A Scientometric Analysis of Transient Patterns in Recommender Systems with Soft Computing Techniques

Charu Gupta¹, Amita Jain², Oscar Castillo³, Nisheeth Joshi⁴

¹ Bhagwan Parshuram Institute of Technology, Department of Computer Science and Engineering, India
² Ambedkar Institute of Advanced Communication Technologies and Research, Department of Computer Science and Engineering, India
³ Tijuana Institute Technology, Mexico
⁴ Banasthali Vidyapith, Department of Computer Science, India

ocastillo@tectijuana.mx

Abstract. Recommender systems recommend items to users based on their interests and have seen tremendous growth due to the use of internet and web services. Recommendation systems have seen escalating growth rate since late 1990’s. A query on Google Scholar (famous research based search engine) gives 175,000 articles for the query “recommender system”. With such a large database of research/application articles, there arises a need to analyses the data so as to fulfill the basic requirements of effectively understanding the potential of the quantum of literature available so far. The study focuses on the topic of recommender system with various soft computing techniques such as fuzzy logic, neural network and genetic algorithm. The major contribution of this work is the demonstration of progressive knowledge for domain visualization and analysis of recommender system with soft computing techniques. The analysis is supported by various scientometric indicators such as Relative Growth Rate (RGR), Doubling Time (DT), Co-Authorship Index (CAI), Author Productivity, Degree of Collaboration, Research Priority Index (RPI), Half Life, Country wise Productivity, Citation Analysis, Page Length Distribution, Source Contributors. This research presents first of its kind scientometric analysis on “recommender system with soft computing techniques”. The present work provides useful parameters for establishing relationships between quantifiable data and intangible contributions in the field of recommender systems.

Keywords. Fuzzy logic, genetic algorithm, neural networks, recommender system, scientometric analysis, web of science.

1 Introduction

A recommender system in today’s context is a valuable tool for analyzing and predicting the user’s interest [1]. A recommender system is a software tool, which provides and recommends items to the users based on their interests and based on the interests of similar users [2, 3]. Recommender systems over the past decade have seen steady growth in various fields of computer science, operations research, telecommunication, automation and many more. The growth percentage of recommender systems in various fields is shown in Figure 1.

A growth of nearly 88% is observed in computer science, followed by a 25% in other engineering...
disciplines. The operations and research management have seen a rate of nearly 12% followed by telecommunications, automation, business economics, mathematics, library science and other areas.

Traditional recommender systems used memory based techniques of collaborative filtering and content based recommendation [4, 5, 6, 7, 8, 9]. With the advent of new technologies and growth of recommender systems in a company’s portfolio; it is observed that Computationally Intelligent (CI) techniques can be used in improving the efficiency and prediction accuracy in recommender systems.

To improve the effectiveness of personalized recommendations, techniques like fuzzy logic [10-19], neural network [19-21], genetic algorithms [22-24], bio-inspired algorithms [25-28], swarm intelligence [29-31] are used.

Recommender system is a popular research topic and the popularity is growing at a fast rate. This is due to the wide range of practical applicability of recommender system. Researchers, apart from focusing on native techniques of recommender system, combine soft computing techniques also.

There is no hard line between these techniques and recommender systems in terms of content and application. However, these techniques have given a new face to the use and application of recommender system. The use of soft computing techniques marked a significant change in the application areas of recommender system. These techniques improved the quality of recommendation thus resulting in increased customer satisfaction. Fig. 2 contains the corresponding data: Left: A word cloud depicting the use of tools and methods of various soft computing techniques in recommender system. Right: The percentage of use of various soft computing techniques with recommender system.

The researchers, scholars, readers can the advantage of this research in many ways. The analysis of this research provides the reader with the following benefits:

– Helps to analyze the state-of-the art techniques and methods for recommender system.
– Concerns with the exhaustive study of the quantum of literature in recommender system

Despite the tremendous growth in the area of recommendation algorithm, it is surprising to note that not much work is devoted to the scientometric study of recommendation algorithms with soft computing techniques.

An analytical/ scientometric study was done in [32, 33], which comprised of estimating the “size of science” in recommendation algorithm but only limited to a few quantitative parameters.

The study lacked the major criteria on which the recommendation works now. It also did not focus on the amalgamation of recommendation system with fuzzy logic, neural networks and genetic algorithms. The studies did not give extensive attention to the domains where recommendation algorithms with soft computing techniques are applied.
This paper analyzes the growth and advent of recommender system along with soft computing techniques in a detailed manner. In this paper, a total of 519 research articles were extracted from Web of Science database and were scrutinized carefully pertaining to a well framed query.

Out of 519, 328 articles were shortlisted with 19,352 references and approx. 1,968 keywords. To the best of our knowledge, this is the first scientometric study on recommendation system with soft computing techniques. This study provides a broad platform based on empirical results in the exploration of the research front in recommendation system.

It gives an insight into the pattern of growth using are Relative Growth Rate (RGR), Doubling Time (DT), Co-Authorship Index (CAI), Author Productivity, Degree of Collaboration, Research Priority Index (RPI), Half Life, Country wise Productivity, Citation Analysis, Page Length Distribution, Source Contributors. Anybody who is new to the field of recommendation system can find huge insights in the area with this scientometric study and can explore the potential growth areas easily. Scholars can find a great objective for diving into the field of recommendation.

This analysis will also help them to finalize the future vision of their research. Till date, the field of recommendation system with soft computing paradigms from year 2002 – till date are an unexplored area by the means of scientometric.

The rest of this paper is organized as follows. Section 2 presents the motivation behind the systematic literature survey on recommender system. Section 3 explains the objective behind this study. The dataset details and methodology is given in section 4. Section 5 discusses the various quantitative scientometric measures like Relative Growth Rate (RGR), Doubling Time (DT), Co-Authorship Index (CAI), Author Productivity, Degree of Collaboration, Research Priority Index (RPI), Half Life, Country wise Productivity, Citation Analysis, Page Length Distribution, Source Contributors for corrective evaluation of the research front. Section 6 presents the conclusion. A discussion on the areas where recommender system can be improvised is summarized in section 7.

2 Motivation

The companies in today’s business oriented environment have to make informed and effective decisions for increasing their customer base. In market there are two scenarios to be considered. First, the cost of retaining a customer is far less
Table 1. Data set details

<table>
<thead>
<tr>
<th>Database Source</th>
<th>Web of Science (WoS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Span</td>
<td>1989-2017</td>
</tr>
<tr>
<td>Citation Indices</td>
<td>SCI-EXPANDED, SSCI, A&amp;HCI, ESCI.</td>
</tr>
<tr>
<td>Query String for extraction of articles</td>
<td>(“fuzzy recommender system or recommender system type 2 fuzzy set or recommender system intuitionistic fuzzy or recommender system fuzzy sets or recommender system deep learning or recommender system convolution neural network or recommender system ant colony or recommender system particle swarm optimization or recommender system evolutionary algorithms or recommender system bee colony algorithm or recommender system digital organisms or invasive weed optimization algorithm recommender system or lifestyle recommender system or recommender system online shopping or recommender system matrix factorization”)</td>
</tr>
<tr>
<td>Total Count</td>
<td>328</td>
</tr>
<tr>
<td>Accessed on</td>
<td>[5-09-17]</td>
</tr>
</tbody>
</table>

Table 2: Sample publication format from WoS database

<table>
<thead>
<tr>
<th>PT</th>
<th>Publication Type (J=Journal; B=Book; S=Series): J</th>
</tr>
</thead>
<tbody>
<tr>
<td>AU</td>
<td>Authors: Ji, K; Shen, H</td>
</tr>
<tr>
<td>AF</td>
<td>Author Full Name: Ji, Ke; Shen, Hong</td>
</tr>
<tr>
<td>TI</td>
<td>Document Title: Jointly modelling content, social network and ratings for explainable and cold-start recommendation</td>
</tr>
<tr>
<td>SO</td>
<td>Publication Name: NEUROCOMPUTING</td>
</tr>
<tr>
<td>LA</td>
<td>Language: English</td>
</tr>
<tr>
<td>DT</td>
<td>Document Type: Article</td>
</tr>
<tr>
<td>DE</td>
<td>Author Keywords: Collaborative filtering; Recommender systems; Explanation; Cold start; Tag-keyword</td>
</tr>
<tr>
<td>ID</td>
<td>Keywords Plus®: PREFERENCES; SYSTEMS</td>
</tr>
<tr>
<td>AB</td>
<td>Abstract: Model-based approach to collaborative filtering (CF), such as latent factor models, has improved both accuracy and efficiency of predictions on large and sparse dataset. However, most existing methods still face two major problems: (1) the recommendation results derived from user and item vectors of a set of unobserved factors are lack of explanation; (2) cold start users and items out of user-item rating matrix cannot be handled accurately. In this paper, we propose a hybrid method for addressing the problems by incorporating content-based information (i.e., users’ tags and items’ keywords) and social information. The main idea behind our method is to build content association based on three factors-user interest in selected tags, tag-keyword relation and item correlation with extracted keywords, and then recommend the items with high similarity in content to users. Two novel methods-neighbor based approach and 3 factor matrix factorization are proposed for building tag-keyword relation matrix and learning user interest vector for selected tags and item correlation vector for extracted keywords. Besides, we introduce a social regularization term to help shape user interest vector. Analysis shows that our method can generate explainable recommendation results with simple descriptions, and experiments on real dataset demonstrate that our method improves recommendation accuracy of state-of-the-art CF models for previous users and items with few ratings, as well as cold start users and items with no rating. (C) 2016 Elsevier B.V. All rights reserved.</td>
</tr>
<tr>
<td>C1</td>
<td>Author Address: [Ji, Ke] Univ Jinan, Sch Informat Sci &amp; Engr, Jinan, Peoples R China; [Shen, Hong] Sun Yat Sen Univ, Sch Informat Sci &amp; Technol, Guangzhou, Guangdong, Peoples R China; [Shen, Hong] Univ Adelaide, Sch Comp Sci, Adelaide, SA 5005, Australia</td>
</tr>
<tr>
<td>RP</td>
<td>Reprint Address: Shen, H (reprint author), Sun Yat Sen Univ, Sch Informat Sci &amp; Technol, Guangzhou, Guangdong, Peoples R China; Shen, H (reprint author), Univ Adelaide, Sch Comp Sci, Adelaide, SA 5005, Australia.</td>
</tr>
<tr>
<td>NR</td>
<td>Cited Reference Count: 38</td>
</tr>
<tr>
<td>TC</td>
<td>Times Cited: 0</td>
</tr>
</tbody>
</table>

ISSN 2007-9737

196 Charu Gupta, Amita Jain, Oscar Castillo, Nisheeth Joshi

doi: 10.13053/CyS-25-1-3891
than attracting new customers for buying a particular product and second, the visitors have to be turned into real buyers.

This implies that the companies/manufacturers should use effective analytics to stay in the business market.

In today’s time, a company pays an escalated attention to the design and modelling of its recommendation system, which provides the user with the most relevant as well as novel items of his interests. The use of recommendation system enhances the sale and purchase of a product directly by a large pool of users.

In order to design an effective recommendation system, it becomes necessary to understand the interpretation and theoretic of the data relative to the quantum of literature already available.

This motives us to present an analytical view of recommendation system, which is currently being used with many computationally intelligent paradigm.
### Table 3. Relative Growth Rate and doubling time

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Year</th>
<th>No. of Published Articles</th>
<th>Cumulative Articles</th>
<th>Relative Growth Rate (RGR)</th>
<th>Doubling Time(DT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2004</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2005</td>
<td>6</td>
<td>7</td>
<td>0.15415068</td>
<td>4.49655493</td>
</tr>
<tr>
<td>3</td>
<td>2006</td>
<td>5</td>
<td>12</td>
<td>0.87546874</td>
<td>0.79174386</td>
</tr>
<tr>
<td>4</td>
<td>2007</td>
<td>5</td>
<td>17</td>
<td>1.22377543</td>
<td>0.56640049</td>
</tr>
<tr>
<td>5</td>
<td>2008</td>
<td>7</td>
<td>24</td>
<td>1.23214368</td>
<td>0.56255371</td>
</tr>
<tr>
<td>6</td>
<td>2009</td>
<td>5</td>
<td>29</td>
<td>1.75785792</td>
<td>0.39431344</td>
</tr>
<tr>
<td>7</td>
<td>2010</td>
<td>10</td>
<td>39</td>
<td>1.36097655</td>
<td>0.50930121</td>
</tr>
<tr>
<td>8</td>
<td>2011</td>
<td>12</td>
<td>51</td>
<td>1.44691898</td>
<td>0.47905032</td>
</tr>
<tr>
<td>9</td>
<td>2012</td>
<td>18</td>
<td>69</td>
<td>1.34373475</td>
<td>0.51583618</td>
</tr>
<tr>
<td>10</td>
<td>2013</td>
<td>18</td>
<td>87</td>
<td>1.57553636</td>
<td>0.43943516</td>
</tr>
<tr>
<td>11</td>
<td>2014</td>
<td>24</td>
<td>111</td>
<td>1.53147637</td>
<td>0.45260052</td>
</tr>
<tr>
<td>12</td>
<td>2015</td>
<td>41</td>
<td>152</td>
<td>1.31030845</td>
<td>0.52899529</td>
</tr>
<tr>
<td>13</td>
<td>2016</td>
<td>50</td>
<td>202</td>
<td>1.39624469</td>
<td>0.49643662</td>
</tr>
<tr>
<td>14</td>
<td>2017</td>
<td>69</td>
<td>271</td>
<td>1.36801232</td>
<td>0.50668184</td>
</tr>
<tr>
<td>15</td>
<td>2018</td>
<td>99</td>
<td>360</td>
<td>1.29098418</td>
<td>0.53691363</td>
</tr>
<tr>
<td>16</td>
<td>2019</td>
<td>113</td>
<td>470</td>
<td>1.42534488</td>
<td>0.48630125</td>
</tr>
</tbody>
</table>

Mean: 1.20580837, 0.73522668

### Table 4. Author productivity

<table>
<thead>
<tr>
<th>Block No.</th>
<th>Year</th>
<th>SA</th>
<th>TA</th>
<th>MuA 3&amp;4</th>
<th>MeA &gt;4</th>
<th>Total</th>
<th>CAI (SA)</th>
<th>CAI (TA)</th>
<th>CAI (MuA)</th>
<th>CAI (MeA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2004</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>188.88889</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2005</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>188.88889</td>
<td>62.962963</td>
<td>113.33333</td>
</tr>
<tr>
<td></td>
<td>2006</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>188.88889</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2007</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>113.33333</td>
<td>37.77778</td>
<td>204</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>3</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>17</td>
<td>0.17647059</td>
<td>0.52941176</td>
<td>0.29411765</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2008</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>44.1558442</td>
<td>121.428571</td>
<td>485.714286</td>
</tr>
<tr>
<td></td>
<td>2009</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>123.63634</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2010</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>10</td>
<td>170</td>
<td>92.7272727</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>2011</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>0</td>
<td>12</td>
<td>141.66667</td>
<td>128.78787</td>
<td>85</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>2</td>
<td>11</td>
<td>20</td>
<td>1</td>
<td>34</td>
<td>0.05882353</td>
<td>0.32352941</td>
<td>0.58823529</td>
<td>0.02941176</td>
</tr>
</tbody>
</table>
3 Objective

This study focuses on the quantitative analysis of the growth of literature in recommender system with Computationally Intelligent (CI) techniques like fuzzy logic, neural networks, genetic algorithms. It aims at the exploratory view of the trajectory of growth of recommender system with the help of various statistical tools and techniques.

The study uses Web of Science (WoS) as the source database due to its large temporal coverage of over 90 million articles.

The analysis of the data is also presented with various scientometric indicators Relative Growth...
Rate (RGR), Doubling Time (DT), Co-Authorship Index (CAI), Author Productivity, Degree of Collaboration, Research Priority Index (RPI), Half Life, Country wise Productivity, Citation Analysis, Page Length Distribution, Source Contributors.

The motive is to provide the reader with the cumulative information pack of recommender system with CI techniques.

4 Data Collection and Data Set Methodology

The data (research papers) on recommendation system have been extracted from Web of Science database for a period of almost 14 years. It starts from the growth of recommender system in the market in years 2002-2003 with all the developments that have occurred in recommender system till 2017.

The database source for extraction of articles is Web of Science (WoS), previously known as ISI Web of Knowledge. It has a temporal coverage since 1900 to present. WoS is chosen due to its large coverage of records (nearly 90 million). The details of the dataset acquired from WoS are presented in Table 1.

A total of 519 articles were obtained. Out of these, the initial screening of the articles gave 83 cleaned articles in the database. Out of 436, 108 articles were not found suitable and cleaned manually.

The articles are discarded and regarded as inappropriate because of irrelevance to the context, incomplete or redundant information or not relevant to the search string. The refined set of 328 articles is given in the reference section [1-240, 243-330].

The downloaded research publications have various fields per publication. Table 2 gives a sample publication format from the WoS Database.

The in-depth scientometric analysis has been done on Authors, Author Keyword, Source, Title, Abstract, Year of Publication, Times cited, Beginning and Ending page count, subject category, Publisher.

5 Analysis

The research articles have been studied under the umbrella of scientometrics indicators. The major indicators, which have been studied, are Relative
Growth Rate (RGR), Doubling Time (DT), Co-authorship Index (CAI), Author Productivity, Degree of Collaboration, Research Priority Index (RPI), Half Life, Country wise Productivity, Citation Analysis, Page Length Distribution, Source Contributors. Recommender systems have been extensively amalgamated with various soft computing techniques.

The year wise growth of recommender system with fuzzy logic, neural network and genetic algorithms is shown in Figure 3. The detailed description of these indicators in given in this section and the results are tabulated for effective visual analysis.

5.1 Relative Growth Rate (RGR) and Doubling Time (DT)

The mean relative growth rate is presented in Equation 1:

$$RGR = \frac{(\ln M_2 - \ln M_1)}{(t_2 - t_1)}.$$  (1)

The relative growth rate is the average number of articles/pages published per unit time [31]:

$$DT = \frac{(t_2 - t_1) \times \ln 2}{\ln c_2 - \ln c_1},$$  (2)

where, $M_1$ and $M_2$ are the amount of published articles in time period $t_1$ and $t_2$ respectively. Doubling Time (DT) is defined as the amount of time required by the research publications to get double of the existing amount. It is calculated from Equation 2 as follows:

The RGR and DT for published articles since 2002 is given in Table 3.

5.2 Co-Authorship Index (CAI)

Co-authorship Index (CAI) is the measure of collaboration between researchers for enhanced research productivity and increased specialization [241]. It is also well noted that effective collaboration leads to meaningful insights into the domain. This study uses 480 research articles for finding the CAI and to explore the relationships between the research output and amount of collaboration. A careful analysis of the articles is done on single author, two authors, and multiple author and mega author publications. Individual CAI’s are calculated.

The results are tabulated in Table 4 (SA: Single Author; TA: Two Author; MuA: Multiple Author; MeA: Mega Author). The author productivity is also computed based on citation count of the articles published by the author. A comprehensive list is given in Table 5. It is also interesting to note that the highest cited research article is authored

---

**Table 6.** Research Priority Index of different countries

<table>
<thead>
<tr>
<th>Country</th>
<th>People R China</th>
<th>USA</th>
<th>Spain</th>
<th>Australia</th>
<th>India</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td>101.79</td>
<td>96.36</td>
<td>109.6</td>
<td>107.1</td>
<td>79.34</td>
</tr>
</tbody>
</table>

**Table 7.** VOS viewer analysis

<table>
<thead>
<tr>
<th>Areas</th>
<th>Item</th>
<th>Clusters</th>
<th>Link</th>
<th>Total Link Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic algorithm based Recommender System</td>
<td>12</td>
<td>3</td>
<td>33</td>
<td>41</td>
</tr>
<tr>
<td>Neural Network based Recommender System</td>
<td>9</td>
<td>2</td>
<td>32</td>
<td>42</td>
</tr>
<tr>
<td>Probabilistic theory based Recommender System</td>
<td>16</td>
<td>15</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>Fuzzy logic based Recommender System</td>
<td>31</td>
<td>7</td>
<td>82</td>
<td>111</td>
</tr>
</tbody>
</table>

**Table 8.** Percentage of distribution of page length

<table>
<thead>
<tr>
<th>Pages</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-6</td>
<td>17</td>
<td>4.59</td>
</tr>
<tr>
<td>6-10</td>
<td>75</td>
<td>20.27</td>
</tr>
<tr>
<td>10-20</td>
<td>207</td>
<td>55.95</td>
</tr>
<tr>
<td>20+</td>
<td>71</td>
<td>19.19</td>
</tr>
<tr>
<td>370</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
by multiple authors and the author has a low research article count. This is shown in figure 5.

5.3 Degree of Collaboration

The degree of collaboration metric is used to indicate the research collaboration in a study Eq. 3:

\[ C = \frac{N_{m}}{N_{m} + N_{s}} , \]  

where C is the collaboration degree, Nm is the multiple authored papers andNs is the number of single authored papers. The value of degree of collaboration is 0.94.
5.4 Research Priority Index (RPI)

Research priority index is used to apply cross national comparisons. It is given by Equation 4:

\[
RPI = \frac{p_{ij}}{p_{io}} \times \frac{p_{oo}}{p_{oj}} \times 100,
\]

where, \( p_{ij} \) is publications of country i in j sub field; \( p_{io} \) is the publications of country i; \( p_{oj} \) is publications of all countries in sub field j; and \( p_{oo} \) is publications of all countries. Figure 6 shows the country wise research output of publication. The top 5 countries have been chosen to indicate the research priority index.

The subfields for RPI are computer science, engineering, operations research and telecommunication. Table 6 shows the RPI of different countries which contribute to the research in recommender system using soft computing techniques in computer science.

Evaluation of RPI=100 indicates average RPI; RPI>100 indicates higher RPI and RPI<100 indicates lower RPI. Table 7 indicates the growth with respect to link strength (VOS viewer analysis) in genetic algorithm, neural network, probabilistic analysis and fuzzy logic.

5.5 Half-Life

The half-life of a quantum of literature is defined as the time by which one half of the article, which is published currently in the literature becomes obsolete. According to [242], the total number of cumulative articles is 380, which give a median of 190. So the half-life of this literature is observed by subtracting the year (which achieved a score card of 190) from the year which held the first article on recommender system.

Therefore, the half-life of this quantum of literature is 12 years. A longer half-life is an indication of stable literature, which indicates that the tools and techniques do not change rapidly. On the contrary, a short half-life is equivalent to rapid obsolescence.

This is also validated from Figures 9 and 10, which indicate that the tools and techniques applied to the designing of recommender system do not change at a very fast pace.

5.6 Control Terms in Research Front

There are terms, which are frequently used with the recommender system. The analysis of these terms and its very existence support the hypothesis that these areas are more influential than others. Figure 8 depicts the unique words from the corpora with the sum of their frequencies. It is observed that the terms corresponding to fuzzy logic are more prominent.

Whereas terms belonging to other soft computing areas such as neural network or genetic algorithm are not very prominent. A careful analysis of the peer-reviewed publication reveal that the use of fuzzy logic is extensive in the area of recommendation system.

The prominent terms were clustered into Figure 8 to obtain a deep insight. Figure 11 depicts the sum of frequencies of unique keywords. These keywords are used to categorize the research publications in various areas. The categorization is done in three different areas – first, algorithms, data mining, e-commerce [243-261] second, fuzzy linguistic modelling, fuzzy logic and fuzzy sets [262-277] and third, personalization, recommendations [278-330].
5.7 Citation Analysis

The term research front was first introduced by Price in 1965 [331]. He studied the tendency of a researcher to cite an article. He found that the researchers cite the most recently published articles first. Learning to analyze the citation pattern of recommendation system which use soft computing techniques show two types of cite patterns – one which have high citation count and two, which have peaks in short interval of time.

Both the patterns have individual capacity to build the growth rate of the domain. Figure 12 shows the citation count has a peak in year 2009, where the published article (although less in number) laid the foundation stone of the recommendation systems using matrix factorization (Koren et al.) method. Since then, the growth is uniform with citation peak falling in year 2009.

5.8 Distribution of Page Length

The distribution of page length is presented in Table 8. The maximum articles are with page length of 10-20 pages and contribute to 55.95% of the total number of articles of 370.

The average length of page varies in four brackets as given in Table 8. The average page length distribution over the years have seen an increasing graph from year 2006-07, then a decrease in 2008 and again a high from 2009 to 2016 with a slight decrease in 2012 (shown in Figure 13).

5.9 Source Contributors

The WoS database was accessed to find out the journals, which published the highest amount of research papers in recommender system with soft computing paradigm. Based on the analysis of the corpus, Expert System with Applications came out to be the highest contributor to the knowledge base with a total paper count of 40 peer-reviewed research articles.

Following it are the Journal of Information Sciences and IEEE Transactions with a count of 26 research articles each. Other journals like ACM Transactions, Neuro-computing, Knowledge and Information System, Engineering Applications of Artificial Intelligence, Fuzzy sets and Systems, Electronic Commerce Research and Applications,
A large volume of literature on Recommender System with soft computing techniques for a period of 2002-2017 recorded in Web of Science has been studied and analyzed. Various scientometric tools have been employed to understand the pattern of growth in this field.

Based on the results and findings, following observations have been made. A total of 370 (Out of 391) were studied for a period of 15 years. Till 2009, the growth had a slow pace. After 2009, this research area has shown an escalating growth in the number of research publications. Significant growth in the literature has been noticed in the last nine years. More than 80% of the growth has been observed after 2009 (Figure 3).

The maximum number of articles have been published from Peoples Republic of China (115 records; 30%), followed by USA, Spain, Australia, India. India has a total 6.6% of share in total output.

The growth rate has turned up linearly till 2007 after which there is a slight decline in year 2008. After 2009 there has been a consistent growth of the output of published articles (Table 3). The mean relative growth rate and doubling time are 1.20 and 0.73 respectively.

With the growth of the research area in recommendation algorithms, it is important to note the nature of co-authorship pattern. Table 5 summarizes the research articles published by single author, two authors, three and four authors and more than four authors.

It can be observed that the CAI for initial block of four years is high for two authors, which gradually shifted to three and four authors in block 2 and 3. Initially, the CAI (two author) is 33% and 55% more than CAI (single author) and CAI (multiple author). The growth has seen significant increase in the co-authorship pattern around the year 2009, with nearly a rise of 70%.

The applications of recommender system are studied on a total of 11 subfields (Computer Science, Engineering, Operations Research Management, Science, Telecommunications, Automation Control Systems, Business Economics, Mathematics, Information Science Library Science, Science Technology and other topics and Physics). Out of which “computer science”, “Engineering”, “Telecommunication” and “operations research” were considered (by understanding the importance of recommender system in these sub fields).

The Research Priority Index (RPI) of Peoples R China, Spain and Australia >100 indicates a higher RPI whereas India and USA have lower RPI (<100).

The collaborative degree is 0.94, which depicts that the collaborative nature of research is growing and shows a positive remark on the research front.

On a more physical analysis of % of distribution of pages over a period of time indicate that 10-20 page length has the maximum value as compared to 3-6, 6-10 and 20+. The citation pattern is analysed with the most cited article referred most of the time. Although the authors of highly cited
research articles have a low document number (co-related from Table 6).

In the last eight years, the problem of recommendation is studied on various specific applications as well as in domains like Cold Start/Convergence/Data Sparsity/Scalability/Noisy Data; modelling and predicting user preferences/temporal dynamics of user preferences; Trust Model; Social Recommendation; Group Recommendation; Concept of Heterogeneous Network/feedback; context aware recommender system; Rating Matrix Optimization; Novel Approaches to Matrix Factorization; Location/Network Based Information; Optimizing Recommendations/Improving Prediction Accuracy/Improving Similarity Measures/ease of access of information; Medical Diagnosis; Ontology Based solution; Quality of Recommendation; Multi-criteria Decision Making; Attacks/Privacy.

This scientometric study comprehensively explains the momentum of recommender system — first, for designing a better recommendation system by carefully analyzing the previous researches and second, for modelling a new recommendation system which overcomes the limitation of previous recommendation systems. In this paper, two fold strategies for understanding recommender system with soft computing techniques is employed.

First, for effectively designing and modelling a recommender system, we present a detailed analysis of the growth of recommender system in the last 15 years.

This study justifies an elegant group of parameters, which help in the micro and macro level analysis of the literature. Second, a comprehensive analysis of the domains in which recommender system is employed with an emphasis on the solution strategy. The analysis further helps in understanding the two specific — areas, first, where there are research work in abundance and second, where is the scope of further improvements.

This study can provide valuable information for the academicians, research scholars, and people interested in scientometric analytical views of recommender system.

7 Discussion on Future Directions

This age of recommendation systems has seen a lot of research focused on improving the prediction of ratings/user preferences/modelling user behavior under uncertainty/improving basic approaches to improve cold start on both user and item, reduce data sparsity and increase the accuracy of recommendation. The idea is to design a computationally affordable recommender system, which takes reasonable storage space, and have low learning complexity.

During this study, the authors realize to highlight some points, which can be of use to the recommender community. The following points are enumerated for discussion, which leads to numerous future directions in field of design of recommender systems:

1. Identify and deal with situations where demands are unclear and data is scarce.
2. Design a recommender system, which has enhanced functionality, usefulness and ease of use for user requirements.
3. Building a more reliable social trust network.
4. Quality of recommendation should be enhanced and precision should be high.
5. For low learning complexity of recommender systems, boosting and co-ordinate descent can be worked out.
6. Using cognitive maps in different layers of analysis can be done with fuzzy rules, dynamic update of weights and formulation of mathematical equations.
7. Designing effective ways to elicit implicit user information left on social media by the user.
8. Finding innovative ways to attract the users to rate more and more web pages.
9. Analyzing qualitative aspects that are important for users such as diversity, coverage, and serendipity in recommendations should also be considered.
10. In Decision making problems – locality awareness, geo demographic analysis, rating sparsity, undetermined nature of ratings should be analyzed carefully.
11. Recommender system for digital products like e-greeting card and medical diagnosis should be investigated more deeply.
12. Analyzing the impact of the socio-cultural
aspect of users’ coupling relationships between users and ratings/ coupling learning has a great influence in business analytics.

13. Evolutionary algorithms for user relevance feedback can be designed with effectiveness in recommendation.


References


71. Yuen, K.K.F. (2017). The fuzzy cognitive pairwise comparisons for ranking and grade clustering to build a recommender system: An application of...


doi: 10.13053/CyS-25-1-3891


ISSN 2007-9737
214 Charu Gupta, Amita Jain, Oscar Castillo, Nisheeth Joshi

doi: 10.13053/CyS-25-1-3891


Knowledge and Data Engineering, Vol. 26, No. 11, pp. 2789–2802.


331. Price, D.J.S. (1979). Networks of Scientific Papers-The pattern of bibliographic references indicates the
A Scientometric Analysis of Transient Patterns in Recommender Systems with Soft Computing Techniques


Article received on 26/07/2020; accepted on 21/12/2020.
Corresponding author is Oscar Castillo.