International Journal of Web Services Research
Volume 20 • Issue 1

A Services Classification Method Based on Heterogeneous Information Networks and Generative Adversarial Networks

Xiang Xie, Hunan University of Science and Technology, China
Jianxun Liu, Hunan University of Science and Technology, China*
Buqing Cao, Hunan University of Science and Technology, China*
Mi Peng, Hunan University of Science and Technology, China
Guosheng Kang, Hunan University of Science and Technology, China
Yiping Wen, Hunan University of Science and Technology, China
Kenneth K. Fletcher, University of Massachusetts, Boston, USA

ABSTRACT

With the rapid development of service computing and software technologies, it is necessary to correctly and efficiently classify web services to promote their discovery and application. The existing service classification methods based on heterogeneous information networks (HIN) achieve better classification performance. However, such methods use negative sampling to randomly select nodes and do not learn the underlying distribution to obtain a robust representation of the nodes. This paper proposes a web services classification method based on HIN and generative adversarial networks (GAN) named SC-GAN. The authors first construct a HIN using the structural relationships between web services and their attribute information. After obtaining the feature embedding of the services based on meta-path random walks, a relationship-aware GAN model is input for adversarial training to obtain high-quality negative samples for optimizing the embedding. Experimental results on real datasets show that SC-GAN improves classification accuracy significantly over state-of-the-art methods.

KEYWORDS

Generative Adversarial Networks, *Heterogeneous Information Network*, *Relationship-Aware*, *Services Classification*, *Web Services*

INTRODUCTION

Web services are gradually becoming the mainstream technology for implementing service-oriented architecture (SOA) applications. With the emergence of many more SOA-based applications, more and more Web services are available on the Internet today. For this reason, the rapid and accurate discovery and selection of required Web services have become a fundamental challenge in service

DOI: 10.4018/IJWSR.319960 *Corresponding Author

computing. In addition, the lack of a formal description model, too little description text, and the irregular description language, further increases the challenge of Web services discovery and selection (Ye, Cao et al. 2019).

Web services classification techniques have been studied and proposed by many researchers previously. The main goal is to reduce the space and time required for Web services search to improve the efficiency and quality of Web service discovery. Most of these studies classify Web services based on their functional attributes (Wang, Yang et al. 2017) (Xia, Fan et al. 2014). They typically employ TF-IDF, cosine similarity, and other similarity measures to determine the functional similarity between Web services based on Web services description language (WSDL) documents (Xia, Fan et al. 2014). Furthermore, several researchers have used LDA topic models or their extensions (Shi, Liu et al. 2019) (Shi, Liu et al. 2017) (Cao, Liu et al. 2017a) (Cao, Liu et al. 2019) to mine hidden topic information in Web services. These topic-model-based works represent Web services using low-dimensional topic vector features and classify Web services by computing similarities based on these topic vectors. However, it is very challenging for these approaches that only consider service content information to achieve good results because of the short length and sparse nature of WSDL documents (Cao, Liu et al. 2016).

Web services are also directly or indirectly related to other information (e.g., Tags, Mashups, etc.), which characterize the functional properties of a Web service from several perspectives(Cao, Liu et al. 2017b). Therefore, several methodologies exist to classify Web services using auxiliary relationships such as tags. Although these methodologies improve the accuracy of Web service classification to a certain extent, they rely on attribute information such as textual description information and labels that do not fully consider the complex structural interactions between Web services (combination and shared labeling relationships).

Several objects link Web services to form a natural heterogeneous information network and provide new ideas for some special Web service classification situations. This has recently led to several researchers focusing on studying node representation learning for heterogeneous information networks (HIN)(Shi, Li et al. 2016) and have applied it for service classification. HIN aims to learn to map input spaces to lower-dimensional spaces while preserving heterogeneous structure and semantics, one of the most promising of such works are Metapath2vec(Dong, Chawla et al. 2017), HERec(Shi, Hu et al. 2018) and Hin2vec(Fu, Lee et al. 2017). Though these methods achieve significant classification accuracy improvement in node classification, they also have limitations. First, they usually sample negatively by randomly selecting existing nodes in the network. Second, they focus on capturing rich semantic information over heterogeneous information networks without paying attention to the underlying node distribution. For these reasons, they do not perform well under real network situations, which tend to be sparse and noisy.

To tackle these challenges, some researchers have adopted generative adversarial network (GAN) models11 to learn potential robust representations for various applications(Ding, Tang et al. 2018) (Yu, Zhang et al. 2017). GAN depends on the idea of adversarial learning, where the discriminator and the generator compete with each other in order to train a better discriminatory model and learn the underlying data distribution. Early adoption of this technique has shown some promising preliminary results on GAN-based network representation learning(Pan, Hu et al. 2018) (Wang et al. 2018) (Fu, Lee et al. 2017). In particular, (Hu, Fang et al. 2019) proposed an adversarial learning approach based on heterogeneous information networks and showed better performance in node classification tasks.

Inspired by the above research, we propose a hybrid Web service classification method based on HIN and GAN called SC-GAN. The proposed method merges heterogeneous information networks and generative adversarial network techniques. Our proposed method uses a heterogeneous information representation model that fully extracts the structural features of Web services and adopts GAN training to capture a more robust representations of Web service nodes. Consequently, the proposed method improves the accuracy of Web service classification. To sum up, the main contributions of this paper are as follows: