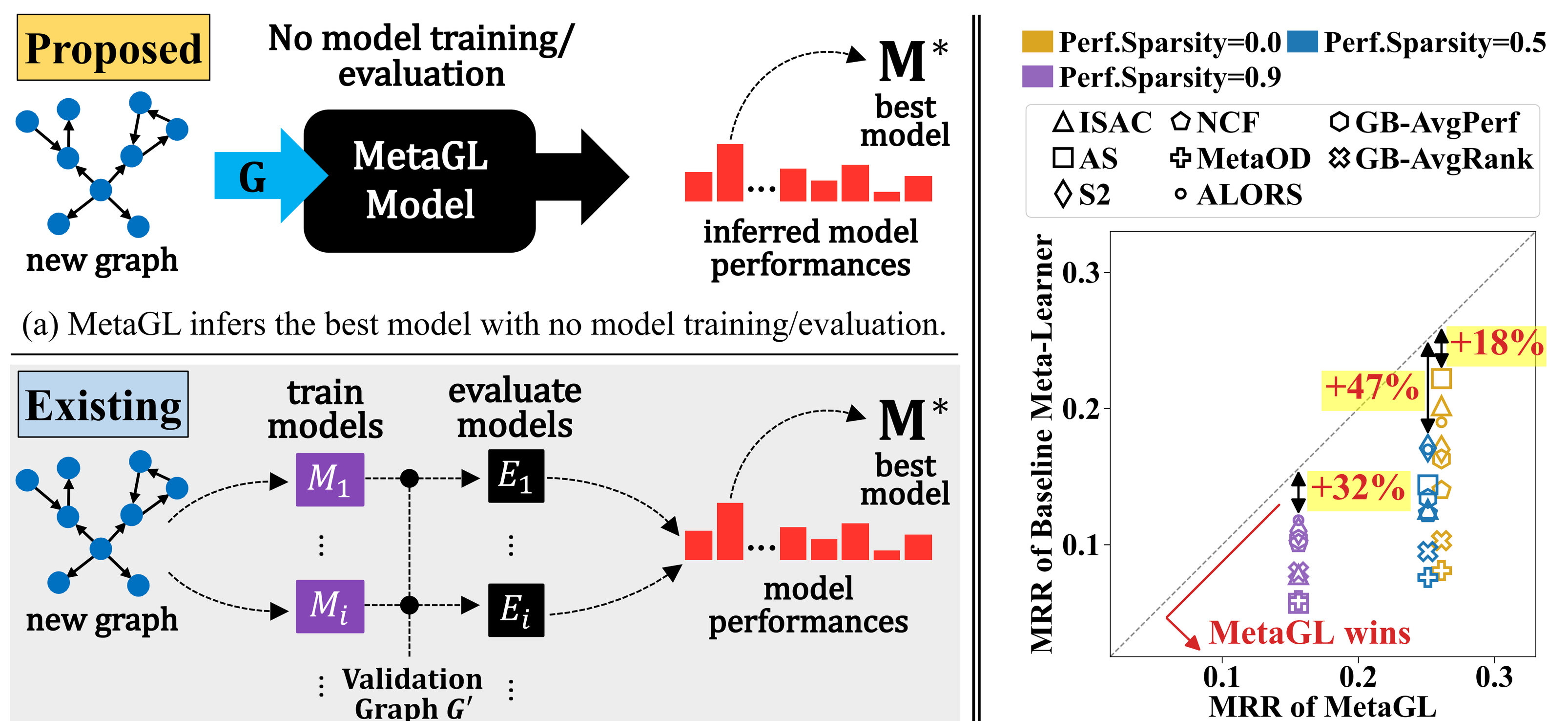
MetaGL: Evaluation-Free Selection of Graph Learning Models via Meta-Learning

Namyong Park, Ryan Rossi, Nesreen Ahmed, Christos Faloutsos

ICLR



(a) MetaGL infers the best model with no model training/evaluation.

(b) Existing model selection methods train/evaluate multiple models.

(c) MetaGL consistently performs the best.

Introduction

Given a graph learning task, such as link prediction, on a new graph, how can we select the best method as well as its hyperparameters (collectively called a model) without training or evaluating any model on the new graph? Model selection for graph learning has been largely ad hoc. A typical approach has been to apply popular methods to new datasets, but this is often suboptimal. On the other hand, systematically comparing models on the new graph quickly becomes too costly, or even impractical. In this work, we develop *the first meta-learning approach for evaluation-free graph learning model selection*, called MetaGL, which utilizes the prior performances of existing methods on various benchmark graph datasets to automatically select an effective model for the new graph, without any model training or evaluations. The following table compares MetaGL with existing model selection paradigms, showing that MetaGL satisfies all desirable features for graph learning (GL) model selection (MS).

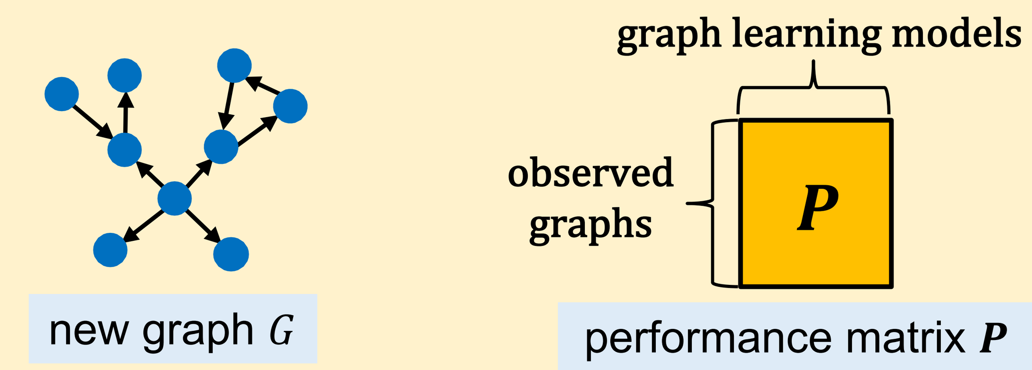
Desiderata for Graph Learning (GL) Model Selection (MS)	No model selection	Naive model selection	Graph HPO/NAS	MetaGL (Ours)
Evaluation-free GL model selection	✓			✓
Capable of MS from among multiple GL algorithms		✓		✓
Capitalizing on graph similarities for GL MS				✓
Estimating model performance based on past observations			✓	✓

Key Contributions of This Work

- **Problem Formulation.** We formulate the problem of selecting effective graph learning models in an evaluation-free manner.
- **Meta-Learning Framework and Features.** We propose MetaGL, the first meta-learning framework for the proposed problem, and meta-graph features that can quantify graph similarities for meta-learning over graphs.
- **Effectiveness.** MetaGL greatly outperforms existing methods (up to 47% better), with negligible runtime overhead at test time (~1 sec).

Problem Formulation

Given

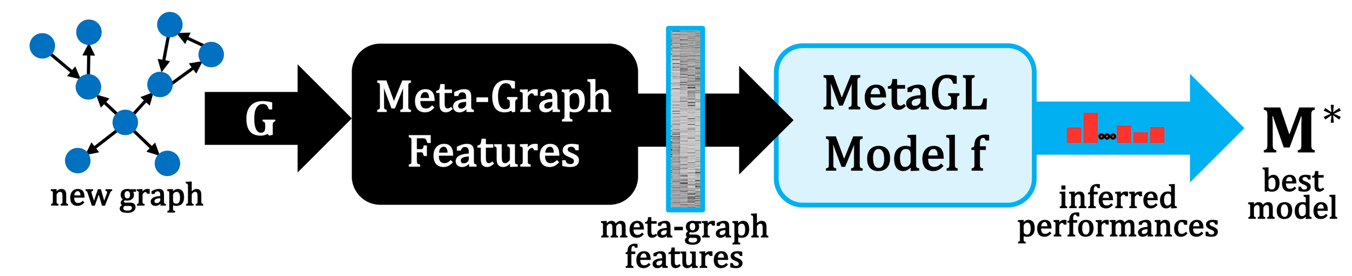


Select

the best model M^* ***without*** training/evaluating any model on the new graph G.

Proposed Framework: MetaGL

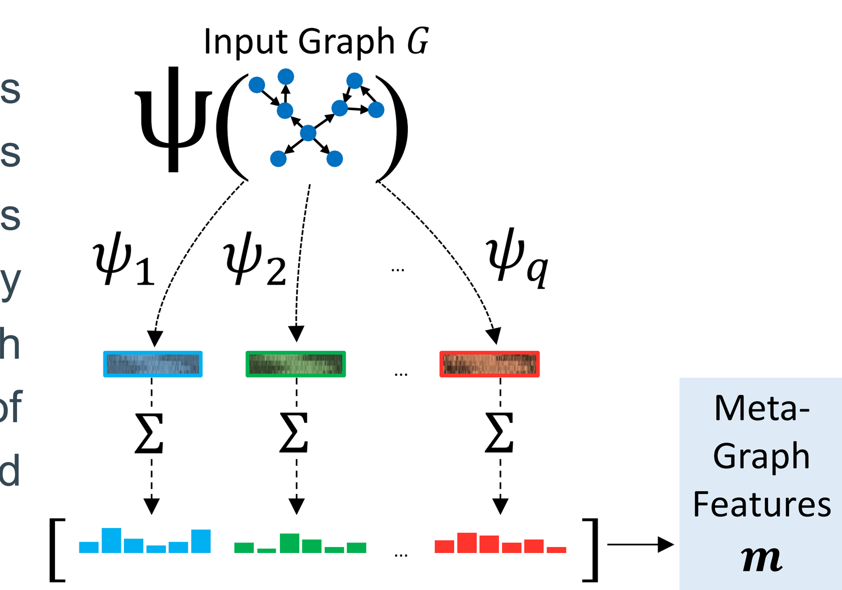
Overview



Given a new graph G, MetaGL extracts meta-graph features, and applies a meta-learned model f to them, which efficiently infers the best model $M^* \in \mathcal{M}$ for the new graph G, with no model evaluation.

Meta-Graph Features

Given a new task, meta-learning leverages prior experience from similar learning tasks to do a better job on the new task. To this end, MetaGL captures the graph similarity by extracting meta-graph features, which summarize the structural characteristics of a graph. Meta-graph features are designed to have the same size for all graphs.



Offline Meta-Training

In offline meta-training phase, MetaGL trains a meta-learner to estimate how well a model would perform on a given graph. The performance p_{ij} of model M_j on graph G_i is estimated to be the dot product between the representations of the corresponding model and graph, learned by a GNN-based embedding function that receives meta-graph features, and latent factors of models and graphs.

$$p_{ij} \approx \hat{p}_{ij} = \langle f(\mathbf{W}[\underbrace{\mathbf{m}_i}_{\text{meta-graph feature}}; \underbrace{\phi(\mathbf{m}_i)}_{\text{graph factor}}]), f(\underbrace{\mathbf{V}_j}_{\text{model factor}})\rangle$$

GNN-based embedding function $f(\cdot)$

Online Model Prediction

In online model prediction phase, MetaGL efficiently infers the best model M^* for the new graph, without any model training and evaluation.

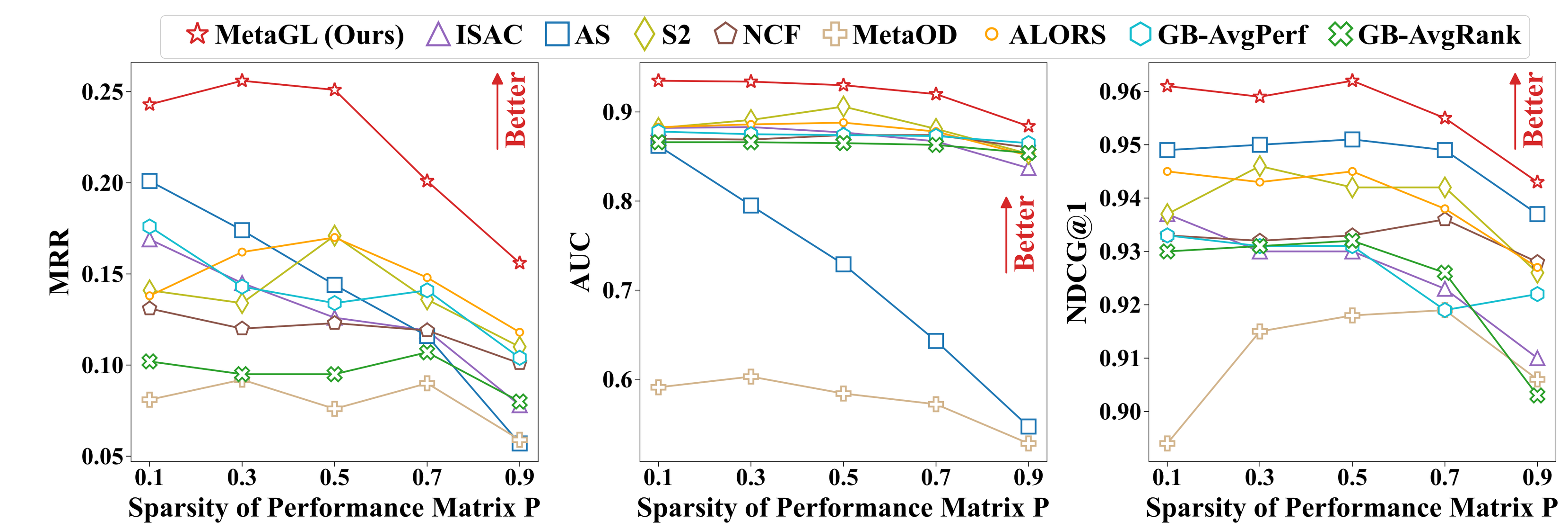
Best model $M^* = \arg\max_{M_j \in \mathcal{M}} \langle f(\mathbf{W}[\mathbf{m}_{\text{test}}; \phi(\mathbf{m}_{\text{test}})]), f(\mathbf{V}_j)\rangle$

Experiments

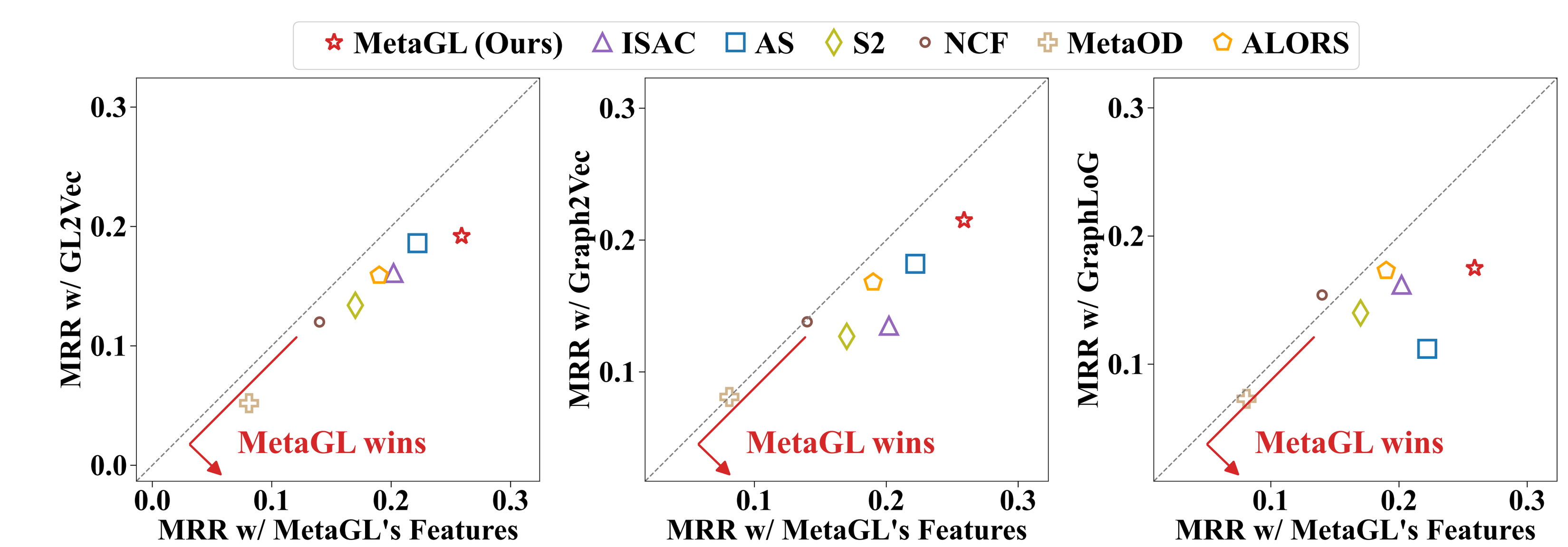
RQ1-a. Model Selection Performance with Fully Observed Perf. Matrix

	Method	MRR	AUC	NDCG@1
	Random Selection	0.011	0.490	0.745
Simple	Global Best-AvgPerf	0.163	0.877	0.932
	Global Best-AvgRank	0.103	0.867	0.930
	MetaGL_AS	<u>0.222</u>	0.905	0.947
	MetaGL_ISAC	0.202	0.887	0.939
Optimization-based	MetaGL_S2	0.170	<u>0.910</u>	0.945
	MetaGL_ALORS	0.190	0.897	<u>0.950</u>
	MetaGL_NCF	0.140	0.869	0.934
	MetaGL_MetaOD	0.075	0.599	0.889
	MetaGL (Ours)	**0.259**	**0.941**	**0.962**

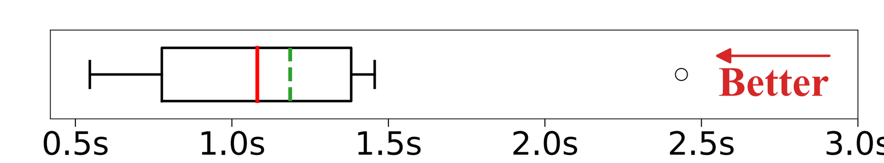
RQ1-b. Model Selection Performance with Partially Observed Perf. Matrix



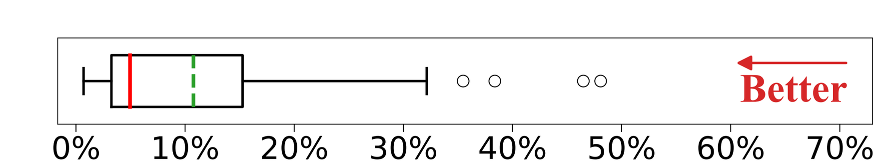
RQ2. Effectiveness of Meta-Graph Features



RQ3. Model Selection Efficiency



MetaGL's runtime (secs) at test time



MetaGL's runtime / time to train GL models

QR Code for Our Paper