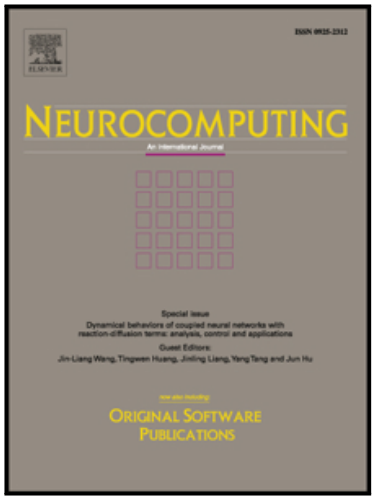
Communicated by Dr Xin Luo

Accepted Manuscript



Folksonomy-based User Profile Enrichment using Clustering and Community Recommended tags in Multiple Levels

Shubham Goel, Ravinder Kumar

PII: S0925-2312(18)30870-1
DOI: https://doi.org/10.1016/j.neucom.2018.07.035
Reference: NEUCOM 19788

To appear in: *Neurocomputing*

Received date: 13 February 2018
Revised date: 14 June 2018
Accepted date: 25 July 2018

Please cite this article as: Shubham Goel, Ravinder Kumar, Folksonomy-based User Profile Enrichment using Clustering and Community Recommended tags in Multiple Levels, *Neurocomputing* (2018), doi: https://doi.org/10.1016/j.neucom.2018.07.035

Folksonomy-based User Profile Enrichment using Clustering and Community Recommended tags in Multiple Levels

Shubham Goel, Ravinder Kumar

Computer Science and Engineering Department,
Thapar Institute of Engineering & Technology, Patiala, Punjab 147004, India

Abstract

Folksonomy (*aka Collaborative tagging*) systems provide a platform to the users where they can annotate a web resource by using any tag of interest. It is a first-hand information directly given by user without any middleman modification, therefore, it is more reliable than any other means. This paper proposes a novel methodology to construct a strong User Interest Profile (UIP) by exploiting user's own activities and other activities occurring in user's social network. UIP will provide a complete list of user preferences along with his level of interest in that preference. The proposed methodology is different from other strategies used for UIP enrichment as user's own tags are not enough to construct a strong UIP. In the current research work, two strategies have been employed for the enrichment of UIP. First one is clustering of tags based on the concept of semantic relatedness between two tags in the real world. This has been measured using Word2vec model. The second one is the utilization of user's real friendship network. It is believed that the present work is the first one to integrate the concept of semantic relatedness for tag clustering. The performance of proposed methodology has been evaluated on the basis of evaluation metrics i.e. *MRR*, *imp*, *completeness* and *P@k* using a dataset of del.icio.us. To analyse the impact of parameters, similarity measure and number of clusters in cluster set, on the performance of UIP constructed by proposed methodology extensive experiments are performed. The results reveal that the proposed methodology outperforms all the state of the art methodologies in terms of accurate and efficient UIP construction for every value of the parameters under consideration.

Keywords: User profiling, Folksonomy, Collaborative tagging, Information retrieval, Personalization, Social media.

1. Introduction

In the recent years, there has been a voluminous increase in the web size with at least 4.6 billion web pages available at present; and still the count is increasing every day with an exponential rate [1]. As well as,

*Corresponding author
Email address: ravinder@thapar.edu (Ravinder Kumar)