Graph Compression with Embedding for Network Alignment

RMIT UNIVERSITY

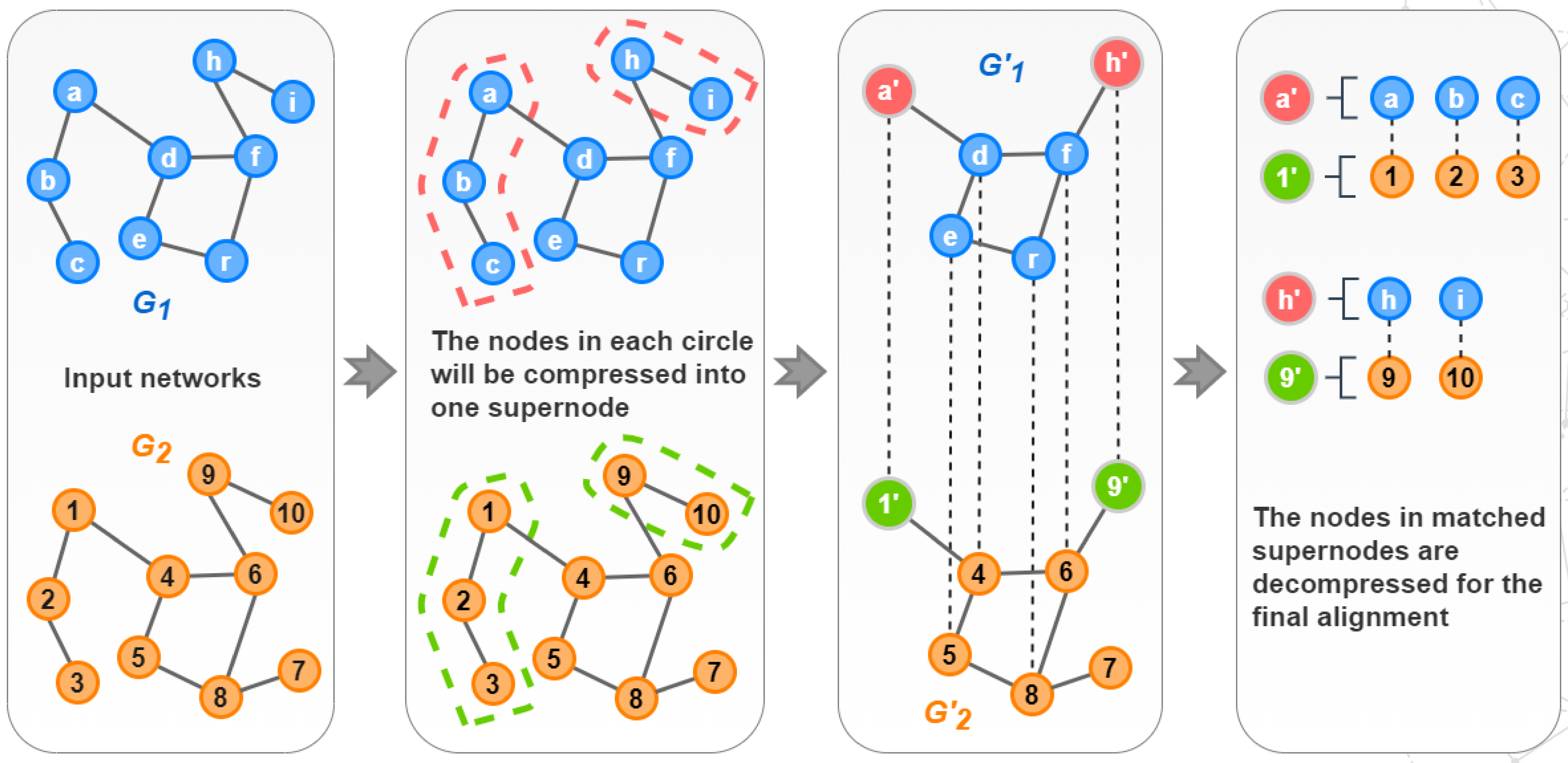
Kyle K. Qin, Flora D. Salim, Yongli Ren, Wei Shao, Mark Heimann, Danai Koutra

Introduction

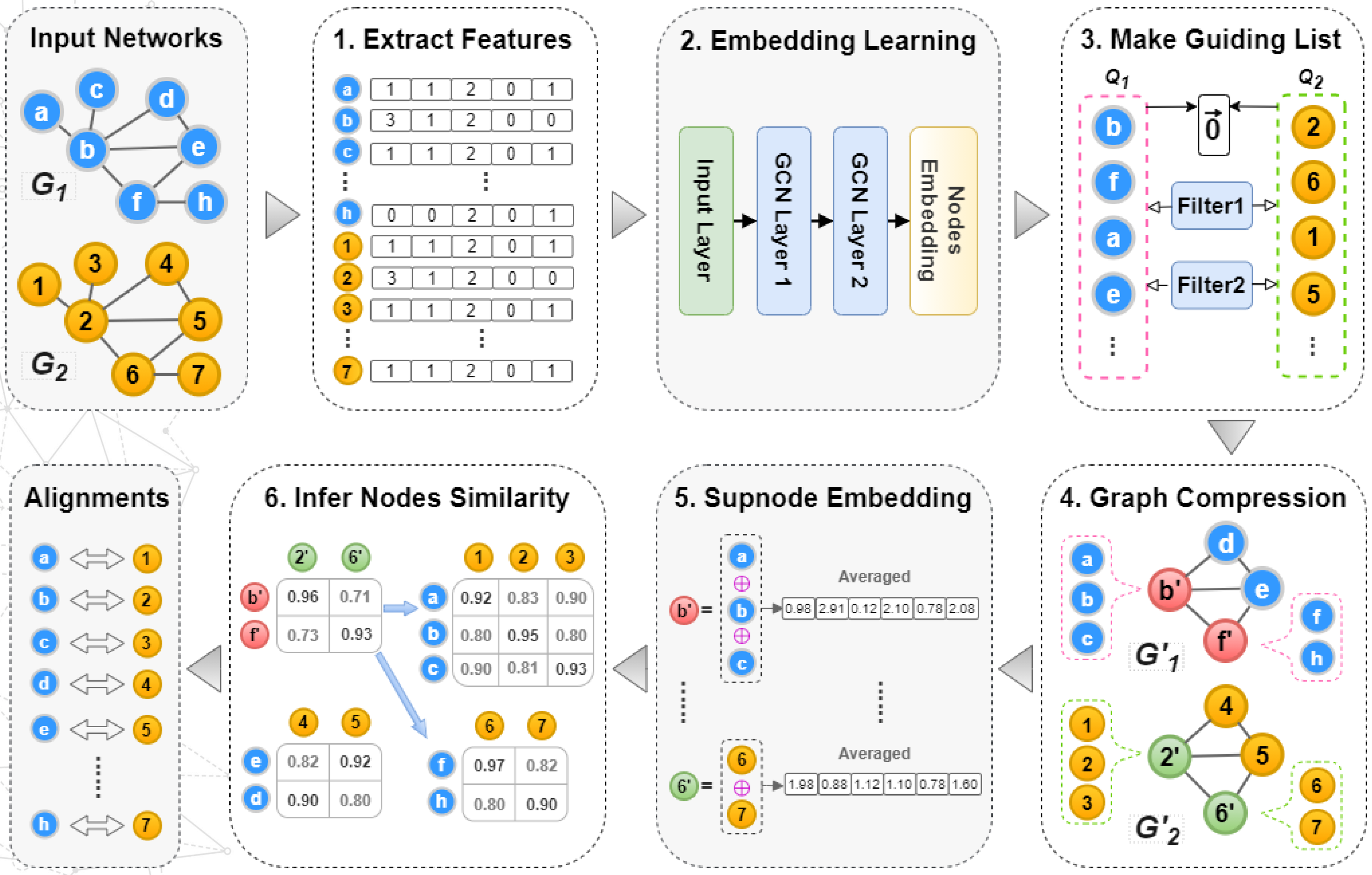
Network alignment is useful for multiple applications that require increasingly large graphs to be processed. the process of aligning every pair of nodes between relatively large networks is time-consuming and resource-intensive. In this paper, we propose a framework, called G-CREWE (Graph CompREssion With Embedding) to solve the network alignment problem. G-CREWE uses node embeddings to align the networks on two levels of resolution, a fine resolution given by the original network and a coarse resolution given by a compressed version, to achieve an efficient and effective network alignment.

Problem

Gi and *Gi ʹ* denote an input network and its compressed version, respectively. A network *Gi* (*Vi*, *Ei*, *Fi*) has a set of nodes *Vi* , a set of the edges *Ei* and the attributes *Fi* for the nodes. *Gi ʹ* (*Vi ʹ*, *Ei ʹ*, *Fi ʹ*) is the compressed version of *Gi*. Particularly, *Gi ʹ* contains a set of uncompressed original nodes *Ci ʹ* and a set of super nodes *Ui ʹ*.



Methods

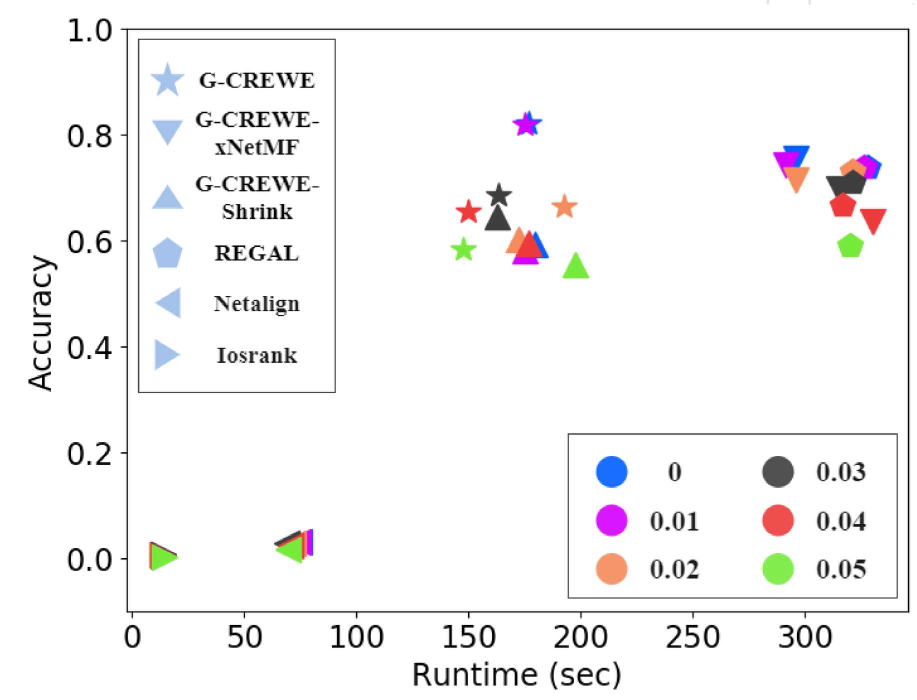


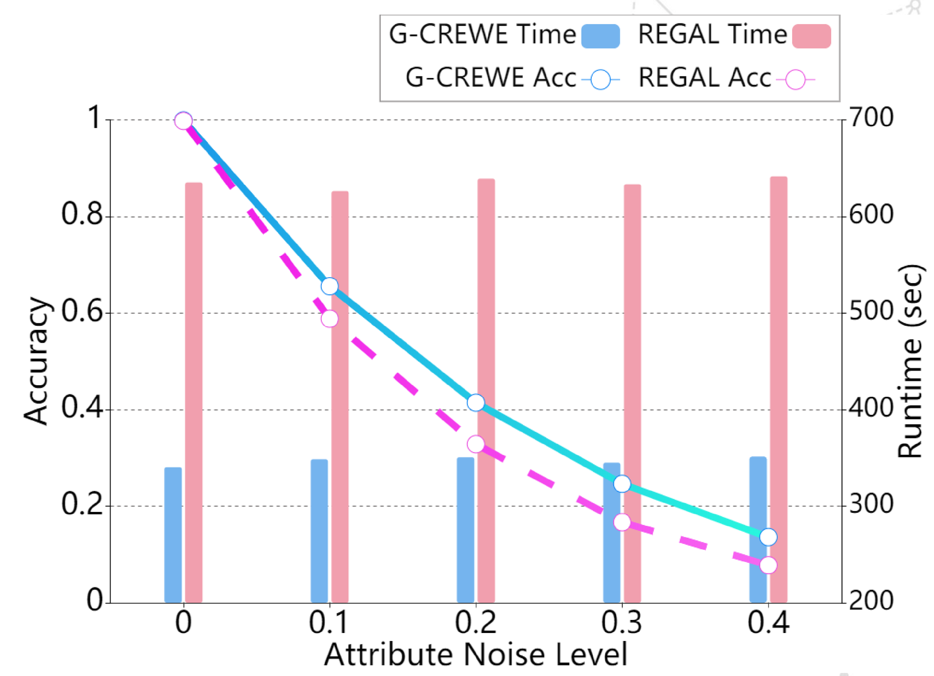
The overview of the framework. (1) Extracting the structural features of nodes in two networks with considering 2-hop neighbors. (2) 2-layer GCN is applied on these features to produce node embedding. (3) Creating guiding-lists G1 and G2 to guide the process of compression. (4) Selecting one node from G1 or G2 in order as the starting point for one compression (equals to generate a supernode) in individual network, it repeats until the compression ratio is met. (5) Calculating the embedding for super nodes via element-wise addition on their sub-nodes with averaging. (6) Inferring similarity scores between nodes.

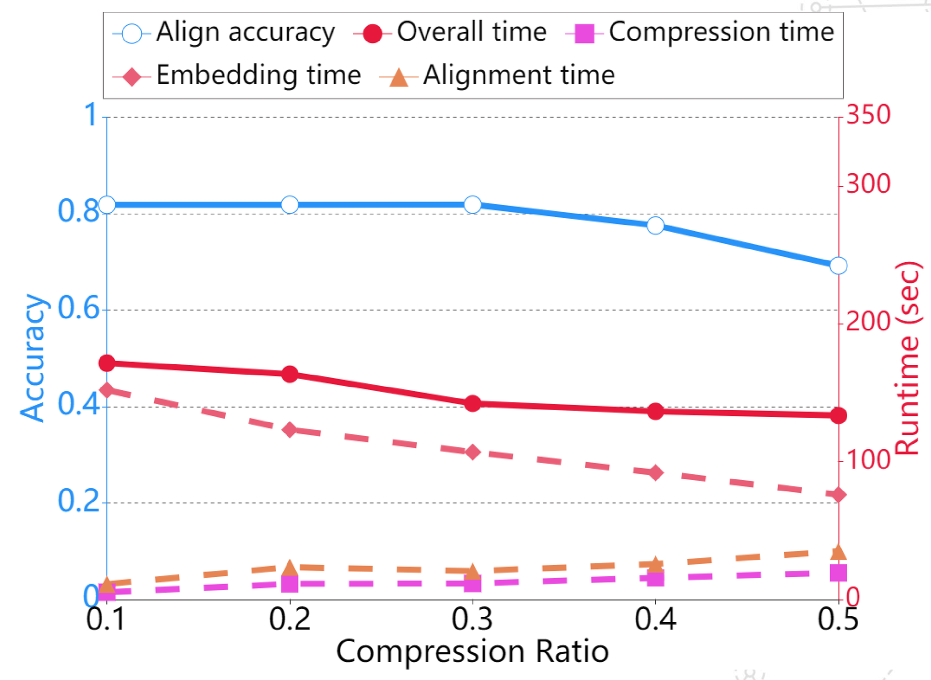
Contributions

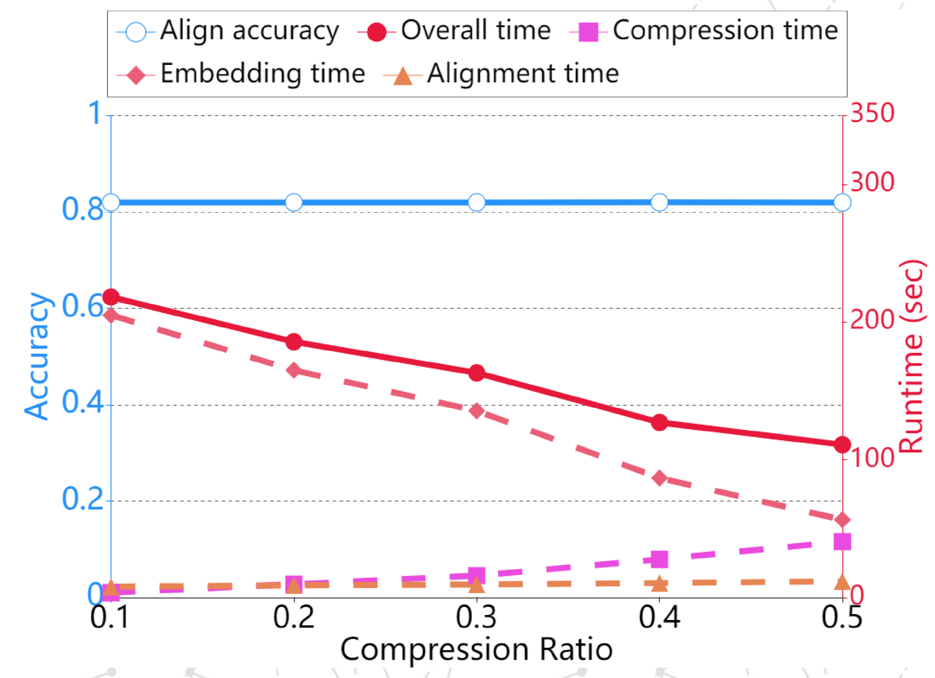
1. We define a problem for network alignment via both graph compression and embedding, to achieve a more rapid alignment for nodes in different coarsened networks while retaining topological consistency during the compression.
2. MERGE is a new compression method that preserves the topological structure of the original graph. Our approach can be applied to attributed and unattributed graphs and is unsupervised.
3. Analyze how the compression algorithm can keep the topological consistency in disjoint networks and maintain the efficiency of the alignment process.

Results









Email: Kai.qin2@rmit.edu.au

Qin, K.K., Salim, F.D., Ren, Y., Shao, W., Heimann, M. and Koutra, D., 2020, October. G-crewe: Graph compression with embedding for network alignment. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*.