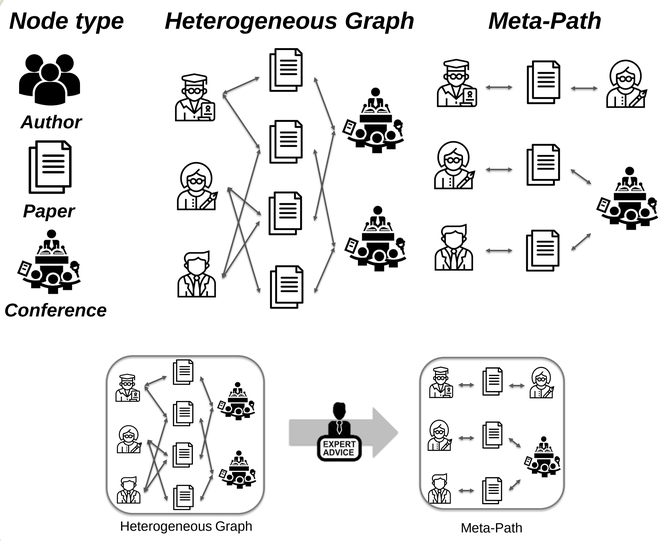
Graph Transformer Networks

Seoungjun Yun, Minbyul Jeong, Raehyun Kim, Jaewoo Kang*, Hyunwoo J. Kim*
Department of Computer Science and Engineering, Korea University, Seoul, Korea
{ ysj5419, minbyuljeong, raehyun, kangj, hyunwoojkim } @ korea.ac.kr

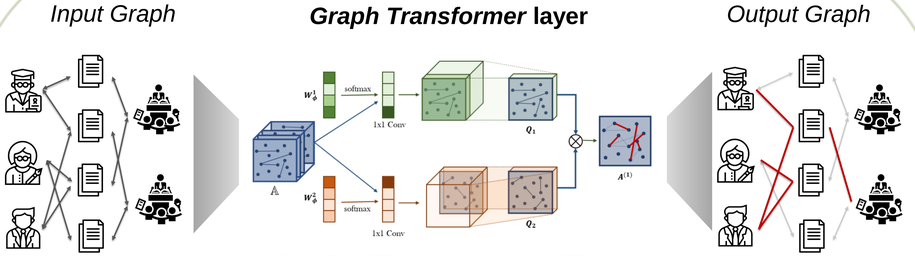
Background & Motivations



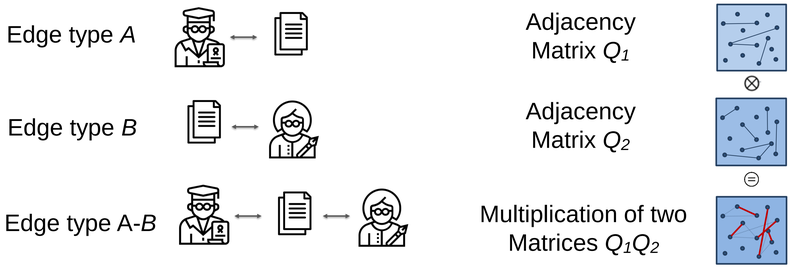
Contributions

1. We propose a novel framework **Graph Transformer Networks**, to **learn a new graph structure** which involves identifying useful meta-paths and multi-hop connections for learning effective node representation on graphs.
2. The graph generation is **interpretable** and the model is able to provide insight on **effective meta-paths** for prediction.
3. We prove the effectiveness of node representation learnt by Graph Transformer Networks resulting in **the best performance** against state-of-the-art methods that additionally use domain knowledge in all three benchmark **node classification on heterogeneous graphs**.

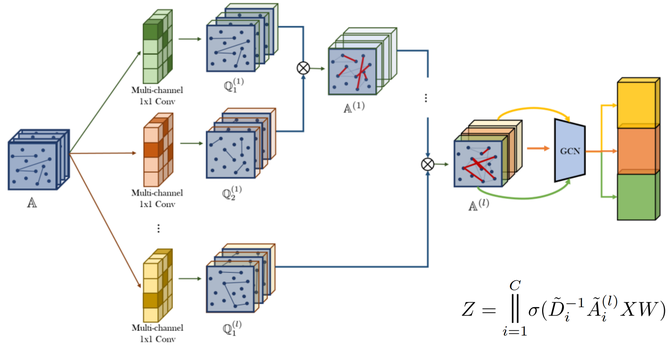
Our Approach



$$Q = F(\mathbb{A}; W_\phi) = \phi(\mathbb{A}; \mathrm{softmax}(W_\phi))$$
$$A^{(l)} = D^{-1} Q_1 Q_2$$



Graph Transformer Networks



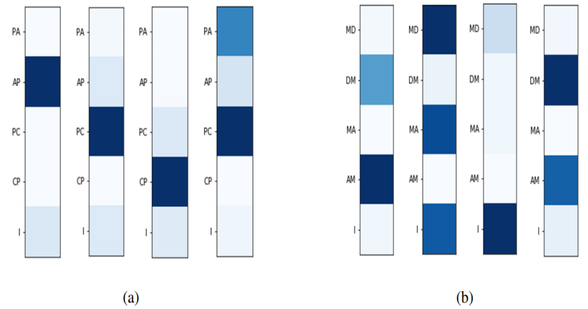
$$Z = \Big\|_{i=1}^{C} \sigma(\tilde{D}_i^{-1} \tilde{A}_i^{(l)} X W)$$

Analysis

Q1) Are the new graph structures generated by GTN effective for learning node representation?

	DeepWalk	metapath2vec	GCN	GAT	HAN	GTN$_{-I}$	**GTN (proposed)**
DBLP	63.18	85.53	87.30	93.71	92.83	93.91	**94.18**
ACM	67.42	87.61	91.60	92.33	90.96	91.13	**92.68**
IMDB	32.08	35.21	56.89	58.14	56.77	52.33	**60.92**

Q2) Can GTN adaptively produce a variable length of meta-paths depending on datasets?



(a) (b)

Q3) Can we interpret the importance of each meta-path from the adjacency matrix generated by GTNs?

Dataset	Predefined Meta-path	Meta-path learnt by GTNs Top 3 (between target nodes)	Top 3 (all)
DBLP	APCPA, APA	APCPA, APAPA, APA	APCPC, APCPA, PC
ACM	PAP, PSP	PAP, PSP	PAPA, APA, PAPS
IMDB	MAM, MDM	MDM, MAM, MDMDM	DM, AM, MDM