RDB2G-Bench: A Comprehensive Benchmark for Automatic Graph Modeling of Relational Databases

RDB2G AUTOMATIC GRAPH MODELING BENCHMARK

Dongwon Choi[1], Sunwoo Kim[1], Juyeon Kim[1], Kyungho Kim[1], Geon Lee[1], Shinhwan Kang[1], Myunghwan Kim[2], and Kijung Shin[1]

[1]{cookie000215, kswoo97, juyeonkim, kkyungho, geonlee0325, shinhwan.kang, kijungs}@kaist.ac.kr, [2]myunghwan@kumo.ai

NEURAL INFORMATION PROCESSING SYSTEMS

KAIST AI Kim Jaechul Graduate School

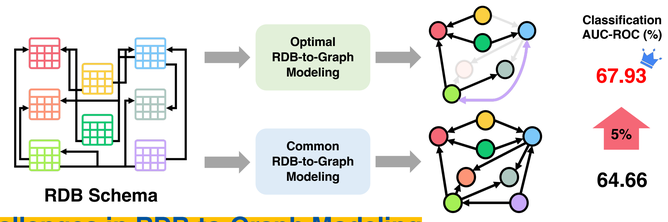
kumo

Summary

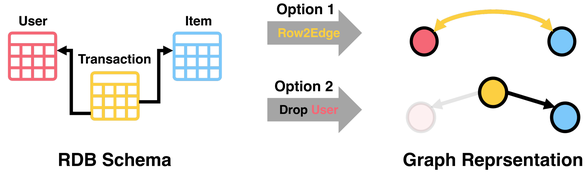
- **Motivations**
  - **[Impact on Performance]** In graph-based ML for relational databases (RDBs) [1], performance heavily depends on the RDB-to-graph modeling.
  - **[Difficult Evaluation]** There are numerous possible graph models, which require costly evaluation with repeated GNN trainings.
  - **[Lack of Benchmark]** No standard datasets/benchmarks exist to evaluate automatic RDB-to-Graph modeling.
- **Contributions**
  - **[Datasets]** We precomputed 50k graph model-performance pairs from 5 real-world RDBs and 12 predictive tasks.
  - **[Benchmarks]** We benchmarked 10 automatic RDB-to-graph modeling, which enables 380x faster evaluation.
  - **[Observations]** Our analysis provides practical insights for graph modeling.

Introduction

- **RDB-to-Graph Modeling**
  - **[Problem Definition]** To find a **graph representation of the given RDB** to maximize the performance of the predictive GNN on the given task.
    - Typically, nodes are table rows and edges are FK relations.
    - Careful RDB-to-Graph modeling is crucial in practice.



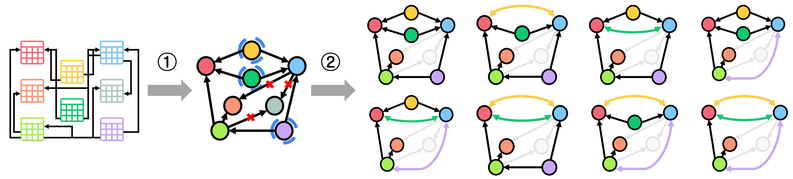
- **Challenges in RDB-to-Graph Modeling**
  - **[Search Space]** There are so many possible graph representations of RDBs.
    - **[Option 1]** We can model table rows not only as nodes but also as edges.
    - **[Option 2]** We can selectively include tables and FK relations to the graph.
  - **[Evaluation Cost]** Evaluating RDB-to-Graph modeling is computationally costly.



Datasets for RDB2G-Bench

- **Proposed Datasets**
  - **[Key Idea]** For **all possible graph representations**, we precomputed GNN performance metrics, enabling fast evaluation of RDB-to-Graph modeling methods.
  - **[Source]** Our datasets were created based on 5 real-world RDBs and 12 predictive tasks provided in RelBench [2].
- **Construction of Graph Representation**
  - **[Step 1]** Selecting which tables and FK relationships to include in the graph.
  - **[Step 2]** Selecting how to represent the rows of each table, as either nodes or edges.

  ⇒ We constructed about **50k distinct graph representations** across RDBs and tasks.



- **Collection of Performance Metrics**
  - **[Predictive Performance]** The performance of each split on the downstream task.
  - **[Runtime]** The total elapsed time per epoch over at least five trials.
  - **[Parameter Size]** The total number of learnable parameters of predictive GNNs.

Benchmarks for RDB2G-Bench

- **Baselines for RDB-to-Graph Modeling**
  - **[Overview]** We benchmarked **10 RDB-to-Graph modeling methods**:
    - **[Heuristic-Based]** Use simple rules (e.g., All-Rows-to-Nodes [1]) to model an RDB as a graph.
    - **[Action-Based Search]** Explore and optimize graph representations by iteratively modifying them using a predefined actions (i.e., set of operations).
    - **[LLM-Based]** Generate action sequences for graph modeling leveraging LLMs.

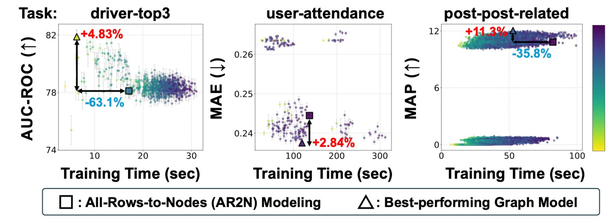
References

[1] Relational Deep Learning: Graph Representation Learning on Relational Databases

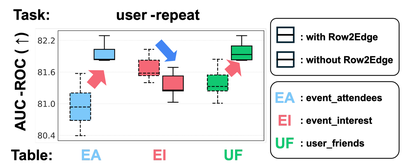
[2] RelBench: A Benchmark for Deep Learning on Relational Databases

Observations from RDB2G-Bench

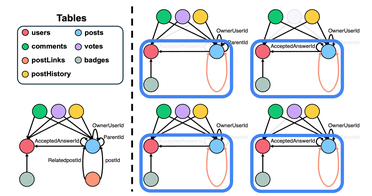
- **Observations from the RDB2G-Bench datasets**
  - **[Obs 1]** Finding the best graph representations is worthwhile with better performance.



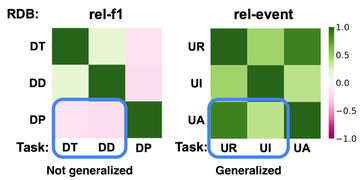
  - **[Obs 2]** Modeling table rows as edges can be crucial, depending on tables & tasks.



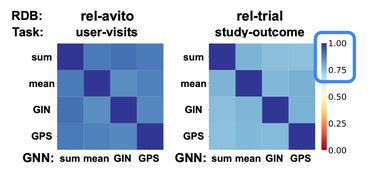
  - **[Obs 3]** Top-performing graph representations share common substructures.



  - **[Obs 4]** Different tasks may prefer different graph representations, even on the same RDB.



  - **[Obs 5]** Effectiveness of graph representations generalizes across GNNs.



- **Benchmark Results on the RDB2G-Bench Datasets**
  - **[Rank Comparison]** Most baselines outperform a widely-adopted AR2N heuristic.
  - **[Efficiency Gain]** Using the precomputed metrics in our RDB2G-Bench speed-ups evaluation by up to **389x**, compared to on-the-fly evaluation with GNN training.