

GNN is a counter?

Revisiting GNN for Question Answering

Kuan Wang, Yuyu Zhang, Diyi Yang, Le Song, Tao Qin



Background

Blue read material outside of his **comfort zone** because he wanted to gain what?

A. **new perspective** B. **understanding** C. **entertained** D. **hunger** E. **tired eyes**

Background

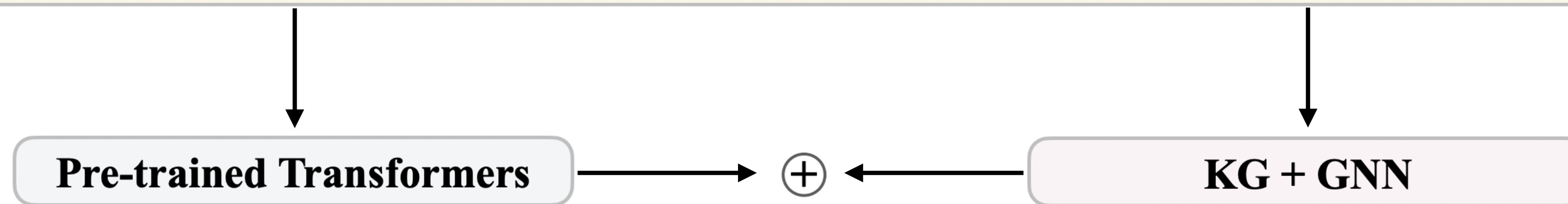
Blue read material outside of his **comfort zone** because he wanted to gain what?
A. **new perspective** B. **understanding** C. **entertained** D. **hunger** E. **tired eyes**



Pre-trained Transformers

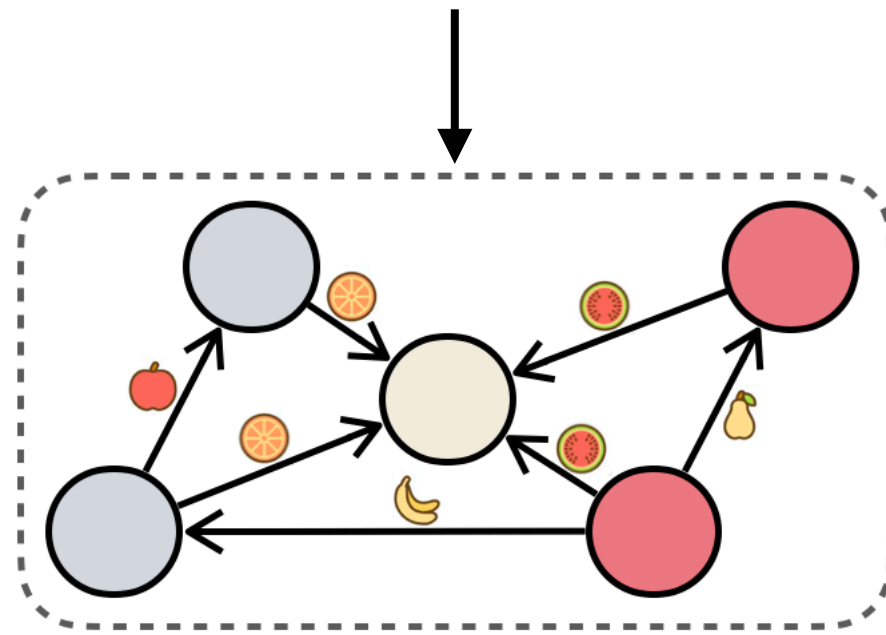
Background

Blue read material outside of his **comfort zone** because he wanted to gain what?
A. **new perspective** B. **understanding** C. **entertained** D. **hunger** E. **tired eyes**



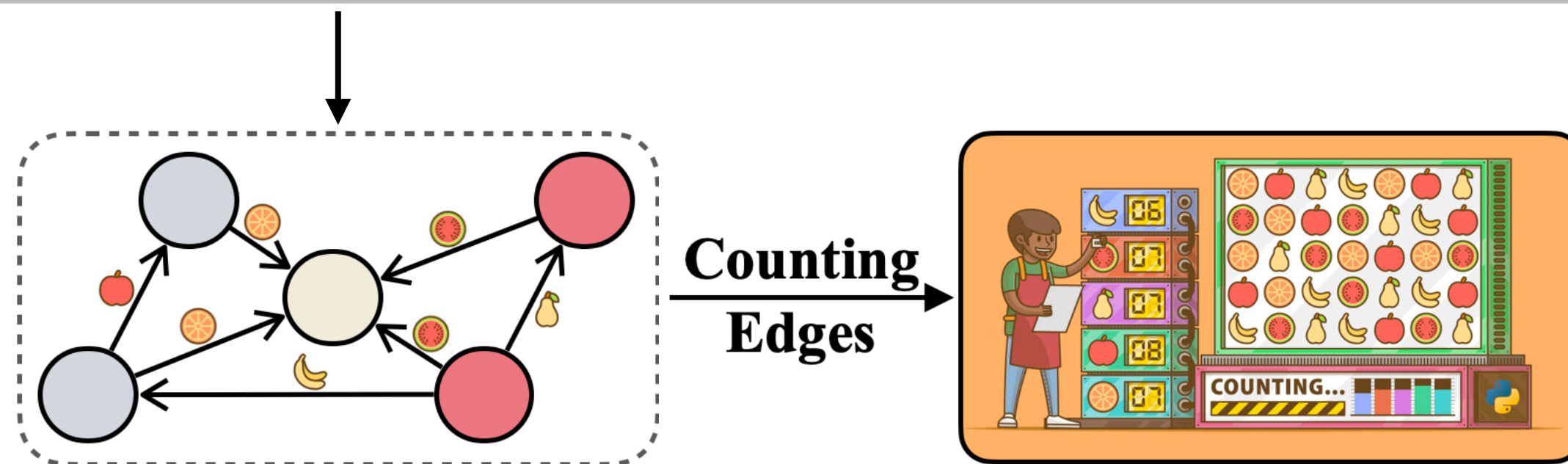
Background

Blue read material outside of his **comfort zone** because he wanted to gain what?
A. **new perspective** B. **understanding** C. **entertained** D. **hunger** E. **tired eyes**



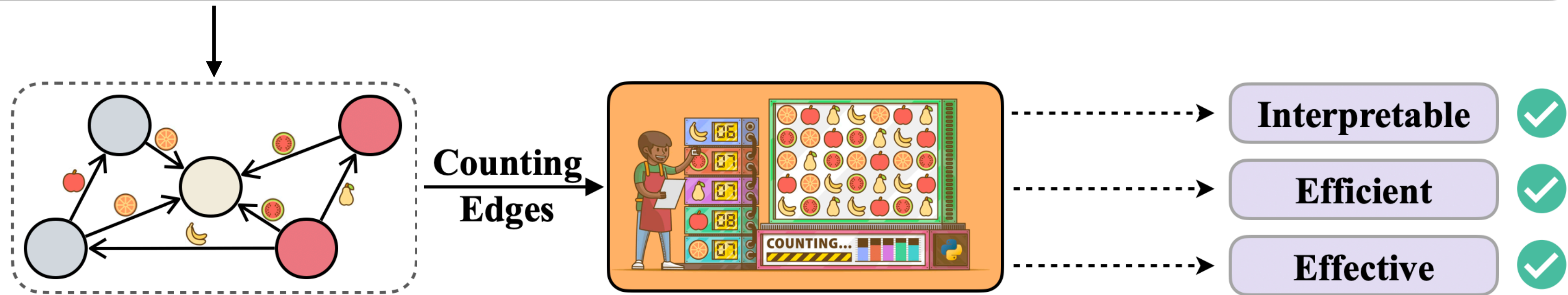
Background

Blue read material outside of his **comfort zone** because he wanted to gain what?
A. **new perspective** B. **understanding** C. **entertained** D. **hunger** E. **tired eyes**



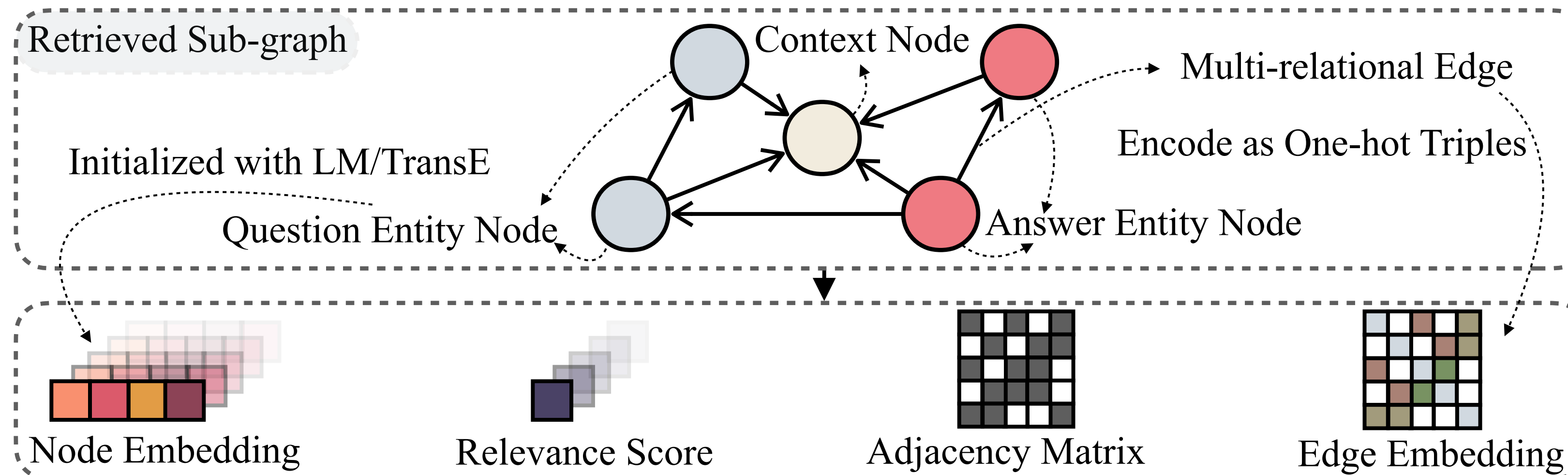
Background

Blue read material outside of his **comfort zone** because he wanted to gain what?
A. **new perspective** B. **understanding** C. **entertained** D. **hunger** E. **tired eyes**



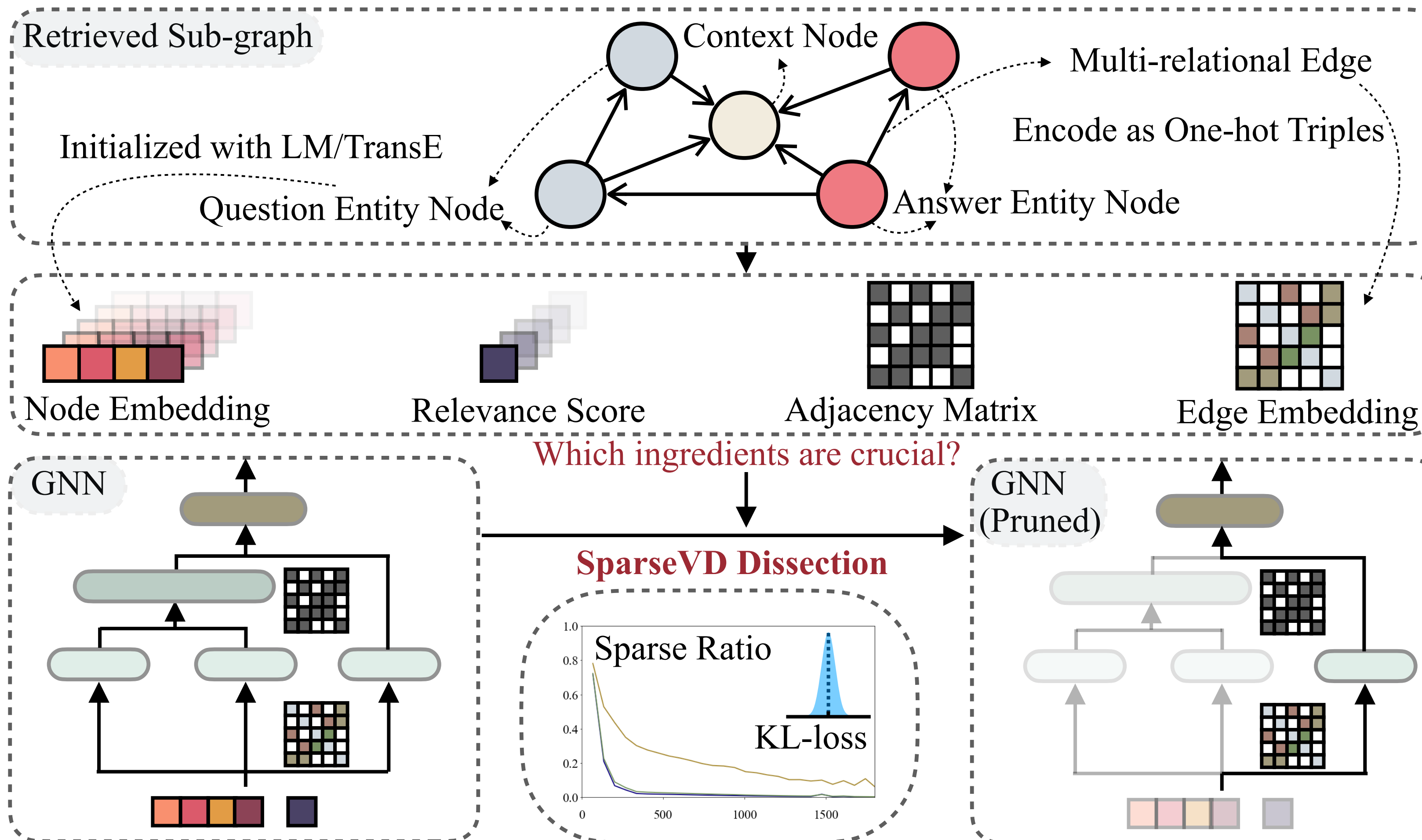
Analysis

Ingredients of Graph input



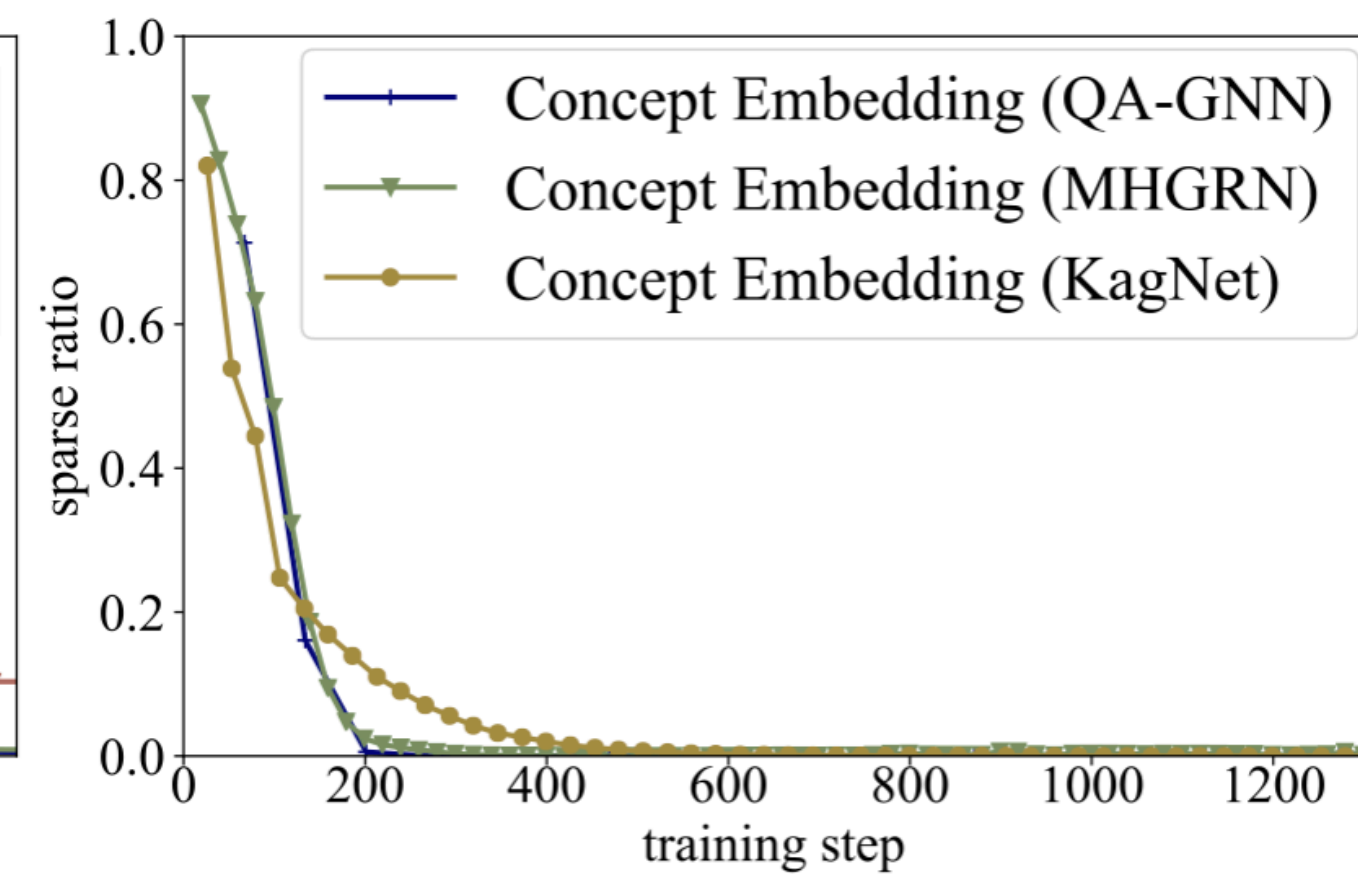
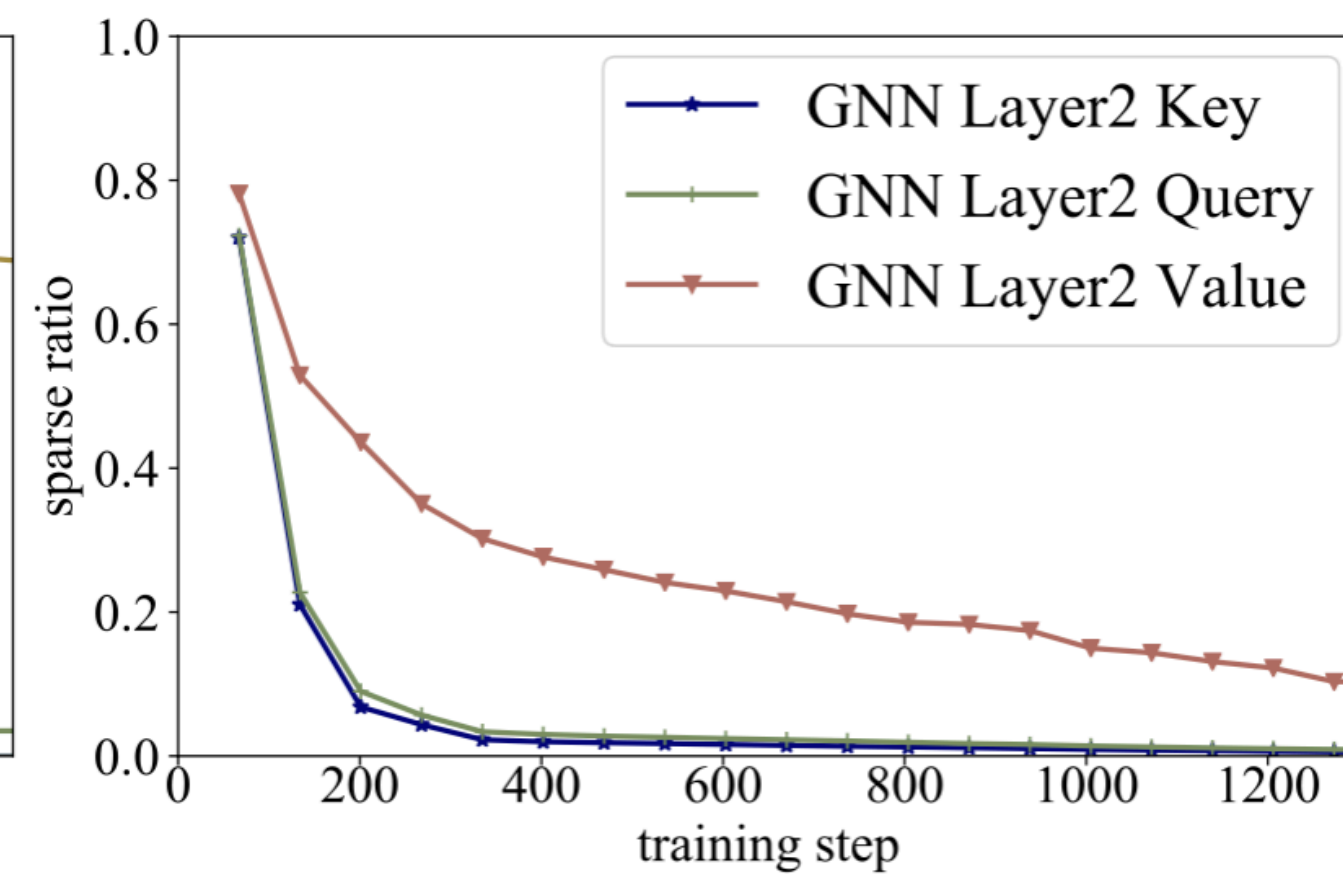
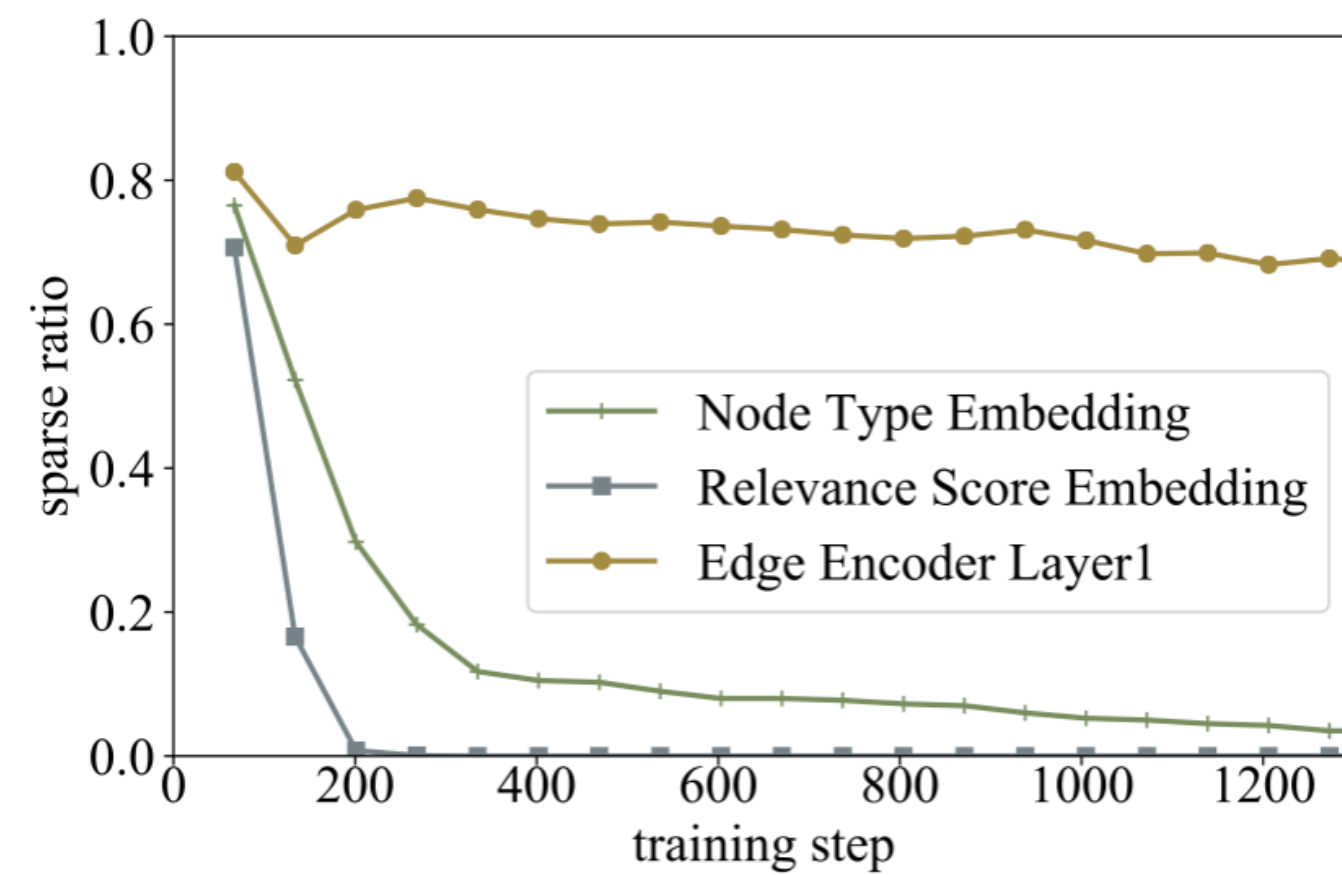
Analysis

Pruning method SparseVD as the dissection tool

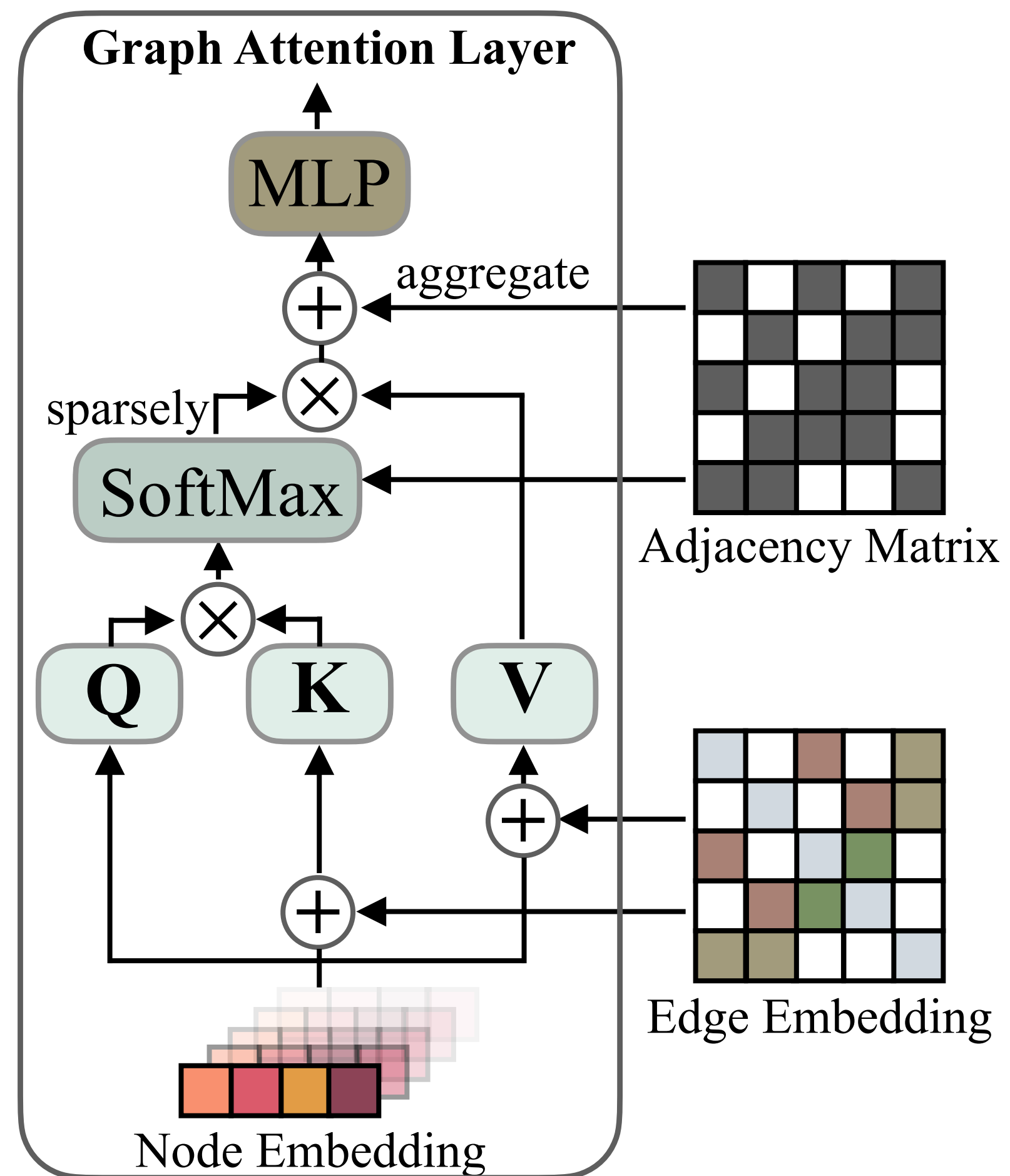


Analysis

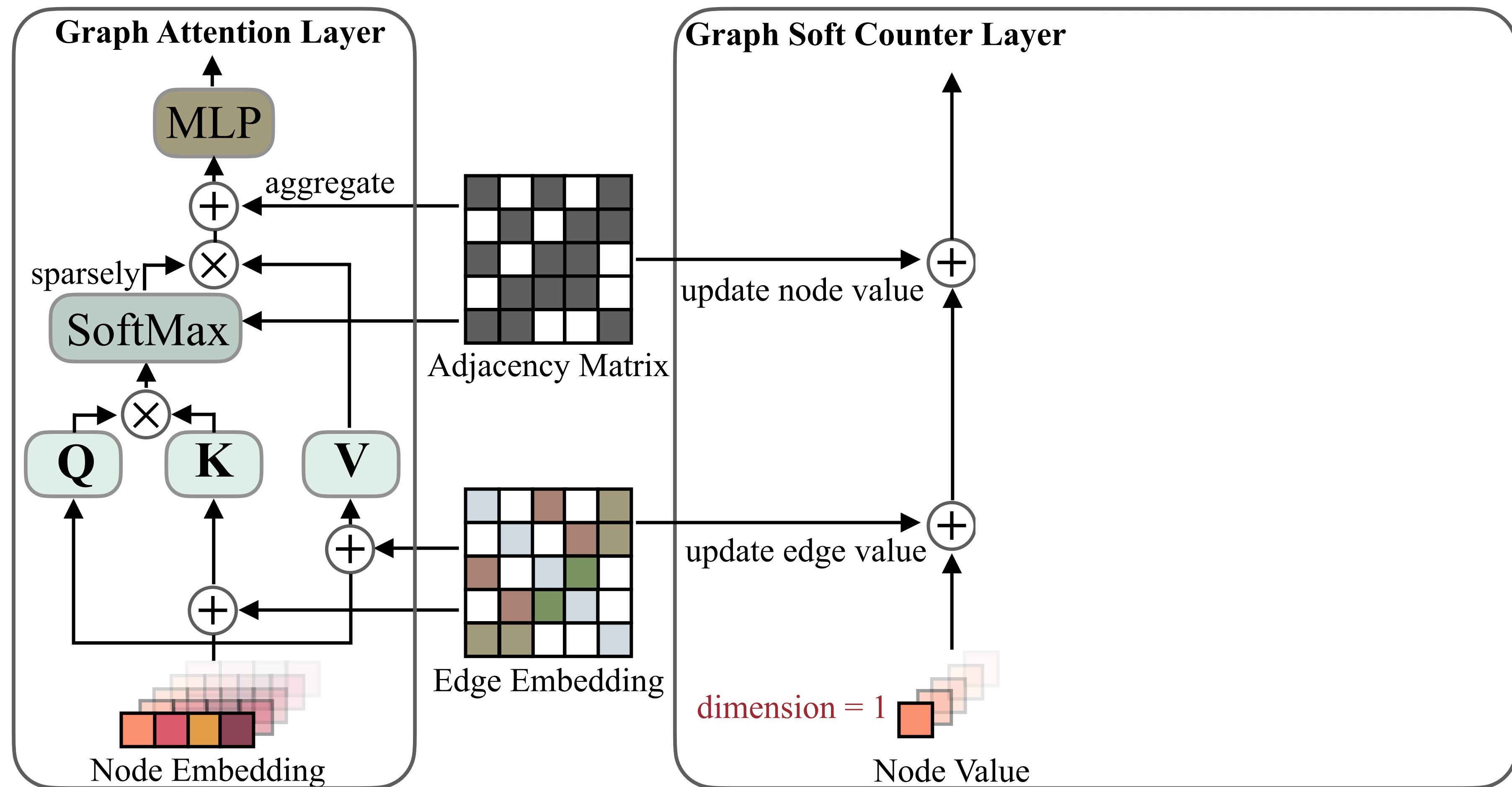
The different sparse ratios show different importance



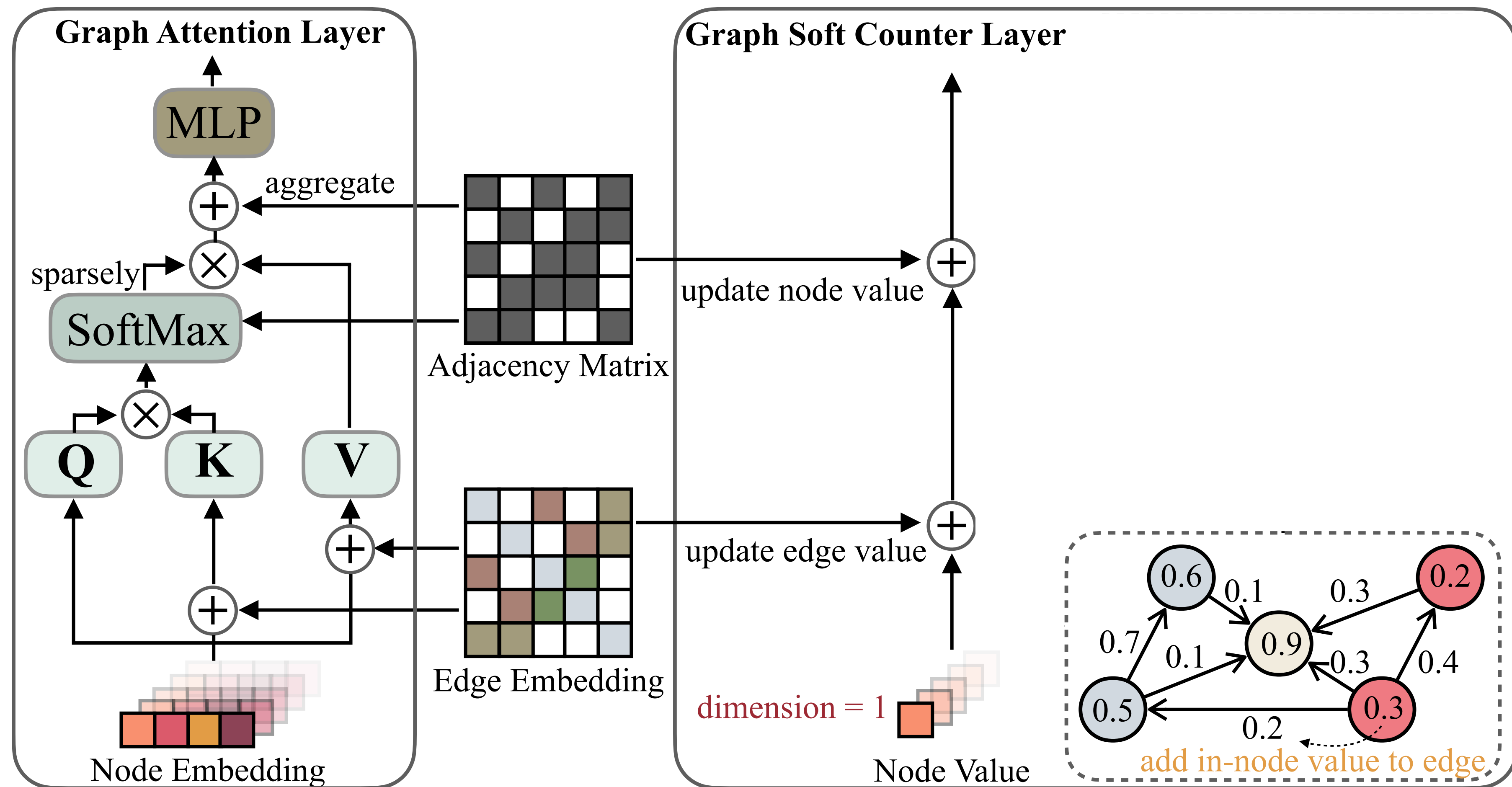
Graph Soft Counter



Graph Soft Counter

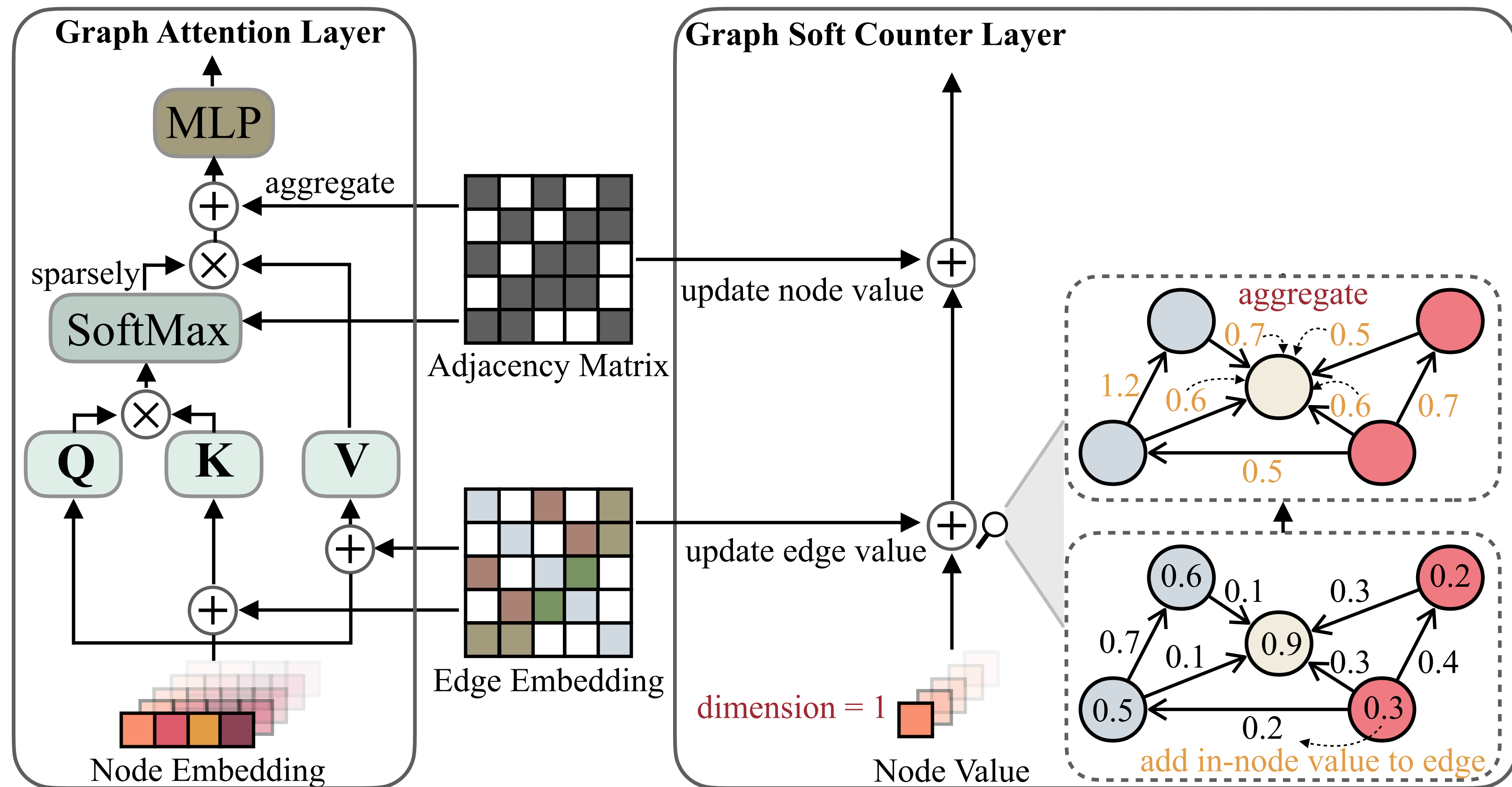


Graph Soft Counter



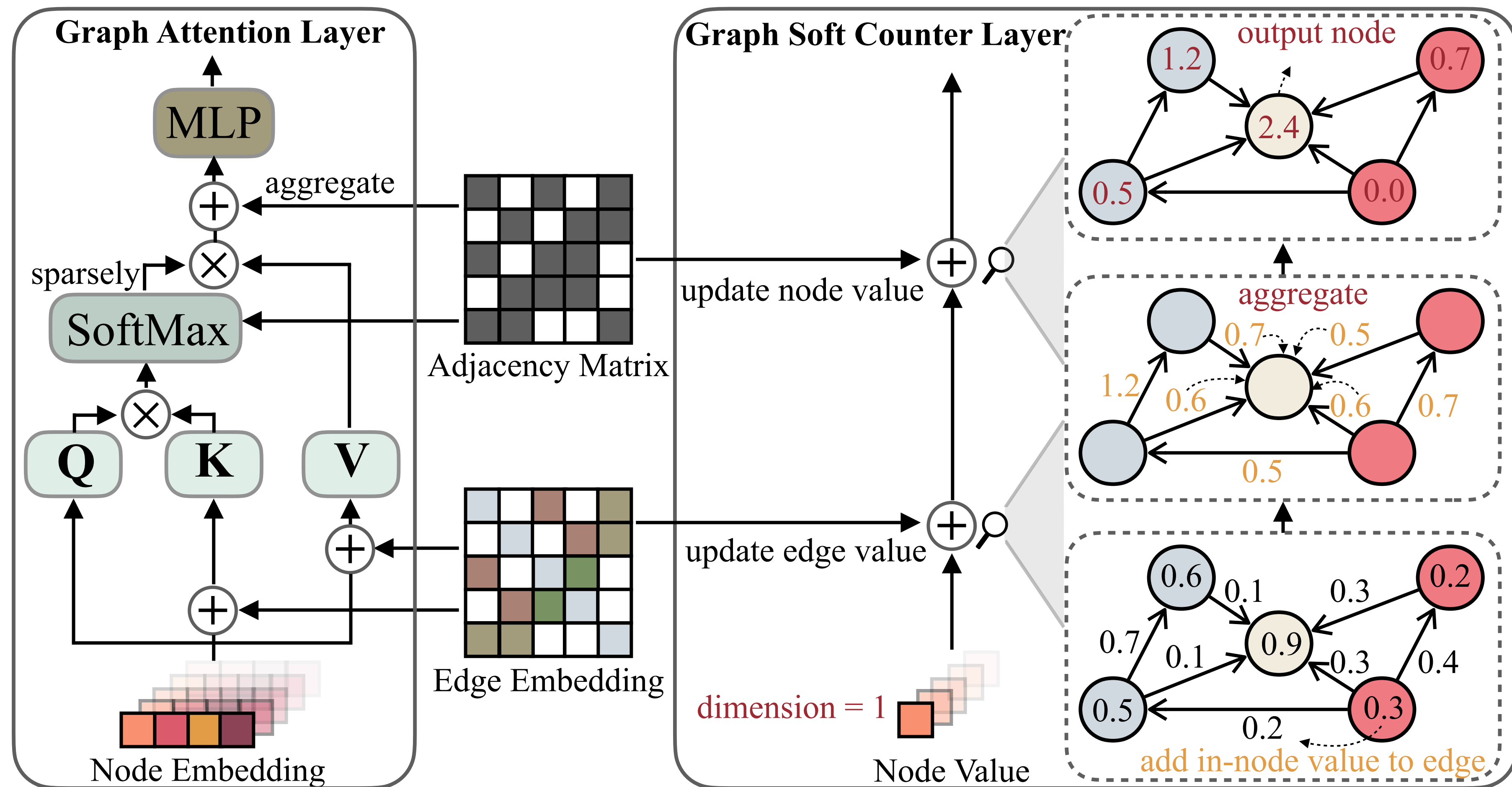
Graph Soft Counter

The first step is updating node value to edge



Graph Soft Counter

The second step is aggregate edge value to node



Results

GSC outperforms other GNNs on the CommonsenseQA leaderboards

Methods	Test
RoBERTa (Liu et al., 2019)	72.1
RoBERTa + FreeLB (Zhu et al., 2019) (ensemble)	73.1
RoBERTa + HyKAS (Ma et al., 2019)	73.2
RoBERTa + KE (ensemble)	73.3
RoBERTa + KEDGN (ensemble)	74.4
XLNet + GraphReason (Lv et al., 2020)	75.3
RoBERTa + MHGRN (Feng et al., 2020)	75.4
ALBERT + PG (Wang et al., 2020)	75.6
RoBERTa + QA-GNN (Yasunaga et al., 2021)	76.1
ALBERT (Lan et al., 2019) (ensemble)	76.5
UnifiedQA (11B)* (Khashabi et al., 2020)	79.1
RoBERTa + GSC (Ours)	76.4

Table 5: Test accuracy on *CommonsenseQA*’s official leaderboard. The previous top system, UnifiedQA (11B params) is 30x larger than our model.

Results

GSC outperforms other GNNs on the OpenBookQA leaderboards

Methods	Test
Careful Selection (Banerjee et al., 2019)	72.0
AristoRoBERTa	77.8
KF + SIR (Banerjee & Baral, 2020)	80.0
AristoRoBERTa + PG (Wang et al., 2020)	80.2
AristoRoBERTa + MHGRN (Feng et al., 2020)	80.6
ALBERT + KB	81.0
AristoRoBERTa + QA-GNN	82.8
T5* (Raffel et al., 2020)	83.2
UnifiedQA(11B)* (Khashabi et al., 2020)	87.2
AristoRoBERTa + GSC (Ours)	87.4

Table 7: Test accuracy on *OpenBookQA* leaderboard. All listed methods use the provided science facts as an additional input to the language context. The previous top 2 systems, UnifiedQA (11B params) and T5 (3B params) are 30x and 8x larger than our model.

Results

GSC outperforms other GNNs on the MetaQA dataset

Methods	1-hop	2-hop	3-hop
KV-Mem (Miller et al., 2016)	96.2	82.7	48.9
VRN (Zhang et al., 2018)	97.5	89.9	62.5
GraftNet (Sun et al., 2018)	97.0	94.8	77.7
GSC (ours)	97.6	99.8	76.8

Table 9: Our GSC achieves comparable performance with the baselines on the MetaQA dataset, which indicates that our observations and hypotheses keep consistent multi-hop QA.

Takeaway

- The existing complex GNN modules may just perform some simple reasoning such as counting in knowledge-aware reasoning.
- Initial node embedding of GNN is not used in the QA reasoning process, It remains a challenging open problem to build comprehensive reasoning modules for knowledge-powered QA.
- Opensource Code: <https://github.com/kuan-wang/graph-soft-counter>