Chair of Network Architectures and Services
Department of Computer Engineering
Technical University of Munich

Synthesizing and Scaling WAN Topologies using Permutation-invariant Graph Generative Models

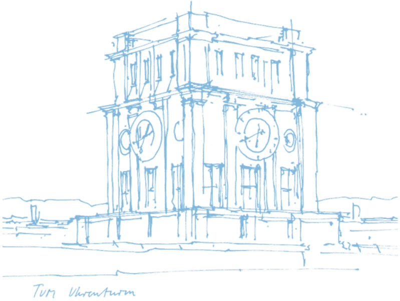
Max Helm, and Georg Carle

October 30, 2023

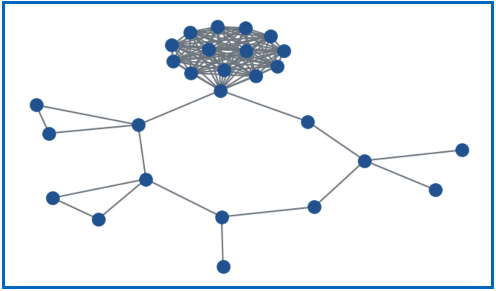
CNSM 19

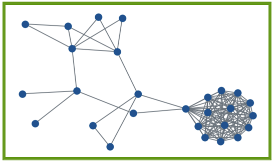
Niagara Falls, Canada

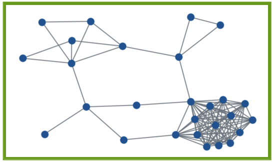
Chair of Network Architectures and Services
Department of Computer Engineering
Technical University of Munich

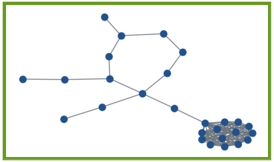


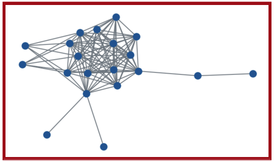
Motivation











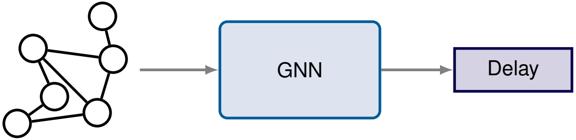
Non-isomorphic

- Graph can not be mapped to the original graph
- Think of two adjacency matrics that are different but describe same graph

Similar structure

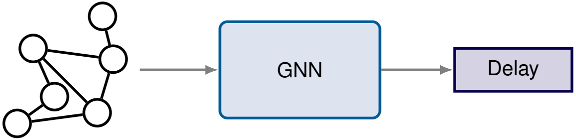
- Graph properties (e.g., number of nodes)
- Graph metrics (e.g., degree centrality)
- Graph distribution metrics (e.g., maximum mean discrepancy of graph mertic)

Why?



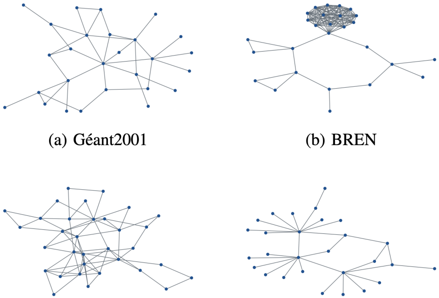
- Machine-learning applications need training data
- Training data needs to be realistic and abundant
- Alternative to random graph generation (e.g., Erdős–Rényi)
- No additional knowledge needed (e.g., Autonomous System graphs approximate powerlaw graphs)

Why?



- Machine-learning applications need training data
- Training data needs to be realistic and abundant
- Alternative to random graph generation (e.g., Erdős–Rényi)
- No additional knowledge needed (e.g., Autonomous System graphs approximate powerlaw graphs)

What?



(a) Géant2001 (b) BREN

(c) BtNorthAmerica (d) GtsSlovakia

- Wide Area Network (WAN) topologies
- Use Internet Topology Zoo (approx. 250 WAN topologies)

Related Work

2022 18th International Conference on Network and Service Management (CNSM)

Comparing Traditional and GAN-based Approaches for the Synthesis of Wide Area Network Topologies

Katharina Dietz, Michael Seufert, Tobias Hoßfeld
University of Würzburg, Germany
{katharina.dietz, michael.seufert, tobias.hossfeld}@uni-wuerzburg.de

- Generative Adverserial Network
- Works on adjacency matrix permutations
- Henceforth referred to as *Dietz*

$\rightarrow$ Permutation invariant: ✗
$\rightarrow$ Used and tested for WANs: ✓

Related Work

2022 18th International Conference on Network and Service Management (CNSM)

Comparing Traditional and GAN-based Approaches for the Synthesis of Wide Area Network Topologies

Katharina Dietz, Michael Seufert, Tobias Hoßfeld
University of Würzburg, Germany
{katharina.dietz, michael.seufert, tobias.hossfeld}@uni-wuerzburg.de

- Generative Adverserial Network
- Works on adjacency matrix permutations
- Henceforth referred to as *Dietz*

$\rightarrow$ Permutation invariant: ✗
$\rightarrow$ Used and tested for WANs: ✓

Scalable Deep Generative Modeling for Sparse Graphs

Hanjun Dai[1] Azade Nazi[1] Yujia Li[2] Bo Dai[1] Dale Schuurmans[1]

GraphGDP: Generative Diffusion Processes for Permutation Invariant Graph Generation

Han Huang[1], Leilei Sun[1*], Bowen Du[1], Yanjie Fu[2], Weifeng Lv[1]

Permutation Invariant Graph Generation via Score-Based Generative Modeling

Chenhao Niu[1], Yang Song[2], Jiaming Song[2], Shengjia Zhao[2], Aditya Grover[2], Stefano Ermon[2]

- Graph based models
- Work directly on graphs

$\rightarrow$ Permutation invariant: ✓
$\rightarrow$ Used and tested for WANs: ✗

$\Rightarrow$ Adapt and benchmark permutation-invariant approaches on the task of WAN topology generation

Methodology

Name	Year	Approach	Permutation Invariant	Directed	Edge Weights	Node Feat.
Netgan	2018	Random walks	✓	✗	✗	✗
GraphRNN	2018	Autoregr. model	BFS-order	✗	✗	✗
GSM	2020	Score match w/ GNNs	✓	✗	✗	✗
BiGG	2020	Autoregr. tree model	DFS-/BFS-order	✓	✗[1]	✓
GraphGDP	2022	Diffusion model	✓	✗	✗	✗

[1] Possible to indirectly include edge weights as node features

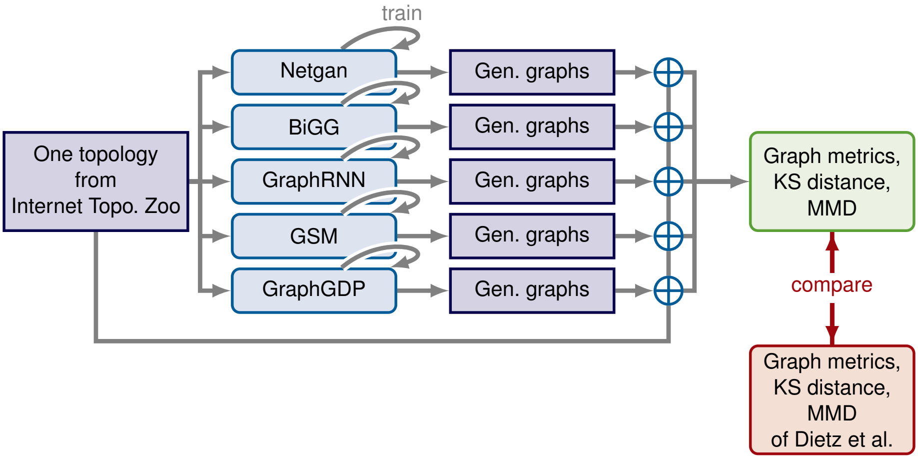
Methodology

Name	Year	Approach	Permutation Invariant	Directed	Edge Weights	Node Feat.
Netgan	2018	Random walks	✓	✗	✗	✗
GraphRNN	2018	Autoregr. model	BFS-order	✗	✗	✗
GSM	2020	Score match w/ GNNs	✓	✗	✗	✗
BiGG	2020	Autoregr. tree model	DFS-/BFS-order	✓	✗[1]	✓
GraphGDP	2022	Diffusion model	✓	✗	✗	✗

[1] Possible to indirectly include edge weights as node features

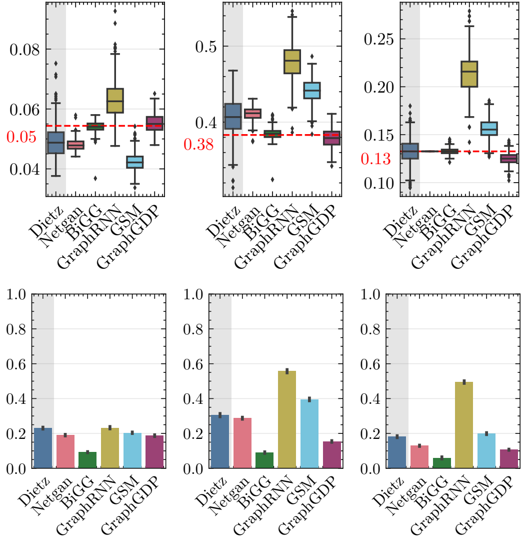
⇒ Compare and benchmark these five approaches against each other and related work

Methodology



Results

One Topology Example



Direct graph metric comparison to original graph:

- Betweenness Centrality (BC)
- Closeness Centrality (CC)
- Degree Centrality (DC)

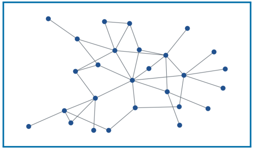
Distribution distance of graph metric:

- Kolmogorov-Smirnov (KS) distance of BC
- Kolmogorov-Smirnov (KS) distance of CC
- Kolmogorov-Smirnov (KS) distance of DC

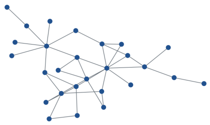
Results

Good and Bad Fits

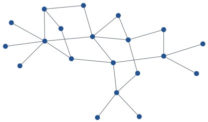
Good Scores:



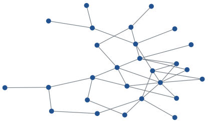
(a) Original



(b) Netgan



(d) GraphRNN

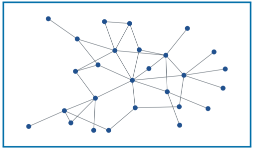


(e) GSM

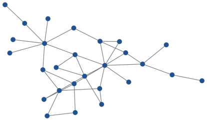
Results

Good and Bad Fits

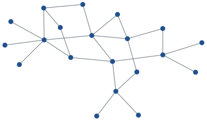
Good Scores:



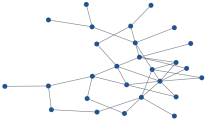
(a) Original



(b) Netgan

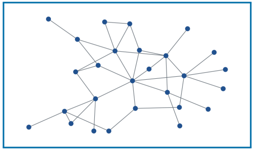


(d) GraphRNN

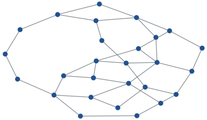


(e) GSM

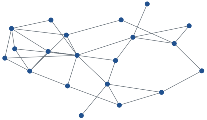
Bad Scores:



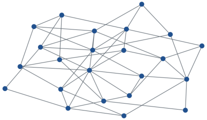
(a) Original



(b) Netgan

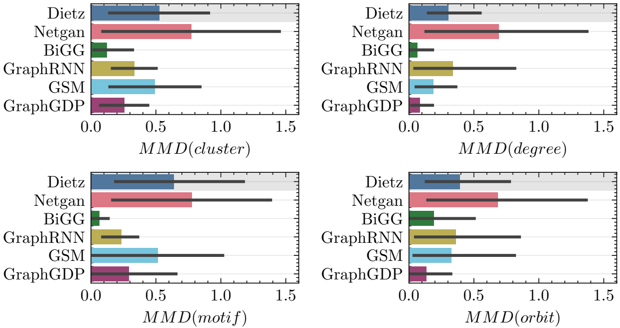


(d) GraphRNN



(e) GSM

Does it work consistently for all topologies?



- Maximum Mean Discrepancy (MMD)
- Measures similarity of all momentums
- Often used to evaluate graph generative models
- **Over all graphs**

Results

⇒ Specifically **BiGG** and **GraphGDP** are able to generate high quality WAN topologies

$\Rightarrow$ Specifically **BiGG** and **GraphGDP** are able to generate high quality WAN topologies

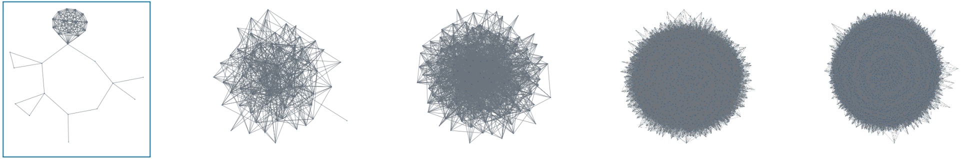
What about changing the size of the network?

⇒ Specifically **BiGG** and **GraphGDP** are able to generate high quality WAN topologies

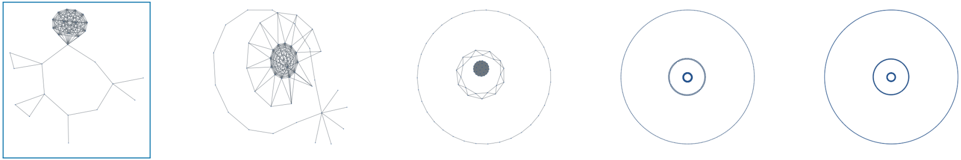
What about changing the size of the network?

Name	Year	Approach
Kronecker	2010	Kronecker product matrix operation
Gscaler	2016	DNA shotgun sequencing variation
EvoGraph	2018	Preferential edge attachment

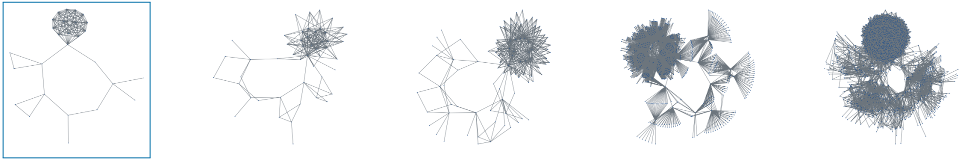
Results



(a) Original (b) Kronecker scale=2 (c) Kronecker scale=4 (d) Kronecker scale=32 (e) Kronecker scale=64



(f) Original (g) Gscaler sacle=2 (h) Gscaler scale=4 (i) Gscaler scale=32 (j) Gscaler scale=64



(k) Original (l) EvoGraph scale=2 (m) EvoGraph scale=4 (n) EvoGraph scale=32 (o) EvoGraph scale=64

Conclusion

Summary:

- Permutation-invariant models outperform adjacency-matrix-based approaches
- Scaling WAN topologies works well with preferential edge attachment, but not with DNA shotgun sequencing or Kronecker products

```
import pickle
import networkx as nx

with open(filename, 'rb') as f:
    graphs = pickle.load(f)
```

In the paper:

- Description of generation and scaling methods
- Dataset available

Paper:

Dataset (interactive):

Dataset: