# Semantic Term-Term Coupling-Based Feature Enhancement of User Profiles in Recommendation Systems

Mona Tanwar, Amity Institute of Information Technology, Amity University, Noida, India

Sunil Kumar Khatri, Amity University Tashkent, Tashkent City, Uzbekistan

 https://orcid.org/0000-0003-4373-9000

Ravi Pendse, University of Michigan, USA

## ABSTRACT

A content-based recommender system is a subclass of information systems that recommends an item to the user based on its description. It suggests items such as news, documents, articles, webpages, journals, and more to users as per their inclination by comparing the key features of the items with key terms or features of user interest profiles. This paper proposes the new methodology using Non-IIDness-based semantic term-term coupling from the content referred by users to enhance recommendation results. In the proposed methodology, the semantic relationship is analyzed by estimating the explicit and implicit relationship between terms. It associates terms that are semantically related in the real world or are used interchangeably such as synonyms. The underestimated features of user profiles have been enhanced after term-term relation analysis, which results in improved similarity estimation of relevant items with the user profiles. The experimentation result proves that the proposed methodology improves the overall search and retrieval results as compared to the state-of-the-art algorithms.

## KEYWORDS

Collaborative Filtering, Content-Filtering, Feature Enhancement, Non-IIDness Learning, Recommender Systems, Semantic Term-Term Analysis

## INTRODUCTION

In present scenario, information is continuously being generated with a velocity which is much higher than our processing capacity. If utilized properly, it can make a great difference to the world in all spheres. Various e-commerce and information sites have an abundance of products or items such as apparels, footwear, accessories, books, journals, web pages, movies, songs, hotels, restaurants, grocery and so on. From millions of items space, it is tedious and time consuming to find suitable items for the users. Recommendation systems suggest personalized items from the possible options to the users by understanding their requirement, preferences and inclinations.

Recommendation system is a significant application of big data and is involved in online e-commerce sites and business, news sites, social media, mobile applications, online journals and

digital libraries, etc. Exploration of recommender systems has gained attention in many fields such as information retrieval, social networking, data mining and machine learning. Efforts are being made to improve the recommendation's accuracy by considering factors such as social relationships, user reviews/comments, cross-domain recommendations, etc. apart from grouping similar categories of users or items.

The user's behavioural information such as purchase patterns, ratings, likes/dislikes and feedback/comments/reviews are analysed to make recommendations. Useful recommendations are essential not only to ensure good user experience but also to ensure good business from the vendor's perspective. An efficient recommender system should recommend appropriate products or services to appropriate people.

In content-based recommender systems, for content-based filtering of text-based contents like journals on the Web and Digital Libraries, news, articles, web pages, etc. keywords or terms are used for item description and for building a user profile to indicate the user's inclination. The user's feature vector is built considering the content that has been of interest to the user in the past. Content-based filtering or cognitive filtering recommends items by comparing the content of the items to a user profile. In the state-of-the-art approach, the user inclination features are retrieved based on the occurrence frequency of terms. Better representation strategies are required to incorporate content-based recommender systems with semantic intelligence which goes beyond the simple syntactic evidence of user inclinations provided by terms (Lops et al., 2011).

The proposed approach is based on Non-IIDness learning which refers to understanding, modelling, analysing and representing non-IID data (not independent and identically distributed data). Coupling and heterogeneity are the important aspects of Non-IIDness. Usually, the terms of the feature vectors are considered to be independent but there are couplings between the terms. If the couplings between the features or terms are analysed, unravelled, mathematically formulated and estimated, better recommendation results are expected. There are intricate semantic couplings between the terms which if incorporated in the user profiles, better recommendations can be achieved. For instance, many terms are semantically related in the real world though on similarity estimation simply based on co-occurrence of terms, the relation may not be inferred (Cheng et al., 2013). Also, there are few terms which are synonyms or are used interchangeably but have differing feature weights in the user profiles as the occurrence of these terms in the items of interest of the users differs. In that case, the items which should semantically match better to the user profiles do not get their due weightage since the relevant features have underestimated weights in the user profiles. Hence, along with the co-occurrence frequency (explicit relation) of terms, the implicit relationship between terms also needs to be estimated to infer the semantic closeness between terms. In the proposed work, the above issues have been addressed by learning the term coupling relationships (i.e., intra-relations and inter-relations) from the items of interest of the users to infer the semantic relationship between features or terms of the user profiles. The highly semantically related terms are coupled, and the semantic relationships are embedded in the user profiles. This is done by enhancing the underestimated relevant features that are highly semantically related to other relevant features based on the assumption that highly semantically related terms should have comparable or similar significance. During experimentation, it has been observed that the appropriate items match better with user profiles after enhancement of features.

## BACKGROUND AND RELATED WORK

The recommendation techniques that exist can be broadly categorized into content-based techniques (Lops et al., 2011), collaborative filtering techniques (Linden et al., 2003), hybrid techniques (Balabanovic & Shoham, 1997) and personalized techniques. The content-based recommendations are based on the description of items preferred earlier by the user. The attributes of a user specific profile which represents the inclination of the user are matched up with the attributes of the objects

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/semantic-term-term-coupling-based-feature-enhancement-of-user-profiles-in-recommendation-systems/281220

## Related Content

### Multi-Objective Big Data View Materialization Using NSGA-II
Akshay Kumarand T. V. Vijay Kumar (2021). *Information Resources Management Journal (pp. 1-28).*
www.irma-international.org/article/multi-objective-big-data-view-materialization-using-nsga-ii/275722

### Delineating Knowledge Flows for Enterprise Agility
Mark E. Nissen (2005). *Encyclopedia of Information Science and Technology, First Edition (pp. 779-785).*
www.irma-international.org/chapter/delineating-knowledge-flows-enterprise-agility/14335

### Developing and Implementing Machine Learning Software at Home Depot
Herbert Remidezand Sri Beldona (2021). *Journal of Cases on Information Technology (pp. 1-10).*
www.irma-international.org/article/developing-and-implementing-machine-learning-software-at-home-depot/293284

### EDI Adoption and Implementation: A Focus on Interorganizational Linkages
Carol Stoak Saundersand Sharon Clark (1992). *Information Resources Management Journal (pp. 9-20).*
www.irma-international.org/article/edi-adoption-implementation/50956

### A Case Study of IT Chargeback in a Government Agency
Dana Edbergand William L. Kuechler (2004). *Annals of Cases on Information Technology: Volume 6 (pp. 522-539).*
www.irma-international.org/chapter/case-study-chargeback-government-agency/44596