GAP: Differentially Private Graph Neural Networks with Aggregation Perturbation

Sina Sajadmanesh

Idiap & EPFL

Ali Shahin Shamsabadi

The Alan Turing Institute

Aurélien Bellet

Inria

Daniel Gatica-Perez

Idiap & EPFL

32nd USENIX Security Symposium
August 2023

- **Graph Neural Networks** (GNNs) are state-of-the-art algorithms for learning on graphs
  - **Tasks:** node classification, link prediction, ...
  - **Applications**: recommendation systems, credit issuing, traffic forecasting, drug discovery, ...

Introduction & Motivation

- **Graph Neural Networks** (GNNs) are state-of-the-art algorithms for learning on graphs
  - **Tasks:** node classification, link prediction, ...
  - **Applications**: recommendation systems, credit issuing, traffic forecasting, drug discovery, ...
- Graph data could be privacy-sensitive
  - e.g., users' personal attributes, financial transactions, medical/biological networks, . . .

- **Graph Neural Networks** (GNNs) are **state-of-the-art** algorithms for learning on graphs
  - **Tasks:** node classification, link prediction, ...
  - **Applications**: recommendation systems, credit issuing, traffic forecasting, drug discovery, ...
- Graph data could be **privacy-sensitive**
  - e.g., users' personal attributes, financial transactions, medical/biological networks, . . .
- GNNs are vulnerable to **privacy attacks**
  - e.g., **link stealing attack** [He et al., 2021] or **membership inference attack** [Olatunji et al., 2021]

Introduction & Motivation

- **Graph Neural Networks** (GNNs) are **state-of-the-art** algorithms for learning on graphs
  - **Tasks:** node classification, link prediction, ...
  - **Applications**: recommendation systems, credit issuing, traffic forecasting, drug discovery, ...
- Graph data could be **privacy-sensitive**
  - e.g., users' personal attributes, financial transactions, medical/biological networks, ...
- GNNs are vulnerable to **privacy attacks**
  - e.g., **link stealing attack** [He et al., 2021] or **membership inference attack** [Olatunji et al., 2021]

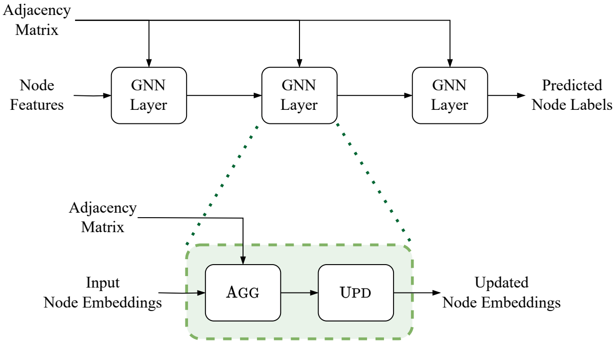
How to preserve privacy of individuals when learning over graph data?

Our Contributions

- GAP: a novel **GNN with differential privacy** (DP) guarantees
  - **Aggregation Perturbation** to preserve privacy of graph edges
  - **Tailored Architecture** to maintain privacy budget
  - **Formal Privacy Analysis** for both edge-level and node-level DP

Graph Neural Networks

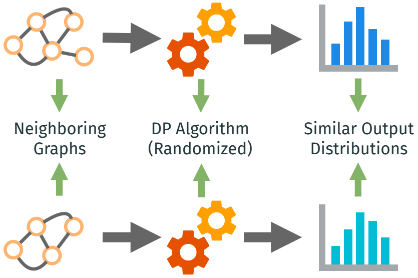
- Graph Neural Networks (GNNs) learn node representations based on node features and the graph structure



Differential Privacy [Dwork et al., 2006]

Randomized algorithm A is ϵ-DP if for all neighboring datasets $G \simeq G'$ and all sets of outputs S:

$$\frac{\Pr[A(G) \in S]}{\Pr[A(G') \in S]} \leq e^{\epsilon}$$

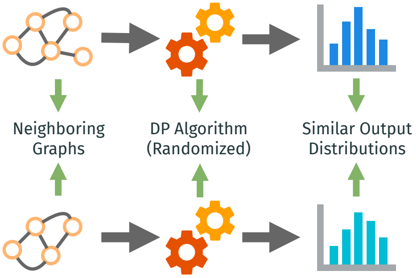


Differential Privacy [Dwork et al., 2006]

Randomized algorithm A is ϵ-DP if for all neighboring datasets $G \simeq G'$ and all sets of outputs S:

$$\frac{\Pr[A(G) \in S]}{\Pr[A(G') \in S]} \leq e^{\epsilon}$$

- **Edge-Level DP**
  Neighboring graph datasets differ by at most one edge

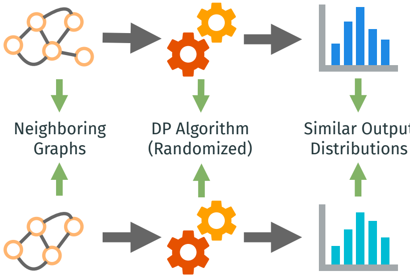


Differential Privacy [Dwork et al., 2006]

Randomized algorithm A is ϵ-DP if for all **neighboring datasets** $G \simeq G'$ and all sets of outputs S:

$$\frac{\Pr[A(G) \in S]}{\Pr[A(G') \in S]} \leq e^{\epsilon}$$

- **Edge-Level DP**
  Neighboring graph datasets differ by at most **one edge**
- **Node-Level DP**
  Neighboring graph datasets differ by at most **one node** (and all adjacent edges)



- Exploding Sensitivity
  - With a K-layer GNN, each node affects the embedding of all the nodes in its K-hop neighborhood
  - $O(D^K)$ gradient terms change at once (D is maximum degree)

Challenges of Learning GNNs with DP: Why Not DP-SGD?

- Exploding Sensitivity
  - With a K-layer GNN, each node affects the embedding of all the nodes in its K-hop neighborhood
  - $O(D^K)$ gradient terms change at once (D is maximum degree)
- Inference Privacy
  - GNNs query the graph structure during inference
  - Private information leaks at inference, even with a private model

Challenges of Learning GNNs with DP: Why Not DP-SGD?

- Exploding Sensitivity
  - With a K-layer GNN, each node affects the embedding of all the nodes in its K-hop neighborhood
  - $O(D^K)$ gradient terms change at once (D is maximum degree)
- Inference Privacy
  - GNNs query the graph structure during inference
  - Private information leaks at inference, even with a private model

DP-SGD cannot be directly applied to GNNs

Our Approach: Aggregation Perturbation

- **Aggregation Perturbation:** adding noise to output of the aggregation step
  - Prevents the exploding sensitivity problem by composing differentially private aggregation steps
  - Ensures inference privacy

Our Approach: Aggregation Perturbation

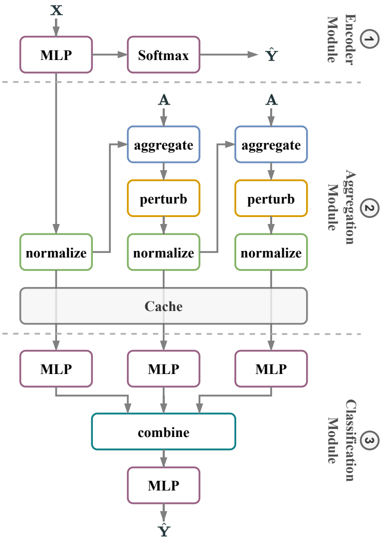
- **Aggregation Perturbation:** adding noise to output of the aggregation step
  - Prevents the exploding sensitivity problem by composing differentially private aggregation steps
  - Ensures inference privacy
- Applying aggregation perturbation to the conventional GNNs is **costly**
  - Every forward pass of the model consumes privacy budget
  - The excessive noise results in poor performance

Our Approach: Aggregation Perturbation

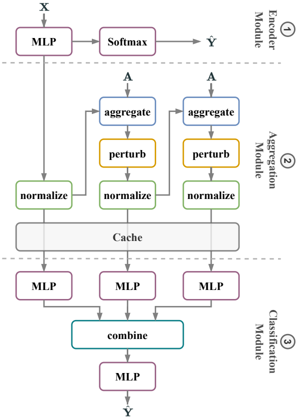
- **Aggregation Perturbation:** adding noise to output of the aggregation step
  - Prevents the exploding sensitivity problem by composing differentially private aggregation steps
  - Ensures inference privacy
- Applying aggregation perturbation to the conventional GNNs is **costly**
  - Every forward pass of the model consumes privacy budget
  - The excessive noise results in poor performance

We decouple the aggregation operations from the model parameters to maintain the privacy budget

GNN with Aggregation Perturbation (GAP)

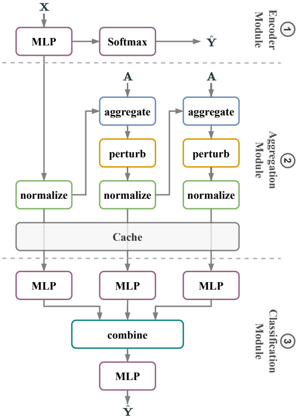


✓ Edge-level DP



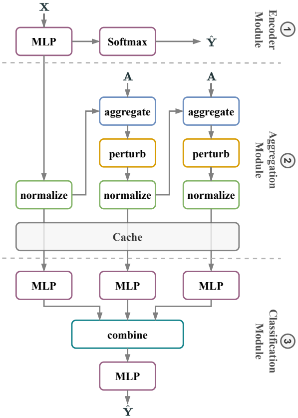
GAP's Advantages

- ✓ Edge-level DP
- ✓ Node-level DP through combination with DP-SGD
  - For bounded-degree graphs

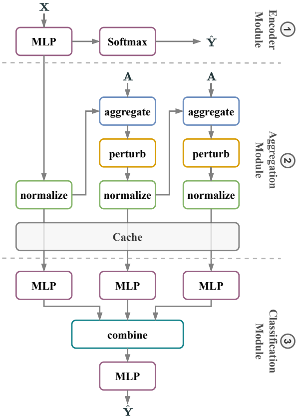


GAP's Advantages

- ✓ Edge-level DP
- ✓ Node-level DP through combination with DP-SGD
  - For bounded-degree graphs
- ✓ Multi-hop aggregations



- ✓ Edge-level DP
- ✓ Node-level DP through combination with DP-SGD
    - For bounded-degree graphs
- ✓ Multi-hop aggregations
- ✓ Zero-cost inference privacy



Experiment Settings

- **Task:** Node Classification
- **Baselines:** MLP, GraphSAGE

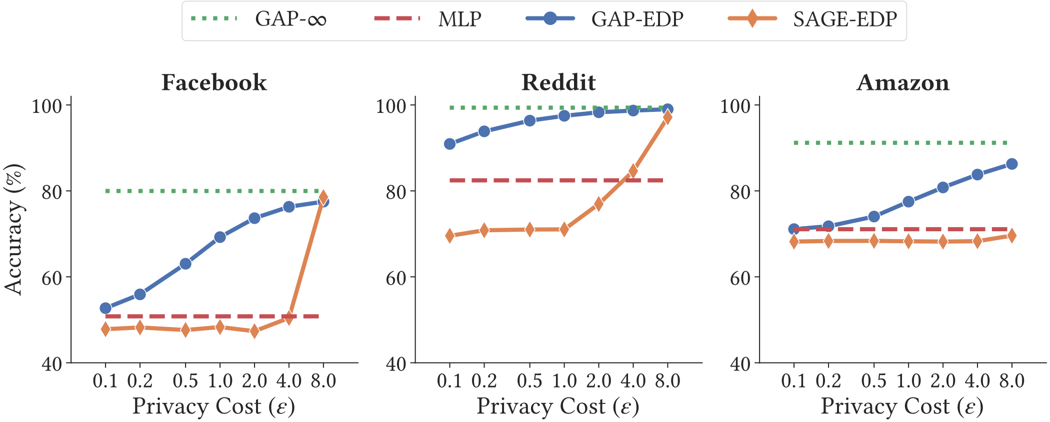
Dataset	Classes	Nodes	Edges	Features	Med. Degree
Facebook	6 Year	26,406 User	2,117,924 Friendship	501	62
Reddit	8 Community	116,713 Post	46,233,380 Mutual User	602	209
Amazon	10 Category	1,790,731 Product	80,966,832 Mutual Purchase	100	22

Comparison of Non-Private Methods

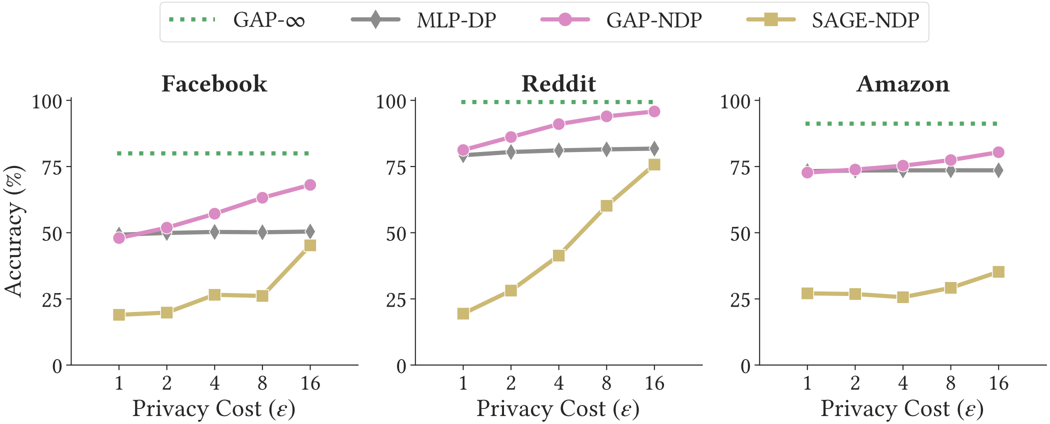
Accuracy of Non-Private Methods

Method	Facebook	Reddit	Amazon
GAP-∞	80.0 ± 0.48	**99.4 ± 0.02**	91.2 ± 0.07
SAGE-∞	**83.2 ± 0.68**	99.1 ± 0.01	**92.7 ± 0.09**

Edge-Level DP Accuracy-Privacy Trade-Off



Node-Level DP Accuracy-Privacy Trade-Off



Mean AUC of node-level membership inference attack.

Dataset	Method	$\epsilon = 1$	$\epsilon = 2$	$\epsilon = 4$	$\epsilon = 8$	$\epsilon = 16$	$\epsilon = \infty$
Facebook	GAP-NDP	50.16	50.25	50.61	51.11	52.66	81.67
Reddit	GAP-NDP	50.04	50.39	51.20	52.23	52.54	54.97
Amazon	GAP-NDP	50.06	50.23	50.54	51.53	51.72	66.68

Conclusion

- GNNs leak private information
  - They are vulnerable to privacy attacks
- Implementing DP in GNNs is challenging
  - Exploding sensitivity
  - Inference privacy
- Our Differentially Private GNN: GAP
  - Ensures both edge-level and node-level DP
  - Supports multi-hop aggregations
  - Provides inference privacy

Thank You!

Questions: sina.sajadmanesh@epfl.ch

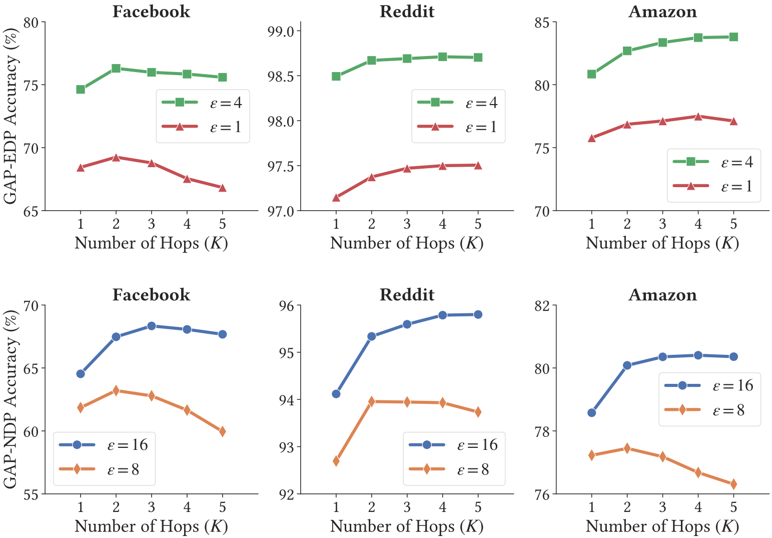
Code: github.com/sisaman/GAP

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006).
Calibrating noise to sensitivity in private data analysis.
In *Theory of cryptography conference*, pages 265–284. Springer.

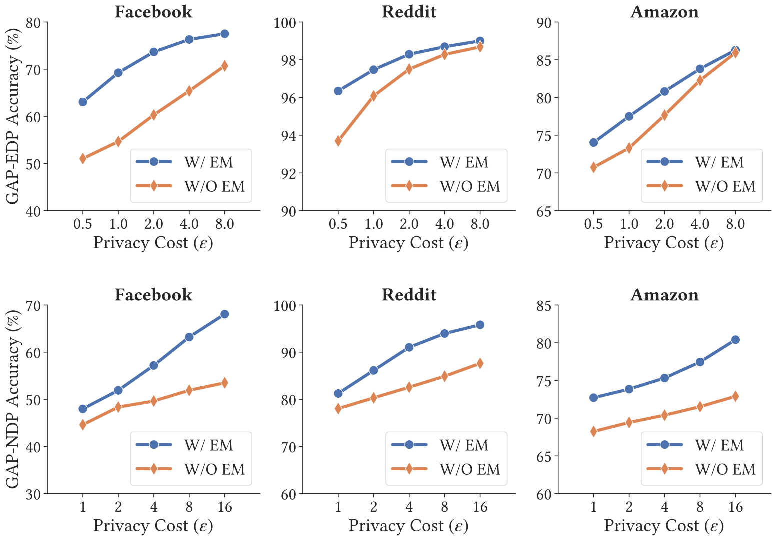
He, X., Jia, J., Backes, M., Gong, N. Z., and Zhang, Y. (2021).
Stealing links from graph neural networks.
In *30th {USENIX} Security Symposium ({USENIX} Security 21)*.

Olatunji, I. E., Nejdl, W., and Khosla, M. (2021).
Membership inference attack on graph neural networks.
arXiv preprint arXiv:2101.06570.

Effect of the Number of Hops



Effect of the Encoder Module



Effect of the Maximum Degree