DeepWSC: A Novel Framework with Deep Neural Network for Web Service Clustering

Guobing Zou[1,2], Zhen Qin[1,2], Qiang He[3], Pengwei Wang[4], Bofeng Zhang[1], Yanglan Gan[4,*]
[1]*School of Computer Engineering and Science, Shanghai University, China*
[2]*Shanghai Institute for Advanced Communication and Data Science, Shanghai University, China*
[3]*School of Software and Electrical Engineering, Swinburne University of Technology, Melbourne, Australia*
[4]*School of Computer Science and Technology, Donghua University, China*
{gbzou, zhenqin, bfzhang}@shu.edu.cn, qhe@swin.edu.au, {wangpengwei, ylgan}@dhu.edu.cn

***Abstract*—Correlative approaches have attempted to cluster web services based on either the explicit information contained in service descriptions or functionality semantic features extracted by probabilistic topic models. However, the implicit contextual information of service descriptions is ignored and has yet to be properly explored and leveraged. To this end, we propose a novel framework with deep neural network, called DeepWSC, which combines the advantages of recurrent neural network and convolutional neural network to cluster web services through automatic feature extraction. The experimental results demonstrate that DeepWSC outperforms state-of-the-art approaches for web service clustering in terms of multiple evaluation metrics.**

***Keywords*-Web service; service clustering; deep learning; probabilistic topic model; word embedding**

I. Introduction

With the rapid advancement of web technology and the increasing demands on service-oriented applications, more and more software vendors publish their applications on the Internet as web services. As of November 13, 2018, the world's largest online web service repository, ProgrammableWeb, has registered more than 20,000 API services. Those web services significantly accelerate the machine-to-machine interactions which widely promotes the development of service-oriented applications. Web service clustering has been proved to be an efficient way to facilitate service discovery [1,2] and recommendation [3]. With an overwhelming number of web services available on the Internet, web service clustering are facing more and more research challenges.

In recent years, many efforts have been devoted to research on web service clustering. Service clustering can be categorized into non-function based and function based clustering. Non-function based clustering takes the QoS vector to calculate the distance for further supporting QoS prediction [4]. Function based clustering initially focused on the similarity in service functionality by Web Service Description Language (WSDL) descriptions [2]. However, traditional clustering approaches are often difficult to obtain a satisfactory performance since the number of terms is limited in WSDL descriptions [1]. As the API mode replaces the traditional way, web service clustering approaches mainly rely on probabilistic topic model [5] and external heuristic information such as service invocation relationship [6] to extract service features. However, those recent approaches based on probabilistic topic model are based on Bag-of-Words statistical model which can mainly extract explicit features of service descriptions. They ignore the implicit contextual information of service descriptions which can boost the clustering accuracy. Therefore, how to design a novel approach for effectively clustering web services has become a key research challenge.

To this end, we propose a novel framework with deep neural network for web service clustering called DeepWSC, where a deep neural network trained based on probabilistic topic model in an unsupervised manner is performed to acquire implicit contextual information of web services. Our framework takes the advantage of recurrent convolutional neural network (RCNN) [7] and achieves better web service clustering accuracy.

The contributions of this paper are summarized as below:

- We propose a novel framework with deep neural network, DeepWSC, for web service clustering. It fundamentally boosts service discovery, composition and recommendation by effectively providing the desired web services.
- We propose an expansion strategy for more accurately extracting features of web services, which integrates an improved recurrent convolutional neural network (RCNN), an augmented probabilistic topic model (WE-LDA) and a trained word embedding model (Word2vec).
- Extensive experiments are conducted on large number of real-world web services crawled from ProgrammableWeb, which demonstrate that DeepWSC outperforms state-of-the-art approaches for web service clustering.

The remainder of this paper is organized as follows. Section II presents the details of DeepWSC. Section III evaluates DeepWSC experimentally. Finally, Section IV concludes the paper and discusses the future work.

II. Implicit Feature Extraction for Service Clustering

The implicit feature of web services is extracted by our improved RCNN that is a deep neural network service feature extractor trained based on probabilistic topic model. Figure 1 illustrates the framework of training process of our improved RCNN. The framework consists of three crucial and correlative layers. In the provision layer, as shown in Figure 1(a), the trained WE-LDA model provides a probabilistic topic distribution θ of each service as its label to guide the

978-1-7281-2717-0/19/$31.00 ©2019 IEEE
DOI 10.1109/ICWS.2019.00077

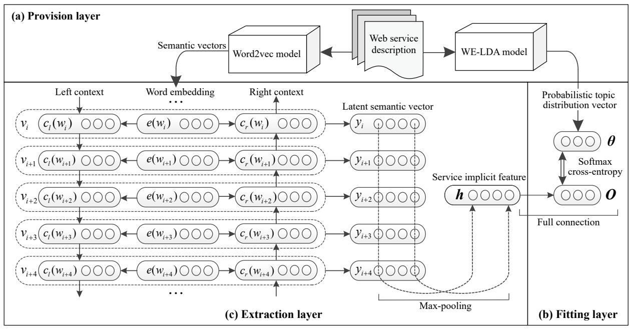

Figure 1. The training process of the improved RCNN

learning of RCNN. The fitting layer, as shown in Figure 1(b), acts as the bridge for RCNN learning to fit the probabilistic topic distribution vectors of web services. In the extraction layer, as shown in Figure 1(c), an improved RCNN is learned by taking embedded service descriptions as inputs.

To get the embedded service descriptions, each word w in a service description is projected into a vector $e(w)$ by a trained Word2vec model. Then, we learn the enhanced representation of each word by combining its contextual information with the original vector $e(w)$. Let w_i denote the i-th word of a web service description, $c_l\left(w_i\right)$ and $c_r\left(w_i\right)$ are the left and right contextual information of w_i, respectively, which are calculated as in as in (1)(2), where f is a non-linear activation function, $W^{(l)}$, $W^{(sl)}$, $W^{(r)}$ and $W^{(sr)}$ are all matrices for linear transformation. v_i is the enhanced representation of word w_i and is obtained by concatenating $c_l\left(w_i\right)$, $e\left(w_i\right)$ and $c_r\left(w_i\right)$ as in (3).

$$c_l\left(w_i\right) = f\left(W^{(l)} c_l\left(w_{i-1}\right) + W^{(sl)} e\left(w_i\right)\right) \tag{1}$$

$$c_r\left(w_i\right) = f\left(W^{(r)} c_r\left(w_{i+1}\right) + W^{(sr)} e\left(w_i\right)\right) \tag{2}$$

$$v_i = \left[c_l\left(w_i\right); e\left(w_i\right); c_r\left(w_i\right)\right] \tag{3}$$

Compared with the original RCNN in [7], our improved RCNN additionally adds $e\left(w_i\right)$ to calculation end of $c_l\left(w_i\right)$ and $c_r\left(w_i\right)$. Although the main portion of an enhanced word representation vector should be the meaning of the word itself, its contextual information still plays an auxiliary and important role in providing implicit feature of web services.

When v_i is obtained, we apply a linear transformation with the $\mathrm{leaky\ relu}$ activation function to generate the latent semantic vector y_i as in (4), where μ is a slope for negative inputs.

$$y_i = \max\left(W^{(y)} v_i + b^{(y)}, 0\right) + \mu \times \min\left(W^{(y)} v_i + b^{(y)}, 0\right) \tag{4}$$

To represent the implicit feature of a web service, we apply an element-wise max-pooling layer to capture the most important service characteristics as in (5) where L is the max length of all web services.

$$h = \max_{i=1}^{L} y_i \tag{5}$$

After that, we add a full connection layer to transform h to O, which is an unscaled probabilistic topic distribution of a web service. Finally, $\mathrm{softmax}$ is used to normalize O into probabilistic topic distribution $P = \{p_1, p_2, \ldots, p_K\}$. The cross-entropy loss function of all the n web services during the training process is defined as in (6).

$$J = -\sum\nolimits_{i=1}^{n} \left(P \log \theta + (1-P) \log\left(1-\theta\right) | s_i\right) \tag{6}$$

The objective is to minimize the loss function J. We use Adam optimizer [8] to update all the parameters that need to be learned in RCNN service implicit feature extraction.

Once the improved RCNN model is learned, we can obtain the implicit feature vectors of all web services by taking the embedded service descriptions as the inputs of RCNN feature extractor. Since they share the same dimensionality, the K-means++ clustering algorithm can be directly applied to partition a set of services into several clusters.

III. Experiments

A. Experimental Setup and Data Set

We conducted extensive experiments to validate the effectiveness of DeepWSC, where all the experiments have been carried out on a platform with an NVIDIA Tesla P100 GPU, Intel(R) Xeon(R) Gold 6130 CPU@ 2.10GHz.

To evaluate the performance of our proposed approach, we crawled 17,923 real-world web services with domain labels from ProgrammableWeb until July 1, 2018. These web services correspond to 384 categories by their domain labels. The number of services in each category is very uneven, for example, the category $Tools$ contains 887 web services, while $Solar$ only contains 2 web services. To prevent our approach from being jeopardized by extremely small clusters, we conducted the experiments on the top 20 categories. The experimental data set contains 8,459 web services. After deleting the stop words in service descriptions, the average number of words in service descriptions is 38.68.

B. Experimental Results and Analysis

We evaluate clustering results by four widely-used evaluation metrics: Purity, Normalized Mutual Information (NMI), Recall and F-measure.

We compare DeepWSC with seven competing approaches, including four existing state-of-the-art approaches, LDA [9], LDA+K [1], WE-LDA [5] and WE-LDA+K [5] and three our self-developed ones based on DeepWSC framework with RCNN [7] and Text-CNN [10], Text-CNN+LDA+K, Text-CNN+WE-LDA+K and RCNN+LDA+K, where 'K' represents 'K-means++' algorithm. In Text-CNN+LDA+K and Text-CNN+WE-LDA+K, the feature extraction is implemented by a Text-CNN learned based on LDA model and WE-LDA respectively. In RCNN+LDA+K, service implicit feature extraction is implemented jointly by the improved RCNN and LDA model. RCNN+WE-LDA+K is our proposed method based on DeepWSC framework, in which service

Table I. PERFORMANCE COMPARISIONS OF WEB SERVICE CLUSTERING AMONG COMPETING METHODS

Methods	Purity	NMI	Recall	F
LDA+K	0.5200	0.4262	0.3199	0.3383
LDA	0.5285	0.4341	0.3321	0.3503
WE-LDA+K	0.5372	0.4363	0.3282	0.3466
WE-LDA	0.5420	0.4403	0.3370	0.3543
Text-CNN+LDA+K	0.5400	0.4625	0.3484	0.3662
Text-CNN+WE-LDA+K	0.5553	0.4668	0.3572	0.3733
RCNN+LDA+K	0.5492	0.4704	0.3614	0.3784
RCNN+WE-LDA+K	**0.5708**	**0.4856**	**0.3821**	**0.3969**

implicit feature extraction is implemented jointly by the improved RCNN and WE-LDA model. Table I shows the clustering performance of all the competing methods.

It is observed from Table I that our proposed approach RCNN+WE-LDA+K outperforms the existing state-of-the-art methods. Specifically, it has an average advantage of 10.25% over WE-LDA and 12.12% over WE-LDA+K across all the evaluation metrics, which proves the superior clustering performance of our approach. We also compare our self-developed approach Text-CNN+WE-LDA+K with the methods based on WE-LDA, where the results show that it has an average advantage of 4.75% over WE-LDA and 6.72% over WE-LDA+K across all the evaluation metrics. It indicates that our DeepWSC framework implemented with other deep learning model can also extract robust service implicit feature under the guidance of WE-LDA.

Moreover, we replace WE-LDA model with LDA model in the training process of deep neural networks for web service clustering. We find that DeepWSC implemented with different variations of deep learning models outperforms the methods based on LDA among all of the evaluation metrics. Although WE-LDA is superior to LDA when integrated into DeepWSC, it demonstrates the effectiveness of our framework that can adapt to different probabilistic topic models.

To further compare the performance of our framework with different deep learning models, we can see that RCNN+WE-LDA+K has an average advantage of 5.03% over Text-CNN+WE-LDA+K across all the evaluation metrics. That is, DeepWSC with the improved RCNN model receives better performance than that with Text-CNN, since it benefits from more precise extraction of service implicit feature.

From the results, we conclude that DeepWSC implemented with an improved RCNN and WE-LDA outperforms all the state-of-the-art approaches and our self-developed ones.

IV. CONCLUSION AND FUTURE WORK

We proposed DeepWSC, a novel framework based on deep neural network for web service clustering. It first represents a web service with a semantic matrix, and then learns an improved RCNN model for more precise service implicit extraction. Finally, the task of clustering web services is performed by a conventional clustering algorithm. The results demonstrate that DeepWSC outperforms state-of-the-art approaches for web service clustering in terms of multiple widely-used evaluation metrics. In the future, we will investigate pre-processing for noisy service elimination and integrate externally auxiliary information specific to web services to the framework so that DeepWSC can more accurately extract implicit features for service clustering.

ACKNOWLEDGMENT

This work was partially supported by Shanghai Natural Science Foundation (No. 18ZR1414400, 17ZR1400200), National Key Research and Development Program of China (No. 2017YFC0907505), National Natural Science Foundation of China (No. 61772128, 61602109), and Shanghai Sailing Program (No. 16YF1400300).

REFERENCES

[1] B. Cao, X. F. Liu, B. Li, J. Liu, M. Tang, T. Zhang, and M. Shi, "Mashup service clustering based on an integration of service content and network via exploiting a two-level topic model," in *IEEE International Conference on Web Services (ICWS)*, 2016, pp. 212–219.

[2] K. Elgazzar, A. E. Hassan, and P. Martin, "Clustering WSDL documents to bootstrap the discovery of web services," in *IEEE International Conference on Web Services (ICWS)*, 2010, pp. 147–154.

[3] Z. Huang, J. Tang, G. Shan, J. Ni, Y. Chen, and C. Wang, "An efficient passenger-hunting recommendation framework with multi-task deep learning," *IEEE Internet of Things Journal*, 2019, doi: 10.1109/JIOT.2019.2901759.

[4] F. Chen, S. Yuan, and B. Mu, "User-QoS-based web service clustering for QoS prediction," in *IEEE International Conference on Web Services (ICWS)*, 2015, pp. 583–590.

[5] M. Shi, J. Liu, D. Zhou, M. Tang, and B. Cao, "WE-LDA: A word embeddings augmented LDA model for web services clustering," in *IEEE International Conference on Web Services (ICWS)*, 2017, pp. 9–16.

[6] D. He, X. Yang, Z. Feng, S. Chen, K. Huang, Z. Wang, and F. Fogelman-Soulié, "A probabilistic model for service clustering-jointly using service invocation and service characteristics," in *IEEE International Conference on Web Services (ICWS)*, 2018, pp. 302–305.

[7] S. Lai, L. Xu, K. Liu, and J. Zhao, "Recurrent convolutional neural networks for text classification," in *AAAI Conference on Artificial Intelligence (AAAI)*, 2015, pp. 2267–2273.

[8] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)*, 2015, pp. 1–15.

[9] L. Chen, Y. Wang, Q. Yu, Z. Zheng, and J. Wu, "WT-LDA: user tagging augmented LDA for web service clustering," in *International Conference on Service Oriented Computing, (ICSOC)*, 2013, pp. 162–176.

[10] Y. Kim, "Convolutional neural networks for sentence classification," in *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014, pp. 1746–1751.