Egress Initialization for Graph Convolution Network in Recommendation Systems

Sang-min Lee, Kim Heechan, Jeon Janggun and Namgi Kim
Department of Computer Science, Kyonggi University
Suwon, South Korea
[e-mail: d9249@kyonggi.ac.kr, train1212@kyonggi.ac.kr, 201812285@kyonggi.ac.kr, ngkim@kyonggi.ac.kr]
*Corresponding author: Namgi Kim

Abstract

Recommendation Systems (RS) help users quickly find the content they want on the Internet, where there is a lot of information. Graph Convolution Network (GCN) algorithms are widely used to efficiently implement such RIS. Xavier Uniform Initialization has been mainly used as a weight initialization in the learning process of GCN algorithms. However, to optimize the learning of GCN, it is necessary to apply a learning method that fits the graph structure. Therefore, we propose Egress Initialization, which initializes the distribution according to the graph embedding size, rather than Xavier Uniform Initialization, which has been traditionally used.

Keywords: Graph Convolution Network, learning method, Recommendation Systems

1. Introduction

Recommendation systems are a key tool for presenting information or content that users want in a world of information overload, and they effectively present the most relevant information or products by analyzing users' past behavior, preferences, and interests. Mainly utilized in various fields such as e-commerce, entertainment, and social media, recommendation systems are used on platforms such as YouTube, Netflix, and Amazon to present users with the content they want. Recent research on recommendation systems utilizes Graph Convolution Network (GCN) to improve performance, and many studies have been conducted to improve the learning process of graphs. Therefore, in this paper, we propose an optimized initialization function for graph learning.

2. Proposal methods

2.1 Egress Initialization

Equation 1 is the popular Xavier Uniform Initialization, where Nin is the number of input units and Nout is the number of output units (embedding size). The Xavier Initialization is designed to ensure that the variance of the output at each layer of the neural network remains the same as the variance of the input, which prevents the gradient from becoming too small or too large during training.

$$W \sim U\left(-\sqrt{\frac{6}{N_{in}+N_{out}}}, \sqrt{\frac{6}{N_{in}+N_{out}}}\right) \quad (1)$$

Xavier Initialization is especially effective when the activation function is linear or when using a symmetric activation function such as tanh.

Equation 2 is the proposed Egress Initialization, which has an initialization distribution that only considers the output of the layer.

$$W \sim U\left(-\frac{1}{\sqrt{N_{out}}}, \frac{1}{\sqrt{N_{out}}}\right) \quad (2)$$

(Support note) This research was supported by a research grant from NRF(National Research Foundation)…..

If the two initialization functions have 10,000 input units and 64 output units, the Xavier Initialization will initialize in the range (-0.024, 0.024), while the proposed Egress Initialization will initialize in the range (-0.125, 0.125). This is a 5.2x higher range. This can lead to problems such as weight divergence/explosion, but in the case of Xavier Initialization, the embedding value of the graph neural network is converging too small, so having a wider range will improve the learning process of the GNN.

3. Experiments

For the experimental data, we used the Yelp2018 dataset, which is often used in previous studies. The configuration of the dataset is shown in Table 1, and the experimental results of the existing method and the proposed initialization method are shown in Table 2. The experimental results show that by applying the egress initialization method, the learning epoch can be reduced, and the accuracy is slightly improved or degraded within the error range. This confirms that having a wider initialization distribution is superior in terms of learning speed.

Table. 1. Experiments data

	Yelp2018	
	Train	Train
User	31,668	31,668
Item	38,048	38,048
Interaction	1,237,259	1,237,259
Sparsity	99.90%	99.90%

Table. 2. Experiments result

	yelp2018		
	Epoch	**recall**	**NDCG**
BUIR(Xavier)	290	0.04371	0.03557
BUIR(Egress)	79	0.04578	0.03718
LightGCN(Xavier)	231	0.05949	0.04887
LightGCN(Egress)	186	0.06059	0.04993
SGL(Xavier)	27	0.06791	0.05585
SGL(Egress)	20	0.06480	0.05304
SimGCL(Xavier)	18	0.07248	0.05976
SimGCL(Egress)	18	0.07217	0.05932
XSimGCL(Xavier)	14	0.07277	0.06002
XSimGCL(Egress)	13	0.07241	0.05964

4. Conclusions

This paper proposed an egress initialization optimized for the learning process of GNN, and showed an increase in performance and a reduction in the learning process through experiments, and we plan to conduct future research on initialization distributions that consider the number of users and items instead of only considering the embedding size during the initialization process.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) Basic Science Research Program funded by the Ministry of Education (No. 2020R1A6A1A03040583) and this research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) and this work was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0008691, HRD Program for Industrial Innovation)

References

[1] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang *et al.*, "LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation," in *Proc. Special Interest Group in Information Retrieval (SIGIR)*, China, pp. 639–648, 2020.

[2] J. Wu, X. Wang, F. Feng, X. He, L. Chen *et al.*, "Self-supervised Graph Learning for Recommendation," in *Proc. Special Interest Group in Information Retrieval (SIGIR)*, Canada, pp. 726–735, 2021.

[3] D. Lee, S. Kang, H. ju, C. Park and H. Yu "Bootstrapping User and Item Representations for One-Class Collaborative Filtering," in *Proc. Special Interest Group in Information Retrieval (SIGIR)*, Cabada, pp. 317-326, 2021.