7:29 61%

# CIKM'23 notification for paper 946

CIKM'23 Long Papers 详情

Dear Yiping Song,

Congratulations! We are happy to inform you t hat your FULL paper:

946 - "Multi-
scale Graph Pooling Approach with Adaptive K ey Subgraph for Graph Representations"

has been accepted for publication in the CIKM 2023 proceedings and for oral presentation at the conference. This year the review process was especially challenging with 1472 FULL pa per submissions and only 354 accepted, whic h is an acceptance rate of 24%.

Each paper was reviewed by independent revi ewers, in addition to a meta-
reviewer from our Senior Program Committee. SPCs made recommendations for each paper, based on their discussions with reviewers, an d their own assessment of the paper. PC Chair s and SPC members held meetings to discuss borderline cases and any other paper when ne eded. Final acceptance decisions were made

# Multi-scale Graph Pooling Approach with Adaptive Key Subgraph for Graph Representations

Anonymous Author(s)

## ABSTRACT

The recent progress in graph representation learning boosts the development of many graph classification tasks, such as protein classification and social network classification. One of the mainstream approaches for graph representation learning is the hierarchical pooling method. It learns the graph representation by gradually reducing the scale of the graph, so it can be easily adapted to large-scale graphs. However, existing graph pooling methods discard the original graph structure during downsizing the graph, resulting in a lack of graph topological structure. In this paper, we propose a multi-scale graph neural network (MSGNN) model that not only retains the topological information of the graph but also maintains the key-subgraph for better interpretability. MSGNN gradually discards the unimportant nodes and retains the important subgraph structure during the iteration. The key subgraphs are first chosen by experience and then adaptively evolved to tailor specific graph structures for downstream tasks. The extensive experiments on seven datasets show that MSGNN improves the SOTA performance on graph classification and better retains key subgraphs.

## CCS CONCEPTS

• **Computing methodologies** → **Artificial intelligence**.

## KEYWORDS

Graph representation learning; Graph classification; Graph pooling; Graph neural network; Key-subgraph

**ACM Reference Format:**

Anonymous Author(s). 2023. Multi-scale Graph Pooling Approach with Adaptive Key Subgraph for Graph Representations. In *Proceedings of (Conference acronym CIKM).* ACM, New York, NY, USA, 12 pages. https://doi.org/XXXXXXX.XXXXXXX

## 1 INTRODUCTION

Graph neural networks have gained considerable popularity for modeling graph data, exhibiting remarkable performance in tasks such as node classification [26, 28, 45] and link prediction [33, 53]. However, to obtain an informative graph representation, the utilization of a pooling function becomes imperative. This function maps a collection of node representations into a condensed form. Current graph pooling methods can be grossly divided into two categories: *flat pooling* methods and *hierarchical pooling* methods.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

*Conference acronym CIKM, Octorbor 21–25, 2023, Birmingham, UK*

© 2023 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

**Table 1: Differences between various hierarchical graph pooling methods**

| Methods | Each Iteration | | | Multi-scale Interaction |
|---|---|---|---|---|
| | Sample | Cluster | Evolution | |
| Graph U-net [13] | ✓ | | | |
| DiffPool [50] | | ✓ | | |
| GXN [23] | ✓ | | | ✓(Features) |
| MSGNN(Ours) | ✓ | ✓(Last) | ✓ | ✓(Structures) |

The flat pooling averages or sums all node representations in the given graph [1, 46]. These methods have limited ability to capture the graph structure, so there are some works such as DGCNN [54] proposing to read the graph in a meaningful and consistent order to keep the structure information. However, their flat architecture designs restrict their capability toward hierarchical pooling.

The hierarchical pooling retains the graph structure to some extent by recursively reducing the graph size, so it is adaptive to large graphs. One line is the Top-K pooling method, which learns a score vector to annotate the node importance and selects the top-$k$ nodes. A representative work Graph U-net [13] builds an encoder-decoder framework with graph-pooling and unpooling operations, which ensures the downsized graph after pooling can better restore the original graph. SAGPool [22] exploits an attention mechanism to distinguish whether to retain or drop nodes. GIB [51] employs the information bottleneck (IB) [38] to detect informative subgraphs. Top-K pooling discards unimportant nodes during downsizing, so this method may lose partial information of the original graph.

Another line of hierarchical pooling does not discard any nodes but downsizes the graph by clustering nodes. Diffpool [50] employs GNNs to obtain node embeddings and learns an assignment matrix for differential clustering. ASAP [31] further improves the model with a sparse pooling operator so that it can preserve node and edge information effectively. However, graphs often contain diverse substructures that provide distinctive contributions to the overall graph representation. For instance, in protein-protein interaction graphs, specific substructures may signify important functions that greatly impact the prediction of the whole graph characteristics [55]. The hierarchical clustering method does not contain any key subgraphs of the original graph during downsizing, so it can hardly preserve the structural information of the graph.

There is another challenge besides the above issues: we cannot remedy the intermediate graph if we discard some important nodes by mistake. The reason is that the existing downsizing strategies for hierarchical pooling only aggregate nodes or directly discard nodes. Note that the discarding operation is not reversible. To make up for possible errors, GXN [23] leverages cross-feature interaction layers to enable the communication and fusion of multi-scale features. However, it cannot change the structure of the intermediate graph at each scale or interaction step.