence among users, making them more likely to rely on and utilize the recommendation system for their decision-making.

Therefore, our work significantly differentiates itself from existing literature by proposing a novel composite clustering technique tailored for web service QoS prediction, thus
bridging the gap between traditional and innovative methods while addressing challenges like cold start and scalability. The enhanced accuracy and efficiency of our approach make it a noteworthy contribution to the field of web service recommendation.

5.3 Limitations of Proposed Work

Here are some potential limitations of the proposed work:

1. **Data Scalability**: The effectiveness of the proposed composite clustering technique might diminish as the dataset scales up, potentially leading to increased computational complexity and resource requirements.

2. **Generalization**: The proposed approach's performance might vary when applied to different domains or datasets, raising concerns about its generalizability across diverse recommendation scenarios.

3. **Dependency on QoS Attributes**: The proposed clustering technique heavily relies on the availability and accuracy of QoS attributes, which might be limited or unreliable in certain datasets or environments.

4. **Cold Start Problem**: While the study aims to improve recommendation accuracy, the proposed technique's effectiveness in addressing the cold start problem for new users or items could be limited.

6. CONCLUSION

Our research makes a distinct contribution to the field of web service recommendation by enhancing the existing methodologies of Pearson correlation-coefficient (PCC) and Normal Recovery Collaborative Filtering. In this study, we delve into the intricacies of web service Quality of Service (QoS) attributes and introduce a novel clustering technique tailored to identify similar users. This innovative composite clustering approach stands out in its ability to produce more accurate QoS predictions. By leveraging this technique to forecast unknown values, our experiments demonstrate a substantial improvement in the accuracy of QoS value predictions compared to the conventional PCC approach. This research bridges the gap between traditional methods and cutting-edge clustering techniques, enhancing the precision and reliability of web service recommendations.

**Future Research Directions:** The introduction of the composite clustering technique opens avenues for further research in refining and optimizing clustering methods for web service recommendation, potentially leading to breakthroughs in accuracy and scalability.

**REFERENCES**


[10] Yali LI, Shanliang PAN and Xi HUANG,”


