Diffusion-based Negative Sampling on Graphs for Link Prediction

Trung-Kien Nguyen and **Yuan Fang**

In Proceedings of the 2024 ACM Web Conference, May 13-17, 2024

Outline

- **Problem & related work**
- Proposed model: DMNS
- Experiments
- Conclusions

Problem

- Link prediction: fundamental task with many applications

  social networks, recommendation, knowledge graph completion, etc.

- Modern approaches: Contrastive learning
  - Aims to learn robust node representations
  - Requires positive and negative samples for a given query node
  - Negative sampling: huge search space and highly false negatives

→ Negative sampling for contrastive link prediction on graph

Challenges & Related Work

1. *How to flexibly model and control the quality of negative nodes?*
   – Heuristics [1,2,3] or automatic generative [4,5] designs are inflexible

   → Multi-level negative sampling strategy

2. *How do we find sufficient negative examples of variable hardness?*
   – Most negative sampling approaches [1,2,3,4,6] are limited from observed graphs

   → Diffusion models: naturally generate multi-level samples at different steps

[1] Mikolov et al 2013. Distributed representations of words and phrases and their compositionality. Neurips.
[2] Zhang et al. 2013. Optimizing top-n collaborative filtering via dynamic negative item sampling. SIGIR.
[3] Yang et al. 2020. Understanding negative sampling in graph representation learning. KDD..
[4] Wang et al. 2018. Graphgan: Graph representation learning with generative adversarial nets. AAAI.
[5] Pan et al. 2018. Adversarially regularized graph autoencoder for graph embedding. IJCAI.
[6] Zhang et al. 2018. Link Prediction Based on Graph Neural Networks. Neurips.

Outline

- Problem & related work
- **Proposed model: DMNS**
- Experiments
- Conclusions

DMNS: Overall Framework

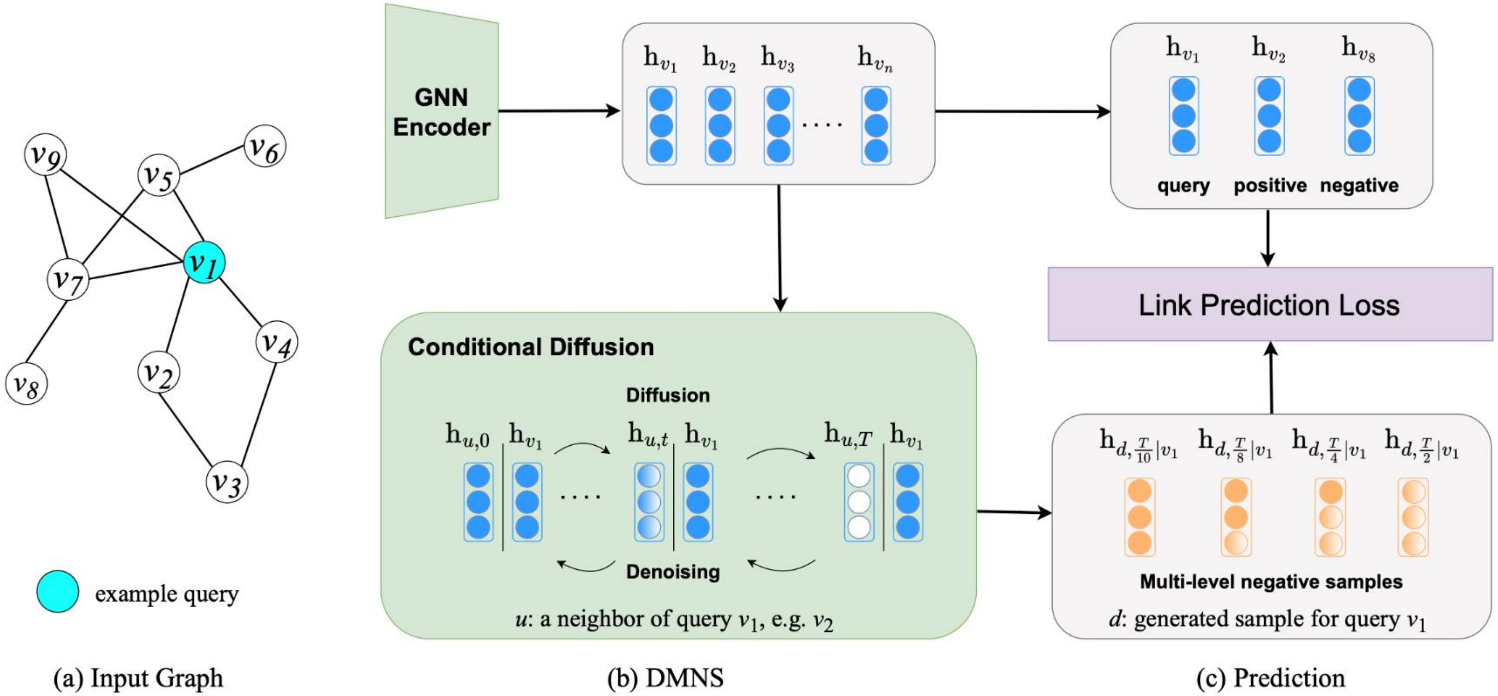


Figure 1: Overall framework of DMNS.

DMNS

GNN Encoder

$$\mathbf{h}_v^l = \sigma\left(\textsc{Aggr}(\mathbf{h}_v^{l-1}, \{\mathbf{h}_i^{l-1} : i \in \mathcal{N}_v\}; \omega^l)\right)$$

Conditional Diffusion

- Forward process

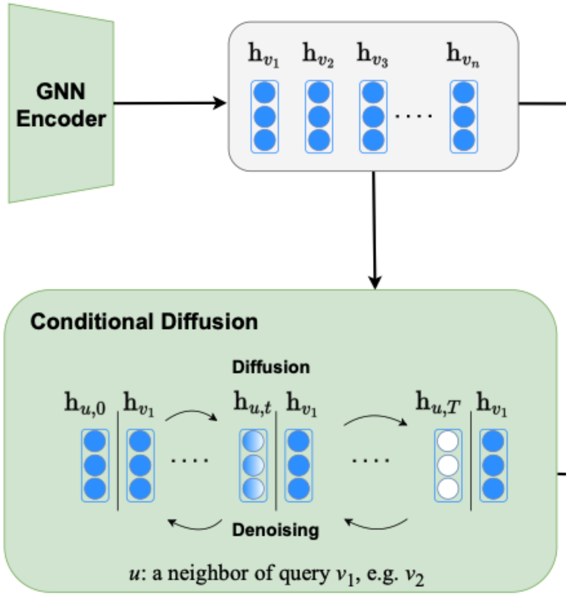
$$\mathbf{h}_{u,t} = \sqrt{\bar{\alpha}_t}\mathbf{h}_u + \sqrt{1-\bar{\alpha}_t}\epsilon_t, \quad \forall u \in \mathcal{N}_v, \quad \epsilon_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

- Reverse process

$$\epsilon_{t,\theta|v} = (\gamma + 1) \odot \mathbf{h}_{u,t} + \eta,$$

$$\gamma = \mathrm{FCL}(\mathbf{t} + \mathbf{h}_v; \theta_\gamma), \quad \eta = \mathrm{FCL}(\mathbf{t} + \mathbf{h}_v; \theta_\eta), \quad [\mathbf{t}]_{2i} = \sin(t/10000^{\frac{2i}{d_h}})$$

$$[\mathbf{t}]_{2i+1} = \cos(t/10000^{\frac{2i}{d_h}})$$



(b) DMNS

Overall Loss

- Diffusion Loss

$$\mathcal{L}_D = \|\epsilon_t - \epsilon_{t,\theta|v}\|^2$$

- Multi-level Negative Sampling

$$\mathbf{h}_{d,T|v} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}),$$

$$\mathbf{h}_{d,t-1|v} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{h}_{d,t|v} - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_{t,\theta|v}\right) + \sigma_t \mathbf{z},$$

- Link Prediction Loss

$$\mathcal{L} = -\log\sigma(\mathbf{h}_v^\top \mathbf{h}_u) - \log\sigma(-\mathbf{h}_v^\top \mathbf{h}_{u'})$$
$$- \textstyle\sum_{d_i \in D_v} w_i \log\sigma(-\mathbf{h}_v^\top \mathbf{h}_{d_i}))$$

$$(v, u, u', D_v),$$

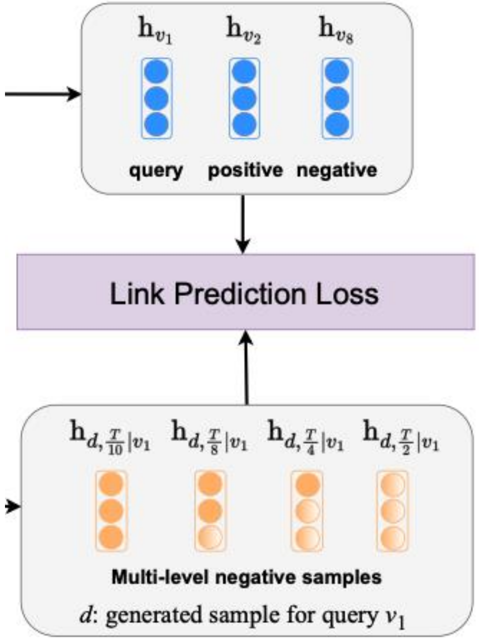
query node

positive node

real negative node

DMNS negative sets:

$D_v = \{\mathbf{h}_{d,t|v} : t = \frac{T}{10}, \frac{T}{8}, \frac{T}{4}, \frac{T}{2}\}$



(c) Prediction

Outline

- Problem & related work
- Proposed model: DMNS
- **Experiments**
- Conclusions

Experimental Setup

Datasets	Nodes	Edges	Features	Property
Cora	2708	5429	1433	homophilous
Citeseer	3327	4732	3703	homophilous
Coauthor-CS	18333	163788	6805	homophilous
Actor	7600	30019	932	heterophilous

Baselines

Classic GNNs
- GCN [1]
- GAT [2]
- SAGE [3]

Heuristic NS
- PNS [4]
- DNS [5]
- MCNS [6]

Generative NS
- GraphGAN [7]
- ARGVA [8]
- KBGAN [9]

Subgraph-based GNNs
- SEAL [10]
- ScaLed [11]

[1] Kipf et al. 2017. Semi-supervised classification with graph convolutional networks. ICLR.
[2] Veličković et al. 2018. Graph attention networks. ICLR.
[3] Hamilton et al. 2017. Inductive representation learning on large graphs. NeurIPS.
[4] Mikotov et al. 2013. Distributed representations of words and phrases and their compositionality. NeurIPS.
[5] Zhang et el. 2013. Optimizing top-n collaborative filtering via dynamic negative item sampling. SIGIR.
[6] Yang et al. 2020. Understanding negative sampling in graph representation learning. KDD.
[7] Wang et al. 2018. Graphgan: Graph representation learning with generative adversarial nets. AAAI.
[8] Pan et al. 2018. Adversarially regularized graph autoencoder for graph embedding. IJCAI.
[9] Cai et al. 2017. Kbgan: Adversarial learning for knowledge graph embeddings. ACL.
[10] Zhang et al. 2018. Link prediction based on graph neural networks. NeurIPS.
[11] Louis et al. 2022. Sampling Enclosing Subgraphs for Link Prediction. CIKM.

Link Prediction

Table 2: Evaluation of link prediction against baselines using GCN as the base encoder.

Methods	Cora		Citeseer		Coauthor-CS		Actor	
	MAP	NDCG	MAP	NDCG	MAP	NDCG	MAP	NDCG
GCN	.742 ± .003	.805 ± .003	.735 ± .011	.799 ± .008	.823 ± .004	.867 ± .003	.521 ± .004	.634 ± .003
GVAE	<u>.783</u> ± .003	<u>.835</u> ± .002	.743 ± .004	.805 ± .003	.843 ± .011	.882 ± .008	<u>.587</u> ± .004	<u>.684</u> ± .003
PNS	.730 ± .008	.795 ± .006	.748 ± .006	.809 ± .005	.817 ± .004	.863 ± .003	.517 ± .006	.631 ± .006
DNS	.735 ± .007	.799 ± .005	<u>.777</u> ± .005	<u>.831</u> ± .004	.845 ± .003	.883 ± .002	.558 ± .006	.663 ± .005
MCNS	.756 ± .004	.815 ± .003	.750 ± .006	.810 ± .004	.824 ± .004	.868 ± .004	.555 ± .005	.659 ± .004
GraphGAN	.739 ± .003	.802 ± .002	.740 ± .011	.803 ± .008	.818 ± .007	.863 ± .005	.534 ± .007	.644 ± .005
ARVGA	.732 ± .011	.797 ± .009	.689 ± .005	.763 ± .004	.811 ± .003	.858 ± .002	.526 ± .012	.638 ± .009
KBGAN	.615 ± .004	.705 ± .003	.568 ± .006	.668 ± .005	<u>.852</u> ± .002	<u>.888</u> ± .002	.472 ± .003	.596 ± .002
SEAL	.751 ± .007	.812 ± .005	.718 ± .002	.784 ± .002	.850 ± .001	.886 ± .001	.536 ± .001	.641 ± .001
ScaLed	.676 ± .004	.752 ± .003	.630 ± .004	.712 ± .003	.828 ± .001	.869 ± .001	.459 ± .001	.558 ± .001
DMNS	**.793** ± .003	**.844** ± .002	**.790** ± .004	**.841** ± .003	**.871** ± .002	**.903** ± .001	**.600** ± .002	**.696** ± .002

*Best is **bolded** and runner-up <u>underlined</u>.

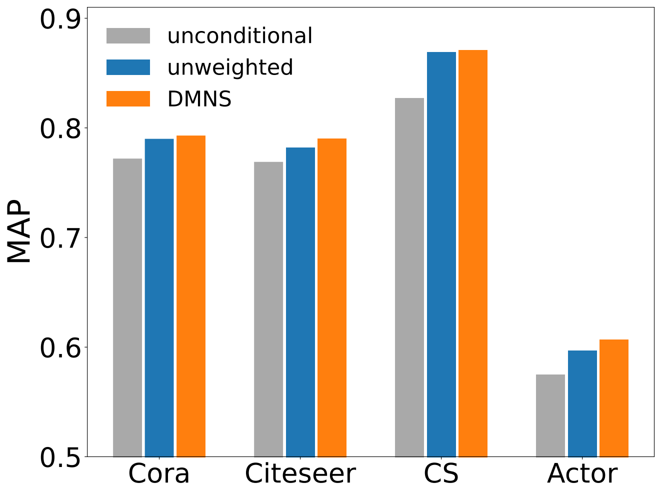
Link Prediction

Table 3: Evaluation of link prediction on DMNS with various base encoders.

Methods	Cora		Citeseer		Coauthor-CS		Actor	
	MAP	NDCG	MAP	NDCG	MAP	NDCG	MAP	NDCG
GAT	.766 ± .006	.824 ± .004	.767 ± .007	.763 ± .062	.833 ± .003	.874 ± .002	.479 ± .004	.603 ± .003
DMNS-GAT	**.813** ± .004	**.859** ± .003	**.788** ± .007	**.840** ± .006	**.851** ± .002	**.889** ± .002	**.573** ± .007	**.675** ± .005
SAGE	.598 ± .014	.668 ± .013	.622 ± .012	.713 ± .009	.768 ± .005	.826 ± .004	.486 ± .004	.604 ± .003
DMNS-SAGE	**.700** ± .007	**.773** ± .005	**.669** ± .013	**.749** ± .010	**.843** ± .004	**.883** ± .003	**.582** ± .017	**.682** ± .013

- DMNS improves performance of various base GNN encoders, demonstrating its flexibility.

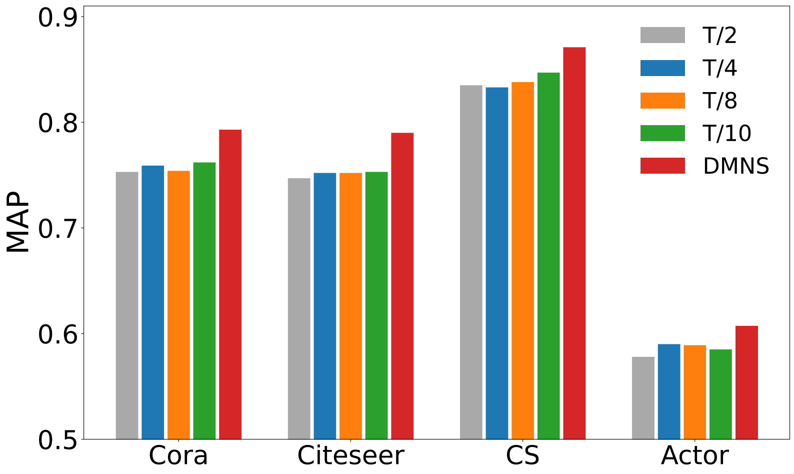
Ablation Studies



(a) On model design

Performance drops on:

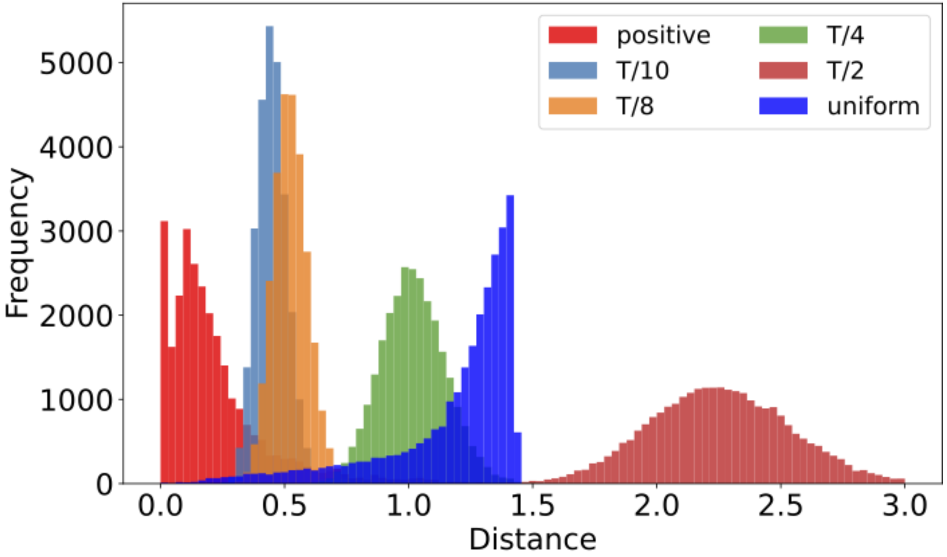
- Unconditional diffusion
- Unweighted negative examples



(b) On sampling choice

- Performance of each single time step varies, but worse than combining them together
- Smaller time steps outperform larger ones

Embedding Visualization



Embedding distance as proxy to hardness: smaller distances from the query node imply harder examples

- Examples of DMNS: generally harder than uniform sampling, but not too hard to impair the performance
- Multi-level samples capture wide range of hardness levels for negative sampling.

Outline

- Problem & related work
- Proposed model: DMNS
- Experiments
- **Conclusions**

Conclusions

- **Problem**
  - Multi-level negative sampling for graph link prediction
- **Proposed model: DMNS**
  - Empowers the sampling of multi-level negative examples, by sampling at different denoised steps of diffusion models
  - Adheres the sub-linear positivity principle for robust negative sampling
- **Experiments**
  - Extensive experiments demonstrate the effectiveness of DMNS

Thanks!

Diffusion-based Negative Sampling on Graphs for Link Prediction

Trung-Kien Nguyen and Yuan Fang

In Proceedings of the 2024 ACM Web Conference, May 13-17, 2024