GNNRank: Learning Global Rankings from Pairwise Comparisons via Directed Graph Neural Networks

Yixuan He [1] Quan Gan [2] David Wipf [2]
Gesine Reinert [1,3] Junchi Yan [4, 5] Mihai Cucuringu [1,3]

[1]Department of Statistics, University of Oxford, United Kingdom
[2]Amazon Web Services AI Shanghai Lablet, China
[3]The Alan Turing Institute, United Kingdom
[4]Department of CSE and MoE Key Lab of AI, Shanghai Jiao Tong University, China
[5]Shanghai AI Laboratory, China

Motivation

- Recovering global rankings from pairwise comparisons is an important problem with many applications, ranging from time synchronization to sports team ranking.
- Pairwise comparisons corresponding to matches in a competition can naturally be construed as edges in a directed graph (digraph), whose nodes represent competitors with an unknown rank or skill strength.
- However, existing methods addressing the rank estimation problem have thus far not utilized powerful neural network architectures to optimize ranking objectives.
- Hence, we augment a certain ranking algorithm with neural networks, in particular graph neural networks (GNN) for its intrinsic connection to the problem at hand.

Without loss of generality, we consider pairwise comparisons in a competition, which can be encoded in a directed graph (digraph) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.

- The node set $\mathcal{V}$: competitors.
- The edge set $\mathcal{E}$: pairwise comparisons.
- The adjacency matrix $\mathbf{A}$: outcomes of the matches.
- Recovering global rankings from pairwise comparisons amounts to assigning an integer R_i to each node $v_i \in \mathcal{V}$, denoting its position among competitors, where the lower the rank, the stronger the node is.
- We attempt by first learning a skill-level vector $\mathbf{r}$ for all nodes, where a higher skill value corresponds to a lower rank ordering.

SerialRank [Fogel et al., 2014]:

- First compute a certain similarity matrix $\mathbf{S}'$
- The corresponding Fiedler vector of $\mathbf{S}'$ then serves as the final ranking estimate, after a global sign reconciliation.
- While often effective in practice, SerialRank heavily depends on the quality of the similarity matrix $\mathbf{S}'$.
- To address this issue, we introduce a parameterized GNN model that allows us to compute trainable measures of similarity that are useful for subsequent ranking.

Framework Overview

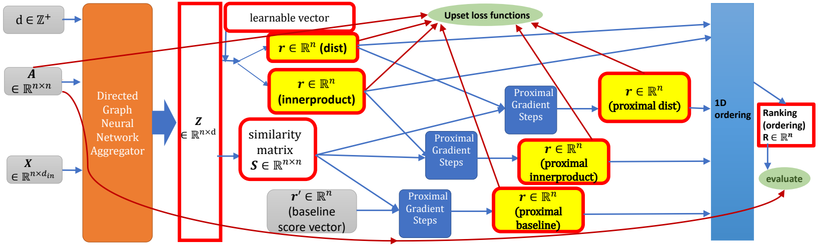

Figure: Overview of GNNRank based on directed graph neural networks and the proximal gradient steps.

Obtaining Directed Graph Embeddings

Any GNN method which is able to take into account directionality and output node embeddings could be applied; examples include:

- DIMPA by [He et al., 2021]
- The inception block model (IB) [Tong et al., 2020]
- MagNet [Zhang et al., 2021]

Denoting the final node embedding $\mathbf{Z} \in \mathbb{R}^{n \times d}$, the embedding vector $\mathbf{z}_i$ for a node v_i is $\mathbf{z}_i = (\mathbf{Z})_{(i,:)} \in \mathbb{R}^d$, the i^{th} row of $\mathbf{Z}$.

- To obtain the final ranking score, we unfold the calculation of a Fiedler vector for the graph constructed from our symmetric similarity matrix $\mathbf{S}$ with proximal gradient steps.

- From the high-dimensional embedding matrix $\mathbf{Z}$, we calculate the symmetric similarity matrix $\mathbf{S}$ with

$$\mathbf{S}_{i,j} = \exp(-|\mathbf{z}_j - \mathbf{z}_i|_2^2/(\sigma^2 d)).$$

- We then apply proximal gradient steps to approximate a Fiedler vector of $\mathbf{S}$, which serves as $\mathbf{r}$.

Main Experimental Results

For certain dense synthetic digraphs, SerialRank (which motivated our proximal gradient steps) attains leading performance, while for some other cases it fails.

We observe across all data sets that our proximal methods:

- usually outperform non-proximal variants;
- can improve on existing baseline methods when using them as initial guesses, and never perform significantly worse, hence enhance existing methods;
- do not rely on baseline methods for an initial guess, but can instead use non-proximal outcomes;
- can outperform SerialRank by unfolding its Fiedler vector calculations with a trainable similarity matrix and proximal gradient steps.

Conclusion

We have proposed a general framework based on directed graph neural networks to recover global rankings from pairwise comparisons. Future directions:

- learning a more powerful model to work for different input digraphs
- minimizing the number of upsets under constraints
- training with some supervision of ground-truth rankings
- exploring the interplay with low-rank matrix completion
- incorporating side information, in the form of node level covariates

Code: `https://github.com/SherylHYX/GNNRank`

More about me: `https://sherylhyx.github.io/`

References

Fogel, F., d'Aspremont, A., and Vojnovic, M. (2014).
Serialrank: Spectral ranking using seriation.
Advances in Neural Information Processing Systems, 27:900–908.

He, Y., Reinert, G., and Cucuringu, M. (2021).
DIGRAC: Digraph Clustering with Flow Imbalance.
arXiv preprint arXiv:2106.05194.

Tong, Z., Liang, Y., Sun, C., Li, X., Rosenblum, D., and Lim, A. (2020).
Digraph inception convolutional networks.
Advances in Neural Information Processing Systems, 33:17907–17918.

Zhang, X., He, Y., Brugnone, N., Perlmutter, M., and Hirn, M. (2021).
Magnet: A neural network for directed graphs.
Advances in Neural Information Processing Systems, 34.