John (Ziyuan) Zhou

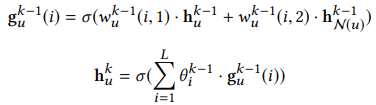
Week 6 Summary

UW ID: 2150115

In the article “IntentGC: a Scalable Graph Convolution Framework Fusing Heterogeneous Information for Recommendation”, the authors point out the big challenge for predicting users’ behaviors on the sparsity if user-item interactions on websites. The recent works didn’t fully exploit the existing rich heterogenous auxiliary relationships, and they heavily relied on linearly combined regularizers and suffered parameter tuning. To solve this problem, the authors propose a novel GCN-based framework named intentGC, to leverage both explicit preferences and heterogeneous relationships by graph convolutional networks. To apply IntentGC to web-scale applications, they also design a faster graph convolutional model called IntentNet which can avoid unnecessary feature interactions. For the methodology, the authors’ approach has three key components:

* Network Translation: translate the original graph into a special type of heterogeneous information network (HIN). The representation of this special HIN for the case of one type of auxiliary nodes is . ⊆ U × U and ⊆ V × V are the sets of generated edges between users and between items, respectively. are the user node, item node, and edge between them. For the case of R types of auxiliary nodes, we have 2R − 4 types of heterogeneous relationships where each can be denoted as () with a weight matrix (or ).
* Faster Convolutional Network: IntentNet, which takes the advantage of vector-wise convolution for scaling up and synthesizes heterogeneous relationships in an optimal sense. Same as the previous section, we also need to consider the two different cases - one type of auxiliary nodes and R types of auxiliary nodes, in vector-wise convolution operation. The final vector-wise convolutional function is same as follow:

A picture containing diagram

Description automatically generated(R type)(one type)

The function is used in the method of building stacked convolutional layers to form a network, which is IntentNet. It consists of two components: vector-wise convolution for learning the neighborhood’s utility, and fully connected layers for extracting the node-level combinatory features.

* Dual Graph Convolution: learn representations for both users and items on the translated HIN. and are used to represent the input feature vectors of user u and item v, respectively. A negative item is sampled for each user-item link to form a complete training tuple as ().



Thus, the entire workflow can be summarized as follow: 1). Network translation. 2). Generate intentNet while training. 3). Process all users and items to get their Z vectors (Dual Graph Convolution) and perform approximate K-nearest neighbors search accordingly for recommendation. For the experiment and result, the authors use two large-scale real-world datasets and online A/B tests in Alibaba demonstrate the superiority of IntentGC over recent state-of-the-art works.