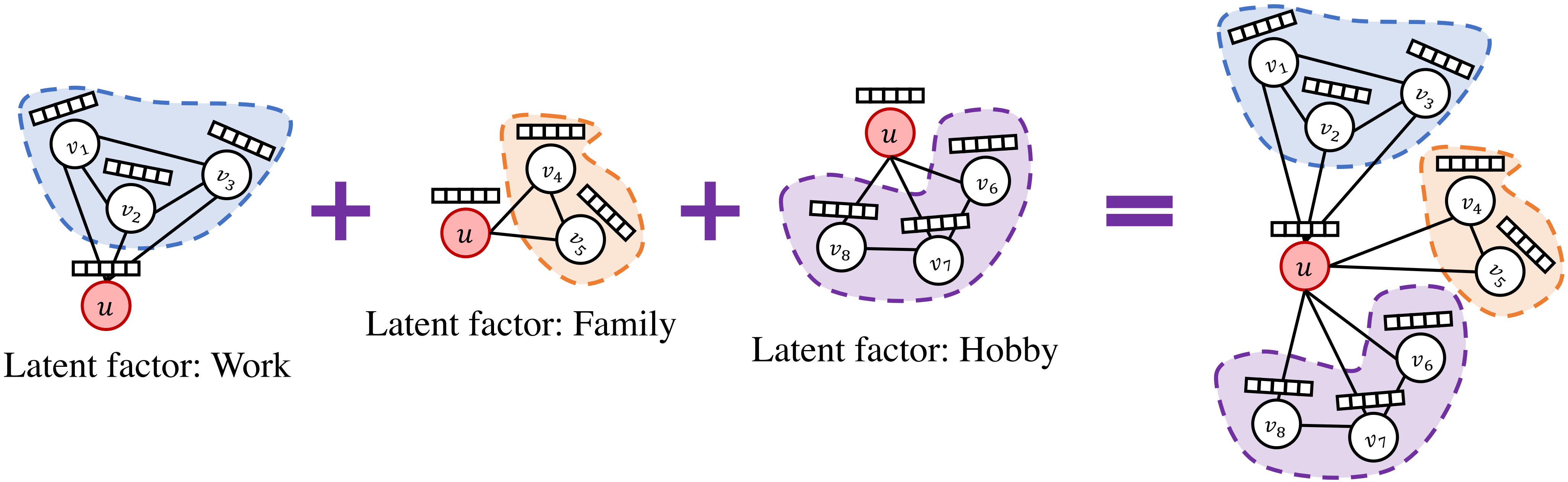
media and network lab

Disentangled Graph Convolutional Networks

Jianxin Ma, Peng Cui, Kun Kuang, Xin Wang, Wenwu Zhu
Tsinghua University

Motivation

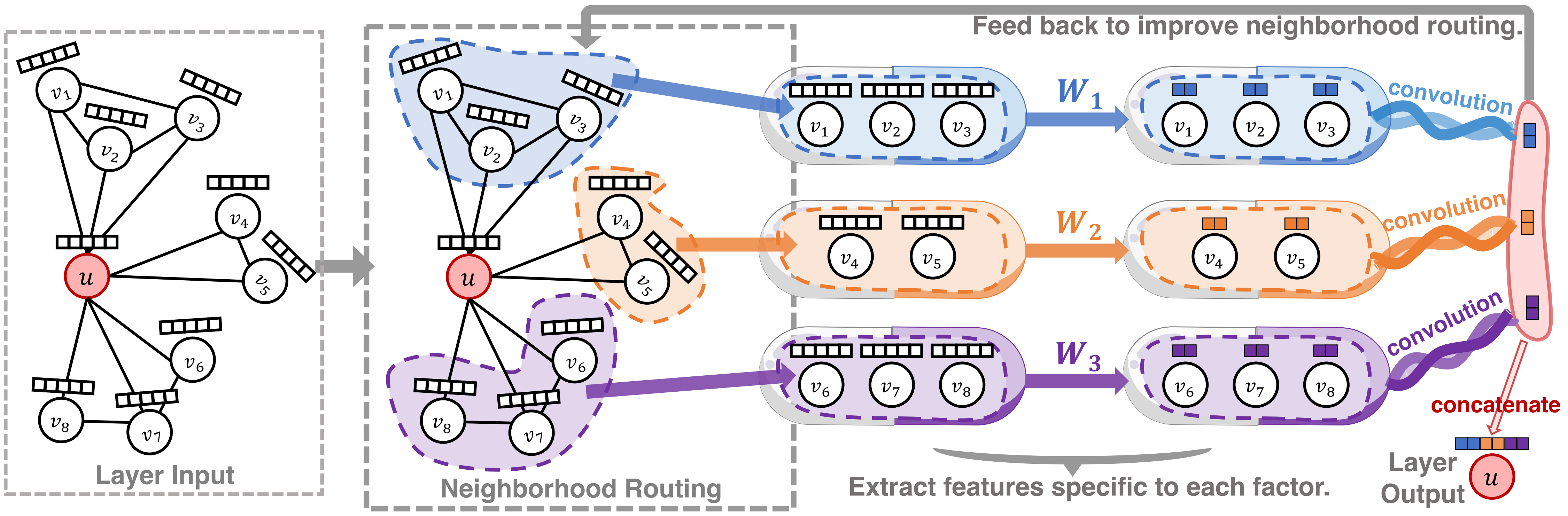
- The neighborhood of a node is formed due to *many latent factors*.



- Existing GCNs convolute the neighborhood as a whole.
  - They do not distinguish between the latent factors.
  - Their node representations are thus not *robust*, and hardly *interpretable*.

Disentangled GCNs

- *Disentangled representation learning* aims to identify and separate the underlying explanatory factors behind the observed data (Bengio et al., 2013).



- We identify the latent factors, and segment the neighborhood accordingly.
- Each segment is related with an isolated factor, and is convoluted separately.

Neighborhood Routing

- We propose *neighborhood routing*, to segment a neighborhood.
  - Dynamic & differentiable. Similar to capsule networks' dynamic routing.

- Phase I:
  - To extract factor-specific features.
  - For node $i \in \{u\} \cup \{v: (v, u) \in G\}$, and factor $k \in \{1,2, \dots, K\}$,
  - $\mathbf{z}_{i,k} = \frac{\sigma(\mathbf{W}_k^\top \mathbf{x}_i + \mathbf{b}_k)}{\left\|\sigma(\mathbf{W}_k^\top \mathbf{x}_i + \mathbf{b}_k)\right\|_2}$
  - which describes node i's aspect k.

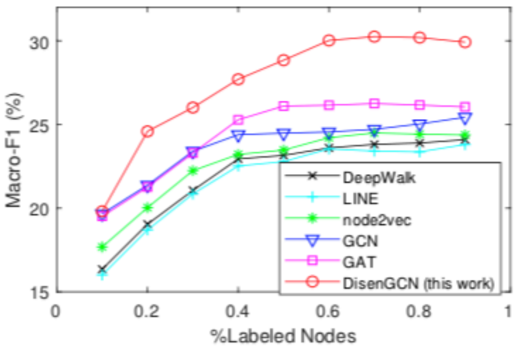
- Phase II:
  - To infer the factor that causes the link between node u and a neighbor v.
  - Initialize $\mathbf{c}_k \leftarrow \mathbf{z}_{u,k}$ for each factor k.
  - Iterate for $T \approx 5$ times,
  - $p_{v,k} \leftarrow \frac{\exp\left(\mathbf{z}_{v,k}^\top \mathbf{c}_k / \tau\right)}{\sum_{k'} \exp\left(\mathbf{z}_{v,k'}^\top \mathbf{c}_{k'} / \tau\right)}$
  - $\mathbf{c}_k \leftarrow \frac{\mathbf{z}_{u,k} + \sum_{v:\,(v,u)\in G} p_{v,k}\, \mathbf{z}_{v,k}}{\left\|\mathbf{z}_{u,k} + \sum_{v:\,(v,u)\in G} p_{v,k}\, \mathbf{z}_{v,k}\right\|_2}$
  - $\mathbf{c}_k$ describes the neighborhood's aspect k.

Intuitions & Theories

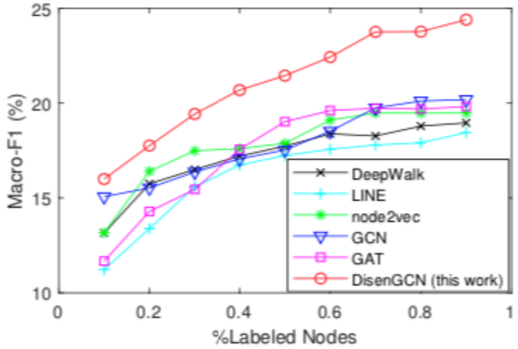
- The two intuitions behind neighborhood routing:
  - p(Factor k is the one that causes the links between node u and a segment) $\propto$ The segment contains a large number of nodes that are similar w.r.t. aspect k.
  - p(Factor k is the one that causes the link between node u and a neighbor) $\propto$ Node u and the neighbor are similar w.r.t. aspect k.

- Neighborhood routing is equivalent to an EM algorithm that performs inference under a von Mises-Fisher subspace clustering model.
  - It finds one large cluster in each of the K subspaces.

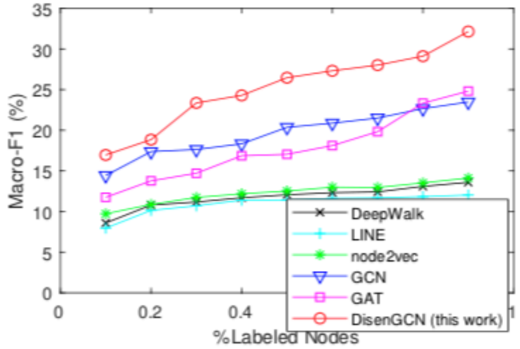
Results: Multi-label Node Classification



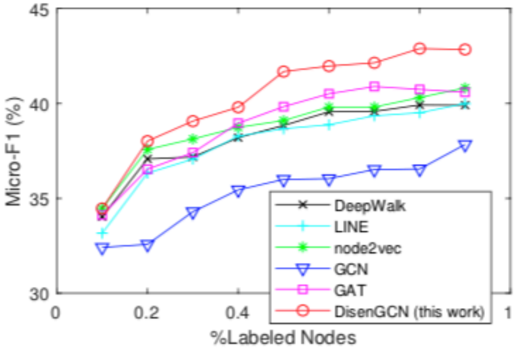
(a) Macro-F1(%), BlogCatalog.



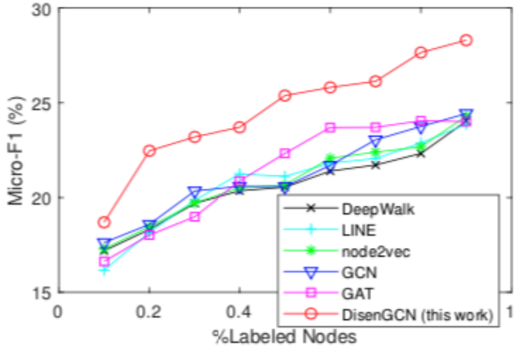
(c) Macro-F1(%), PPI.



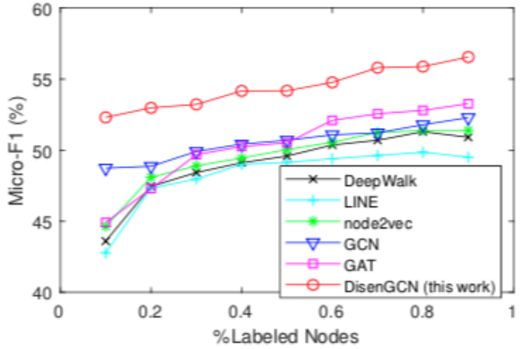
(e) Macro-F1(%), POS.



(b) Micro-F1(%), BlogCatalog.



(d) Micro-F1(%), PPI.

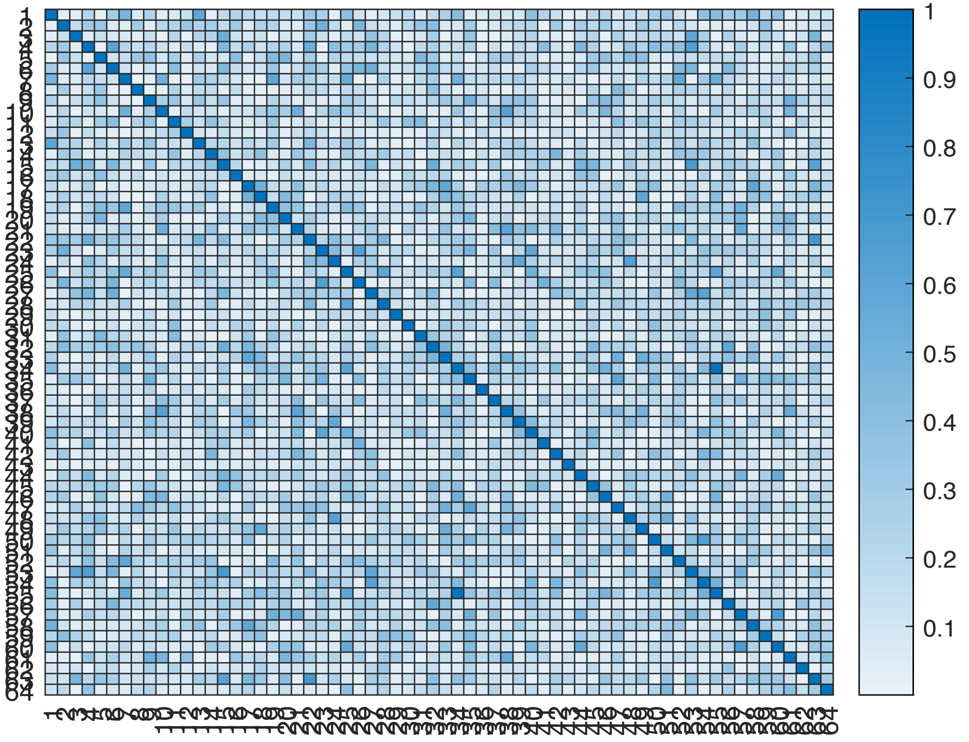


(f) Micro-F1(%), POS.

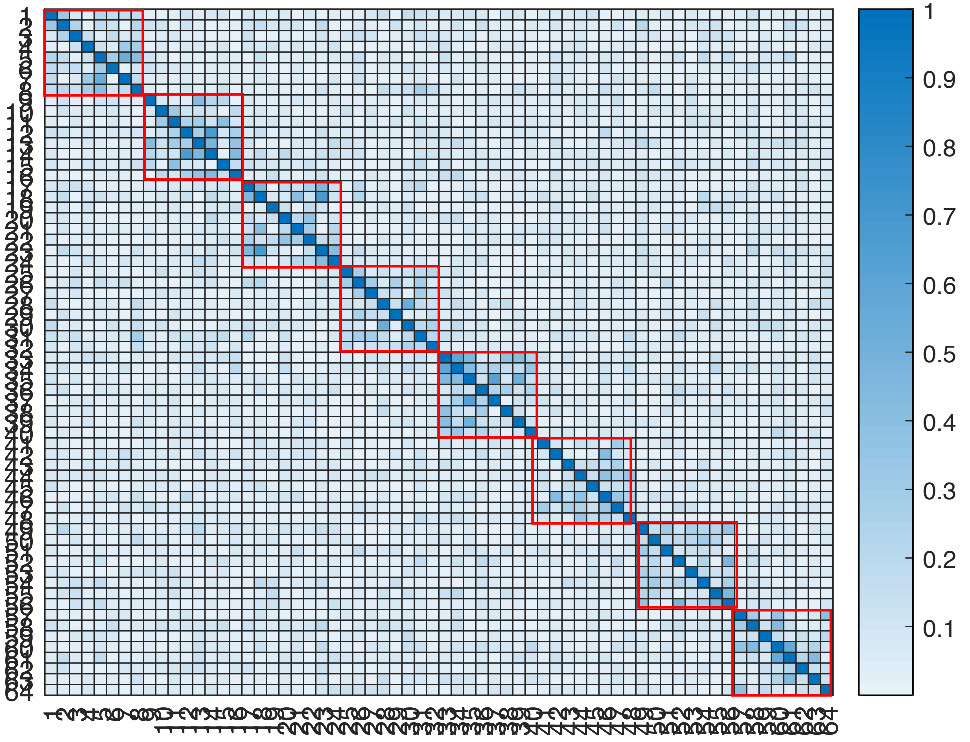
Figure 2. Macro-F1 and Micro-F1 scores on the multi-label classification tasks. Our approach consistently outperforms the best performing baselines by a large margin, reaching 10% to 20% relative improvement in most cases.

Results: Disentangled Node Representations

- Correlations between the 64 dimensions, on a graph with eight factors.



(a) GCN.



(b) DisenGCN (this work).