Micro and Macro Level Graph Modeling for Graph Variational Auto-Encoders

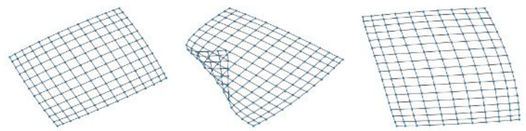
Kiarash Zahirnia, Oliver Schulte, Parmis Naddaf, Ke Li

NEURAL INFORMATION PROCESSING SYSTEMS

Graph Generation

- Given a set of observed graphs **G** = {$\mathscr{G}_1$, ... $\mathscr{G}_S$ } sampled from data distribution p(G), the goal of learning generative models for graphs is to learn a generative model p_θ(G) which is similar to p(G).
- This paper focus on the ***realism*** of generated graphs.

Samples from Protein dataset (real data).

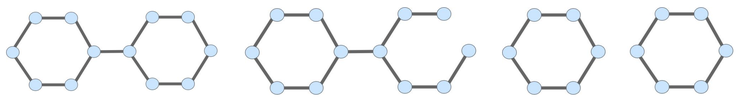


Samples from Grid dataset (synthetic benchmark).

- Deep Graph Generative Models (GGM)s fall into 2 categories (Hamilton 2020):
  1) All-at-once Models
  2) Autoregressive Models
- Current all-at-once methods are faster, but tend to generate less realistic graphs.

Global and Local Graph Properties

- It is common in graph analysis to distinguish two levels of information:
  1) Local node-level properties; e.g degree of a specific node.
  2) Global graph-level properties; e.g degree distribution.
- Deep GGMs are trained with an objective based on ***local properties***.
- ***Local properties do not model different edge importance in the graph global structure.***
- Example:



Original Graph **Generated Graph 1** **Generated Graph 2**

- The two right graphs score the same in terms of the number of reconstructed edges from the original graph, *a node-level score*.
- But Graph 1 is structurally more similar to the Original Graph, at the *global graph-level*.

References

Martin Simonovsky et al. (2018). "GraphVAE: Towards generation of small graphs using variational autoencoders". ICANN.

Jiaxuan You et al. (2018). "GraphRNN: Generating realistic graphs with deep auto-regressive models". ICLR.

Renjie Liao et al. (2019). "Efficient graph generation with graph recurrent attention networks". NeurIPS.

Hanjun Dai et al. (2020). "Scalable deep generative modeling for sparse graphs". ICML.

William L Hamilton. (2020). "Graph representation learning". Synthesis Lectures on Artificial Intelligence and Machine Learning.

Rylee Thompson et al. (2022). "On evaluation metrics for graph generative models". ICLR.

Leslie O'Bray et al. (2022). "Evaluation metrics for graph generative models: Problems, pitfalls, and practical solutions". ICLR.

Micro-macro (MM) Modeling:

- A principled probabilistic ***framework*** that incorporates both local (Micro) and global (Macro) graph properties.
- Assuming a predefined finite set of graph global statistics/properties, calculated by $\varphi_1()$, …, $\varphi_m()$ micro-macro loss is of the form:

$$\mathcal{L}_{\boldsymbol{\theta}}(\boldsymbol{A}) = \mathcal{L}^0_{\boldsymbol{\theta}}(\boldsymbol{A}) + \gamma \mathcal{L}^1_{\boldsymbol{\theta}}(\boldsymbol{F}_1, \ldots, \boldsymbol{F}_m)$$

- L^0: micro loss.
- L^1: macro loss.
- **A**: training sample.
- $\mathbf{F}_u$: random variable defined by $\varphi_u(\mathbf{A})$.
- γ: hyperparameter.

GraphVAE-MM Objective

- This paper works with negative log-likelihood losses:

$$\mathcal{L}^0_{\boldsymbol{\psi}}(\boldsymbol{A}) = -\ln p^0_{\boldsymbol{\psi}}(\boldsymbol{A}) = -\ln \int P(\boldsymbol{A}|\tilde{\boldsymbol{A}}_{\boldsymbol{z}}) p(\boldsymbol{z}) d\boldsymbol{z}$$

$$\mathcal{L}^1_{\boldsymbol{\psi},\boldsymbol{\sigma}}(\boldsymbol{F}_1, \ldots, \boldsymbol{F}_m) = -\sum_{u=1}^{m} \frac{1}{|\boldsymbol{F}_u|} \ln p^1_{\boldsymbol{\psi},\boldsymbol{\sigma}}(\boldsymbol{F}_u)$$

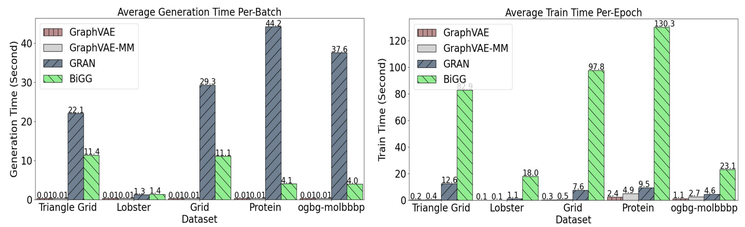
- $\tilde{\mathbf{A}}_z$: probabilistic adjacency matrix computed as a function of graph embedding z.
- $|\mathbf{F}_u|$: dimensionality of $\mathbf{F}_u$.
- Approximating with variational Lower bound,

$$\mathcal{L}_{\boldsymbol{\theta}}(\boldsymbol{A}) \leq E_{\boldsymbol{z}\sim q_{\varphi}(\boldsymbol{z}|\boldsymbol{A})}\big[-\ln p^0_{\boldsymbol{\psi}}(\boldsymbol{A}|\tilde{\boldsymbol{A}}_{\boldsymbol{z}}) - \sum_{u=1}^{m} \frac{1}{|\boldsymbol{F}_u|} \ln p^1_{\boldsymbol{\psi},\boldsymbol{\sigma}}(\boldsymbol{F}_u)\big] + (1+\gamma m) KL(q_{\varphi}(\boldsymbol{z}|\boldsymbol{A})||p(\boldsymbol{z}))$$

Experiments

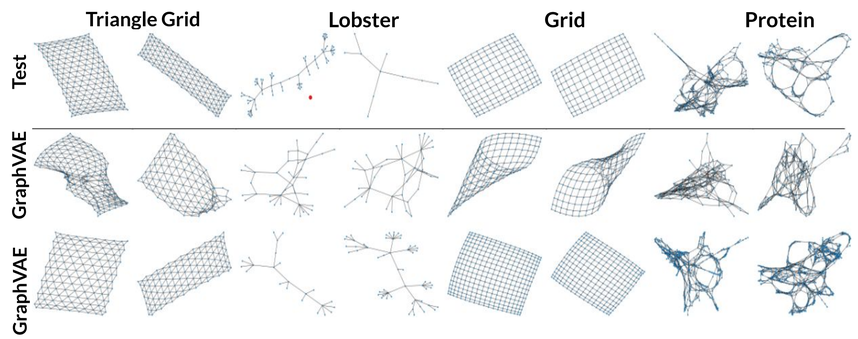
- **GraphVAE** (Dai et al 2018) architecture: All-at-once model with fast training and generation time for medium-sized graphs.
- **GraphVAE-MM:** Uses e micro-macro objective to improve GraphVAE.
- To evaluate the effect of MM modeling, we follow a AB design: utilize GraphVAE architecture as is and change only the training objective.
- In our experiments, we utilize 3 default global graph-level properties:
  - Degree histogram
  - Number of triangles
  - S-Step transition probability for S=2,...,5.

Generation and Train Time



- MM modeling maintains the GraphVAE generation speed advantage.

Qualitative Evaluation



- GraphVAE–MM achieves much better visual match than GraphVAE.

Quantitative Evaluation (GNN-based evaluation metrics

Thompson et al 2022)

Method	Triangle Grid		Lobster		Grid		ogbg-molbbbp		Protein	
	MMD RBF	F1 PR	MMD RBF	F1 PR	MMD RBF	F1 PR	MMD RBF	F1 PR	MMD RBF	F1 PR
50/50 split	0.03 ± 0.00	98.58 ± 0.00	0.04 ± 0.00	98.58 ± 0.00	0.009 ± 0.00	98.70 ± 0.00	0.002 ± 0.00	98.07 ± 0.00	0.04 ± 0.00	98.67 ± 1.11
GraphVAE	0.23 ± 0.01	75.92 ± 8.96	0.36 ± 0.11	78.48 ± 24.13	0.17 ± 0.01	75.52 ± 2.53	0.20 ± 0.07	54.53 ± 6.15	0.10 ± 0.05	84.11 ± 9.56
GraphVAE-MM	$\mathbf{0.17 \pm 0.01}$	$\mathbf{83.58 \pm 5.50}$	$\mathbf{0.10 \pm 0.00}$	$\mathbf{100.00 \pm 0.00}$	$\mathbf{0.13 \pm 0.01}$	$\mathbf{97.09 \pm 6.33}$	$\mathbf{0.02 \pm 0.01}$	93.78 ± 1.33	$\mathbf{0.03 \pm 0.01}$	90.78 ± 3.76
GraphRNN-S (You et al.)	0.72 ± 0.17	33.68 ± 19.44	0.98 ± 0.13	58.72 ± 7.55	0.79 ± 0.08	71.18 ± 2.36	0.48 ± 0.02	81.41 ± 0.71	0.28 ± 0.26	72.36 ± 27.63
GraphRNN (You et al.)	0.64 ± 0.11	25.80 ± 11.75	0.87 ± 0.04	61.97 ± 0.00	0.99 ± 0.03	13.22 ± 0.05	1.45 ± 0.19	$\mathbf{98.94 \pm 0.56}$	0.32 ± 0.14	93.94 ± 0.56
GRAN (Liao et al.)	0.88 ± 0.09	23.71 ± 9.72	0.24 ± 0.04	50.53 ± 12.12	0.40 ± 0.00	78.73 ± 0.02	0.39 ± 0.07	94.06 ± 2.60	0.07 ± 0.00	98.05 ± 0.76
BiGG (Dai et al.)	0.41 ± 0.13	62.08 ± 0.14	0.12 ± 0.00	99.74 ± 0.76	0.35 ± 0.00	92.43 ± 0.00	0.04 ± 0.00	96.16 ± 0.31	0.15 ± 0.00	$\mathbf{98.11 \pm 0.62}$

- **MMD RBF and F1 PR** capture the reality and diversity of generated graphs, respectively.
- **Impact on GraphVAE.** MM modeling provides a large improvement in the realism and diversity of graphs generated by a GraphVAE architecture.
- **GraphVAE-MM vs. Benchmark GGMs.** Micro-macro (MM) modeling greatly improved the GraphVAE, to match or exceed that of benchmark models.

Quantitative Evaluation (statistic-based evaluation metrics

O'Bray et al 2022)

(a) Synthetic Graphs

Method	Triangle Grid					Lobster					Grid				
	Deg.	Clus.	Orbit	Spect	Diam.	Deg.	Clus.	Orbit	Spect	Diam.	Deg.	Clus.	Orbit	Spect	Diam.
50/50 split	$3e^{-5}$	0.002	$8e^{-5}$	0.004	0.014	0.002	0	0.002	0.005	0.032	$1e^{-5}$	0	$2e^{-5}$	0.004	0.014
GraphVAE	0.0821	0.442	0.421	0.020	0.152	0.081	0.739	0.372	0.056	0.129	0.062	0.055	0.515	0.018	0.143
GraphVAE-MM	**0.001**	**0.093**	**0.001**	**0.013**	**0.133**	$\mathbf{2e^{-4}}$	**0**	0.008	0.017	0.187	$\mathbf{5e^{-4}}$	**0**	**0.001**	0.014	**0.065**
GraphRNN-S (You et al.)	0.053	1.094	0.121	0.033	1.124	0.016	0.319	0.285	0.045	0.242	0.014	0.003	0.090	0.112	0.128
GraphRNN (You et al.)	0.033	1.167	0.107	0.030	1.121	0.004	**0**	0.033	0.035	0.384	0.013	0.166	0.019	0.111	0.460
GRAN (Liao et al.)	0.134	0.678	0.673	0.184	1.133	0.005	0.304	0.331	0.043	0.446	0.003	$1e^{-4}$	0.007	**0.012**	0.281
BiGG (Dai et al.)	**0.001**	0.107	0.004	0.020	1.123	0.001	**0**	$\mathbf{6e^{-4}}$	**0.012**	**0.101**	0.002	$3e^{-5}$	0.003	0.018	0.328

(b) Real Graphs

Method	Protein					ogbg-molbbbp				
	Deg.	Clus.	Orbit	Spect	Diam.	Deg.	Clus.	Orbit	Spect	Diam.
50/50 split	$4e^{-5}$	0.004	$5e^{-4}$	$4e^{-4}$	0.003	$2e^{-4}$	$2e^{-5}$	$9e^{-5}$	$5e^{-4}$	0.002
GraphVAE	0.022	0.108	0.577	0.016	0.080	0.028	0.442	0.047	0.015	0.055
GraphVAE-MM	0.006	**0.059**	0.152	0.007	0.091	**0.001**	0.005	$8e^{-4}$	**0.005**	**0.018**
GraphRNN-S (You et al.)	0.046	0.324	0.316	0.028	0.302	0.016	0.572	0.006	0.045	0.199
GraphRNN ((You et al.)	0.012	0.123	0.264	0.018	0.342	0.002	$\mathbf{9e^{-4}}$	$4e^{-4}$	0.135	0.495
GRAN (Liao et al.)	**0.003**	**0.059**	**0.053**	**0.004**	**0.009**	0.008	0.353	0.013	0.056	0.317
BiGG (Dai et al.)	0.007	0.099	0.316	0.012	0.181	0.003	0.001	$\mathbf{5e^{-5}}$	0.007	0.033

- Statistic-based evaluation of MM modeling shows MM modeling improves the reality of graphs generated by GraphVAE, up to 2 orders of magnitude on five benchmark datasets.