Graph Neural Network-based Clustering Enhancement in VANET for Cooperative Driving

Hang Hu, Myung J. Lee

Lab of Advanced Wireless Networks

Department of Electrical Engineering

City University of New York (CUNY), City College

Outline

- Motivation and Background
- GNN-based Clustering Algorithm
- Evaluation Results
- Conclusions and Future Works

Outline

- Motivation and Background
- GNN-based Clustering Algorithm
- Evaluation Results
- Conclusions and Future Works

Motivation and Background

- VANET clustering benefits [1]
  - Transportation network: traffic capacity, safety, and cooperative driving enhancement
  - Air Environment: fuel efficiency improvement and exhaust emissions reduction

Platooning improves traffic capacity, fuel economy, and safety

Cooperative driving facilitates safety in autonomous driving

[1] D. Jia, K. Lu, J. Wang, X. Zhang, and X. Shen, “A Survey on Platoon-Based Vehicular Cyber-Physical Systems,” IEEE Communications Surveys & Tutorials, 18(1), 2016, pp.263-284.

Challenges of VANET Clustering

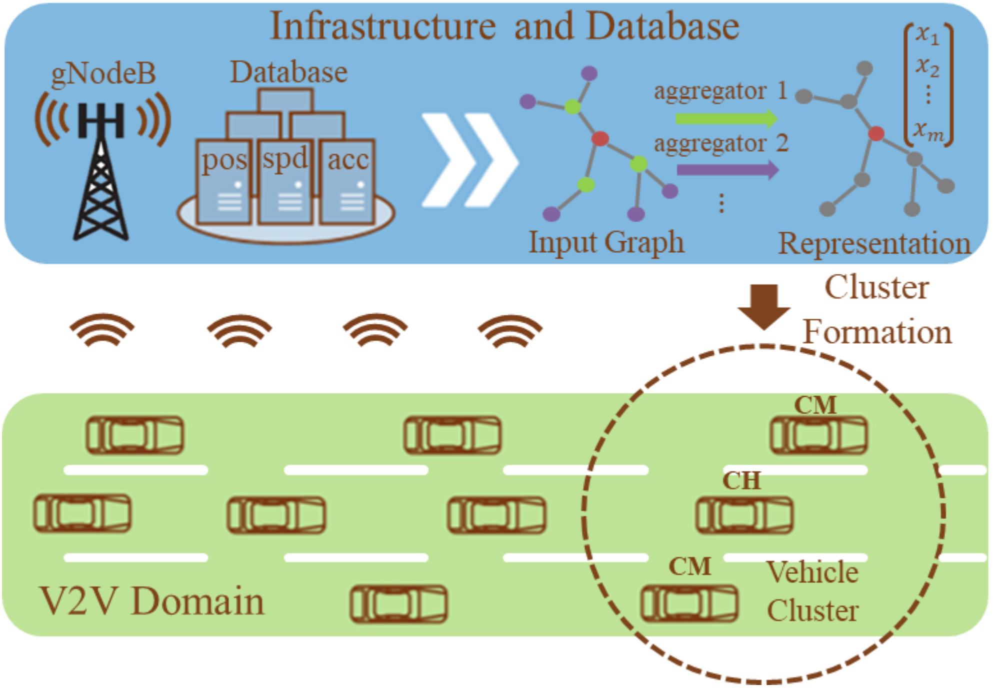
- State-of-the-art VANET clustering Algorithms
  - Distributed clustering approaches
    - High control message overhead
    - Manually select hyperparameters
    - Not intelligent and learnable
  - Machine learning-based clustering approaches
    - Leverage only a single feature
    - Not learnable

Clustering Algorithm	Weight-based Clustering [1]	ML-based Clustering [2]	GNN-based Clustering
Formation Strategy	Distributed	Centralized	Centralized
Complexity	High	Low	Low
Information Utilization	Node Feature	Node or Graph Feature	Node and Graph Feature
Learnability	No	No	Yes

[1] A. Bello Tambawal, et al., "Enhanced weight-based clustering algorithm to provide reliable delivery for VANET safety applications," PloS one, 14(4), 2019, e0214664.
[2] G. Liu, et al., "Enhancing clustering stability in VANET: A spectral clustering based approach. China Communications," 17(4), 2020, pp. 140-151.

Goal and Proposal Summary

- Goal: enhance the vehicle system's stability and optimize the average lifetime of all clusters
- Why we choose GNN
  - Fits naturally to solve clustering type of graph problem
  - Uses both node feature and graph information
  - Centralized approach and offloads the computation to BS
  - It's the very first attempt to apply GNN to solve the clustering problem in VANET

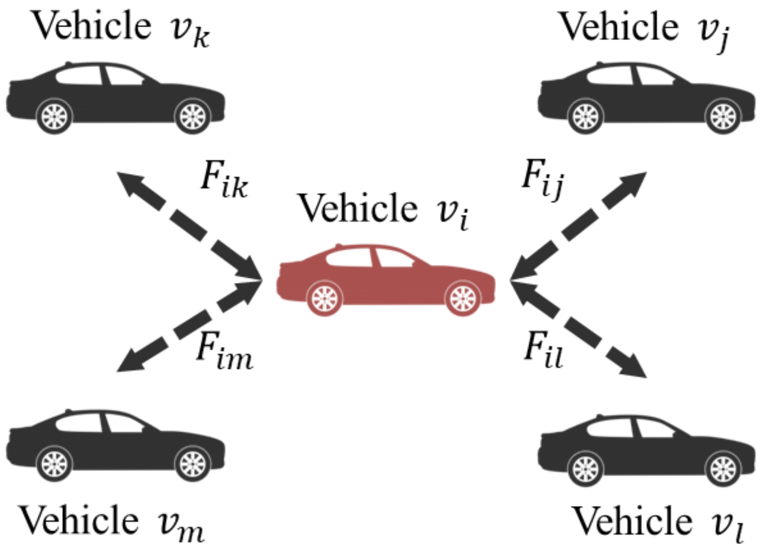


Outline

- Motivation and Background
- GNN-based Clustering Algorithm
- Evaluation Results
- Conclusions and Future Works

Graph Construction

- Graph is modeled by force-directed algorithm [1]
  - Relative force among vehicle interconnection weighs the similarity between the movement patterns of 2 vehicles
  - The greater the positive forces among nodes are, the more similar the moving pattern is



$$F_{ijx} = k_{ijx}\frac{q_i q_j}{D_{ij}^2}; F_{ijy} = k_{ijy}\frac{q_i q_j}{D_{ij}^2} \quad (1)$$

$$D_{ijx}(t) = x_i - x_j; D_{ijx}(t+dt) = x_i + dx_i - x_j - dx_j \quad (2)$$

$$D_{ijy}(t) = y_i - y_j; D_{ijy}(t+dt) = y_i + dy_i - y_j - dy_j \quad (3)$$

$$k_{ijx} = \frac{1}{1 + |D_{ijx}(t+dt) - D_{ijx}(t)|dt} \quad (4)$$

$$k_{ijy} = \frac{1}{1 + |D_{ijy}(t+dt) - D_{ijy}(t)|dt} \quad (5)$$

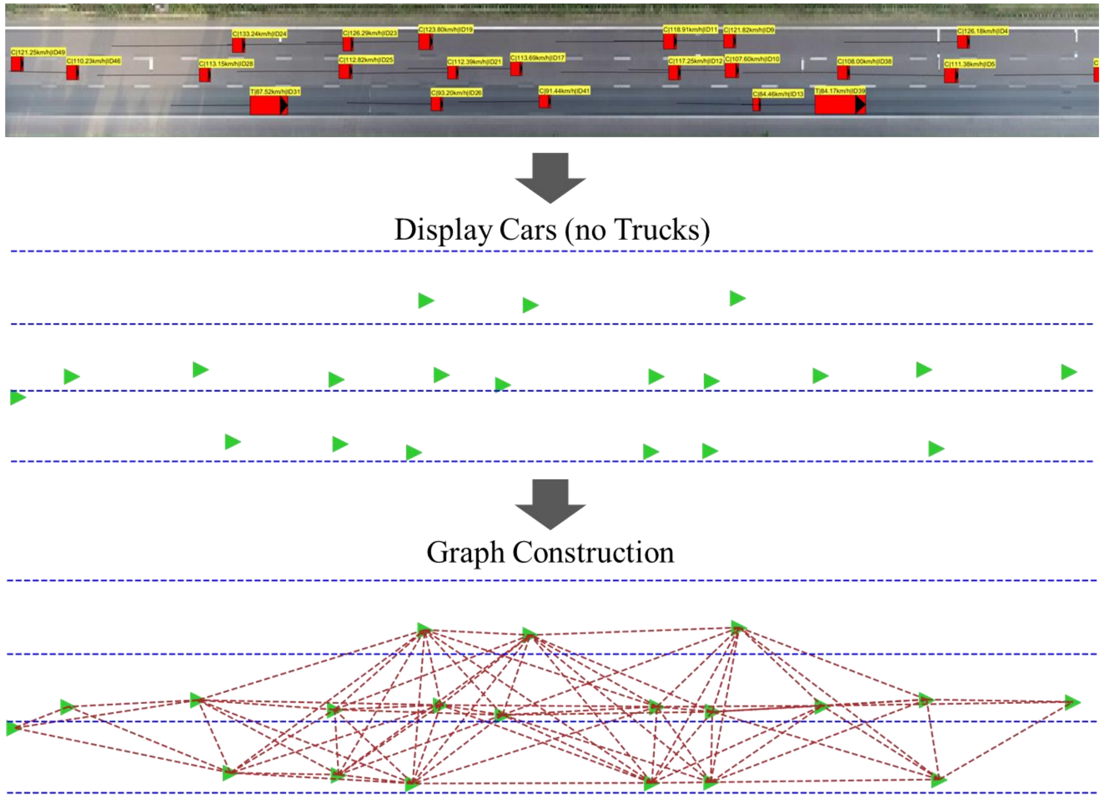
$$q_i = q_j = \begin{cases} R - D_{ijx}(t), & \text{if } D_{ijx}(t) \le D_{ijx}(t+dt) \\ R + D_{ijx}(t), & \text{if } D_{ijx}(t) > D_{ijx}(t+dt) \end{cases} \quad (6)$$

$$||\boldsymbol{F}_{ij}||_2 = \sqrt{F_{ijx}^2 + F_{ijy}^2} \quad (7)$$

[1] L. A. Maglaras, and D. Katsaros, “Distributed clustering in vehicular networks,” IEEE 8th international conference on wireless and mobile computing, networking and communications (WiMob), 2012, pp. 593-599.

Graph Construction Visualization

- Open-source highD traffic dataset [1]
  - Naturalistic vehicle trajectory recordings on German highways
  - Cover about 420 m road segment. The median duration of visible vehicles is13.6s
  - Traffic information includes vehicle trajectory, type, size, etc. The Position error is typically less than 10 cm
- Apply force-directed algorithm to highD dataset to customize our graph dataset

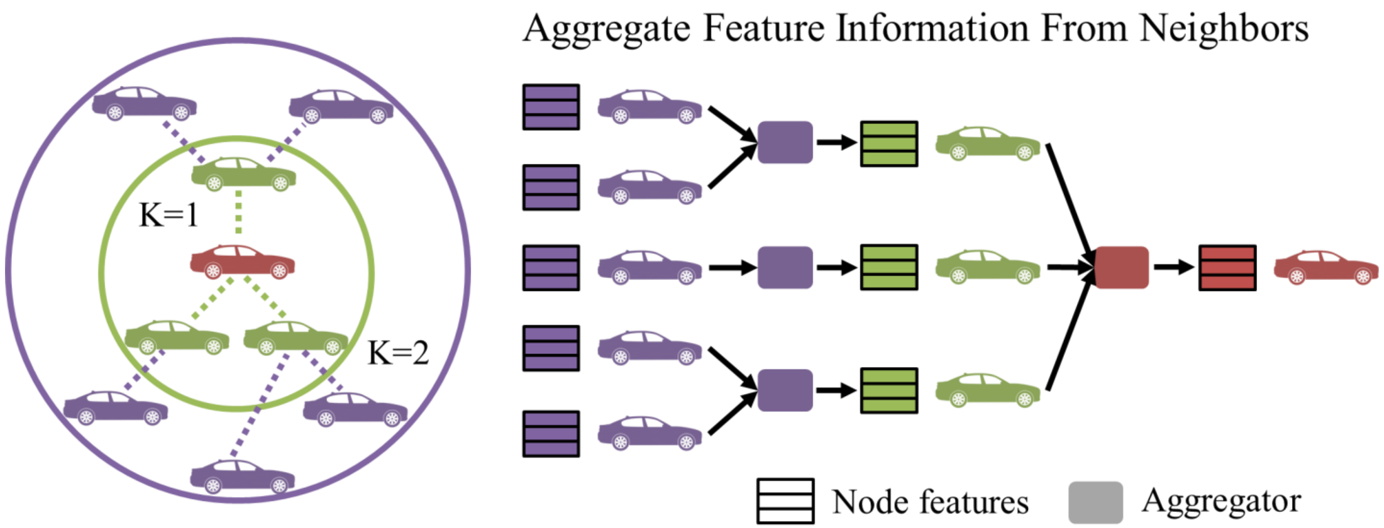


[1] R. Krajewski, et al., “The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems,” International Conference on Intelligent Transportation Systems (ITSC), 2018, pp. 2118-2125.

Design of GNN Clustering Algorithm

- Spatial-based convolutional graph neural network
  - Input is vehicle feature and graph; output is useful node embedding
  - Apply SAGE convolutional layer [1]
  - Apply Mean aggregator and search depth $K = 2$

$$h_i^k = \sigma\left(W^k \cdot \frac{1}{|N_i|} \sum_{j \in N_i} (h_j^{k-1} \cdot F_{ij})\right)$$



[1] W. L. Hamilton, R. Ying, and J. Leskovec, “Inductive representation learning on large graphs,” In Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 1025-1035.

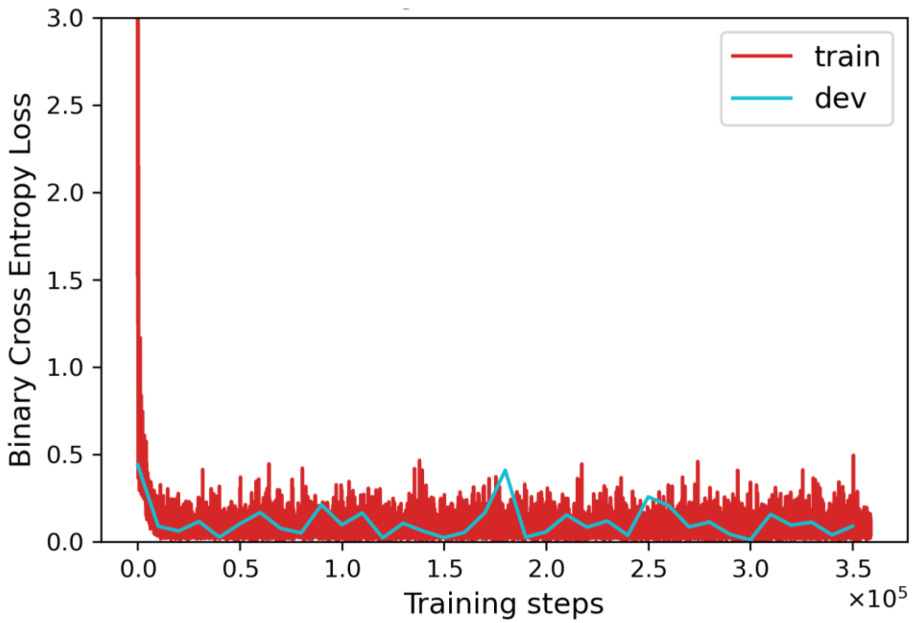
Model Training

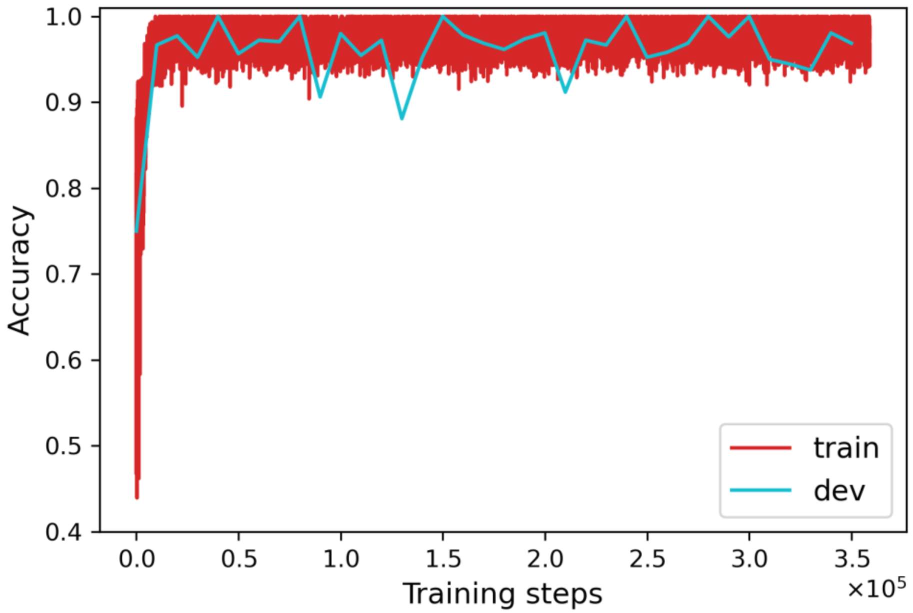
- Graph-based loss function in unsupervised learning
  - Guidance is the edges existent or non-existent
  - Forward propagation
    - Calculate node representations via GNN model
    - Apply node embeddings to compute pairwise probability among nodes
  - Backward propagation
    - Calculate loss and update model parameters via stochastic optimization

$$J_G(z_i) = -\sum_{i,j \in V} \left(y_{ij} \log\left(\hat{y}_{ij}\right) + \left(1 - y_{ij}\right) \log(1 - \hat{y}_{ij})\right)$$

Model Training Results

- 1000 training graphs (train:dev=9:1) and 210 testing graphs





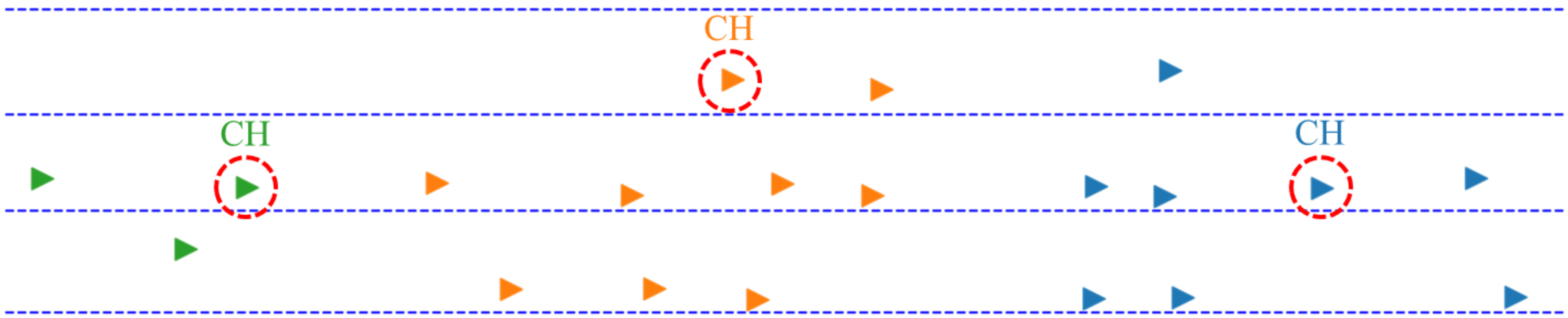
	Training Set	Validation Set	Testing Set
Loss	0.041	0.084	0.063
Acc	0.986	0.969	0.978

*Trained GNN model can learn useful and effective node representation

Clustering Visualization

- GNN-based clustering steps
  - Apply the trained GNN model on a graph to calculate node embeddings
  - Obtain the clustering results by running k-means on node embeddings
- A visual example of clustering results

$$n = \left\lceil \frac{L}{2R} \right\rceil$$

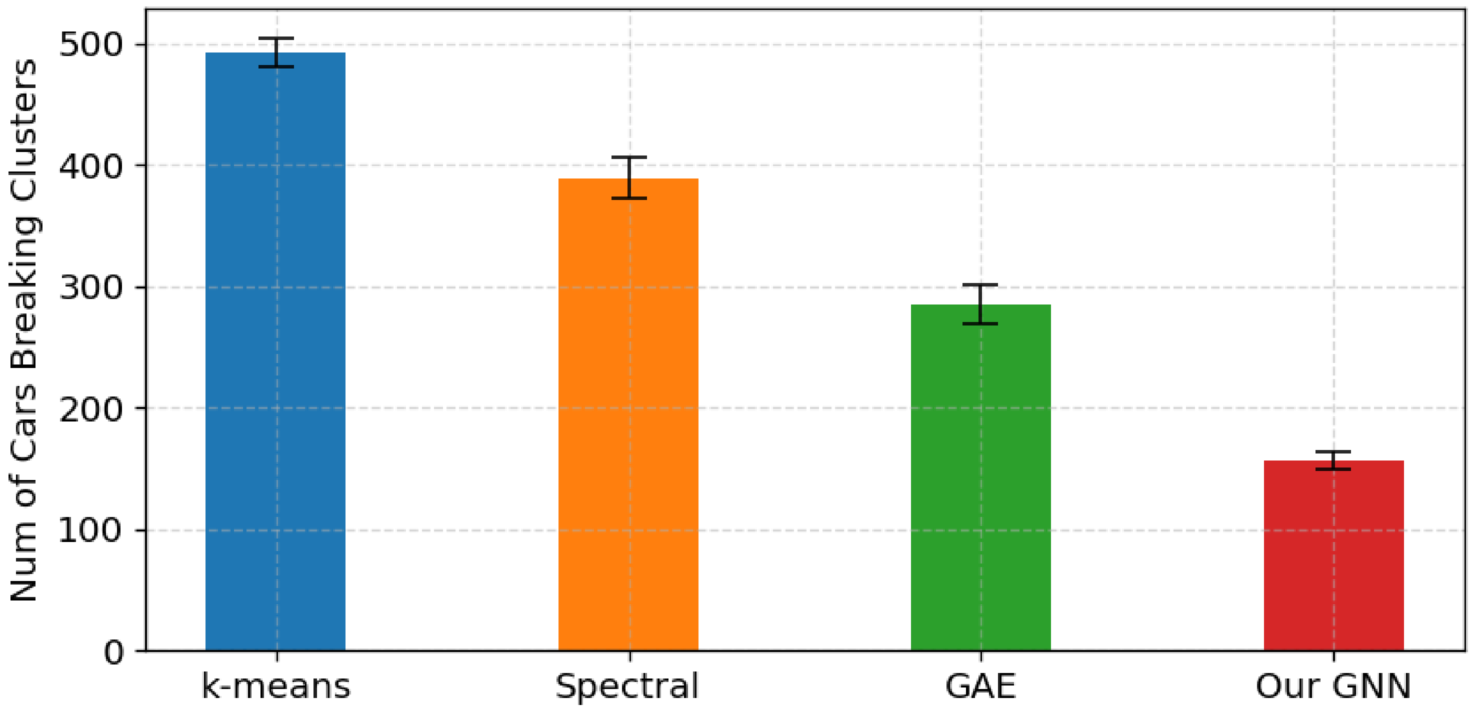


Outline

- Motivation and Background
- GNN-based Clustering Algorithm
- Evaluation Results
- Conclusions and Future Works

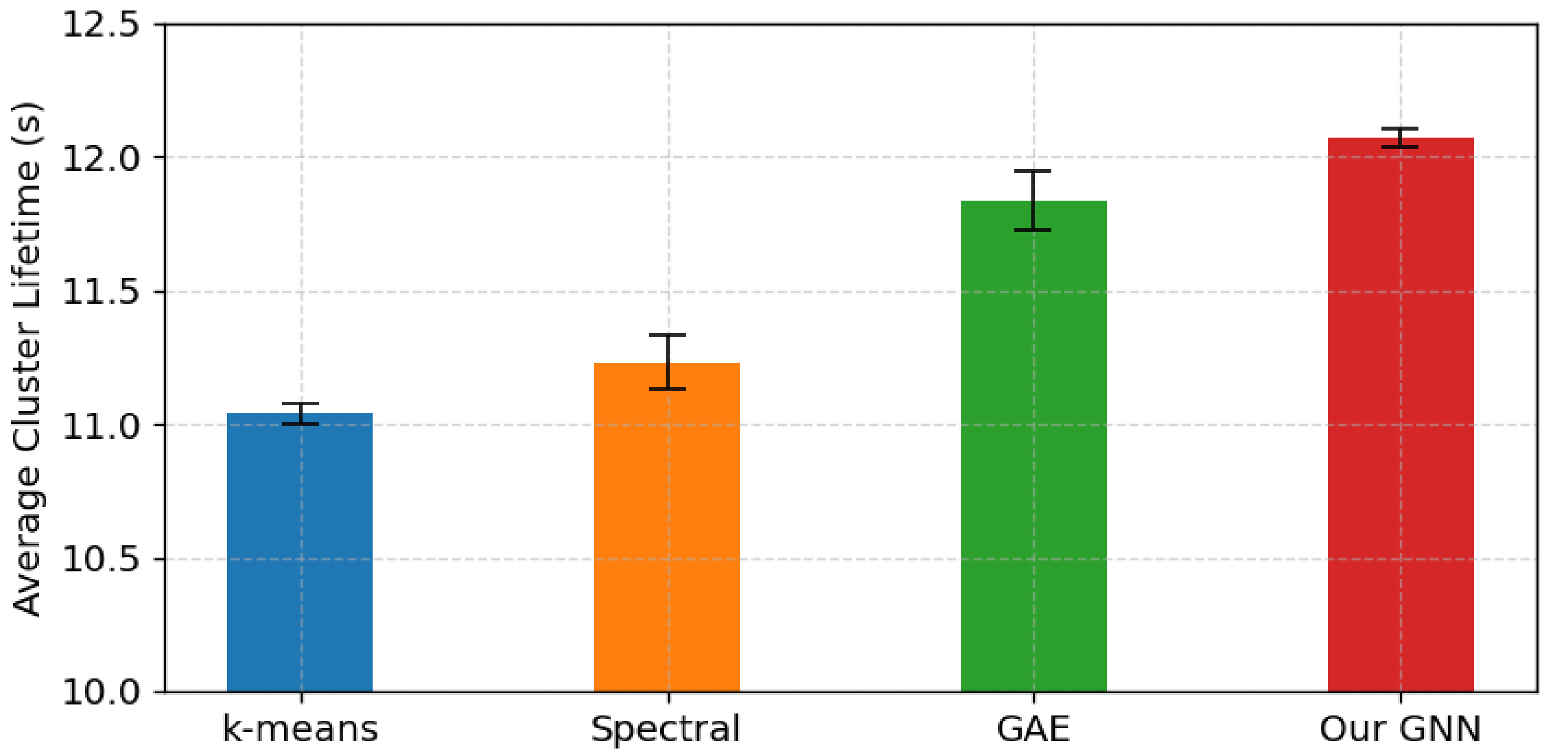
Performance Evaluation

- GNN-based algorithm corresponds to the minimum number of vehicles breaking the initial clusters



Average Cluster Lifetime Evaluation

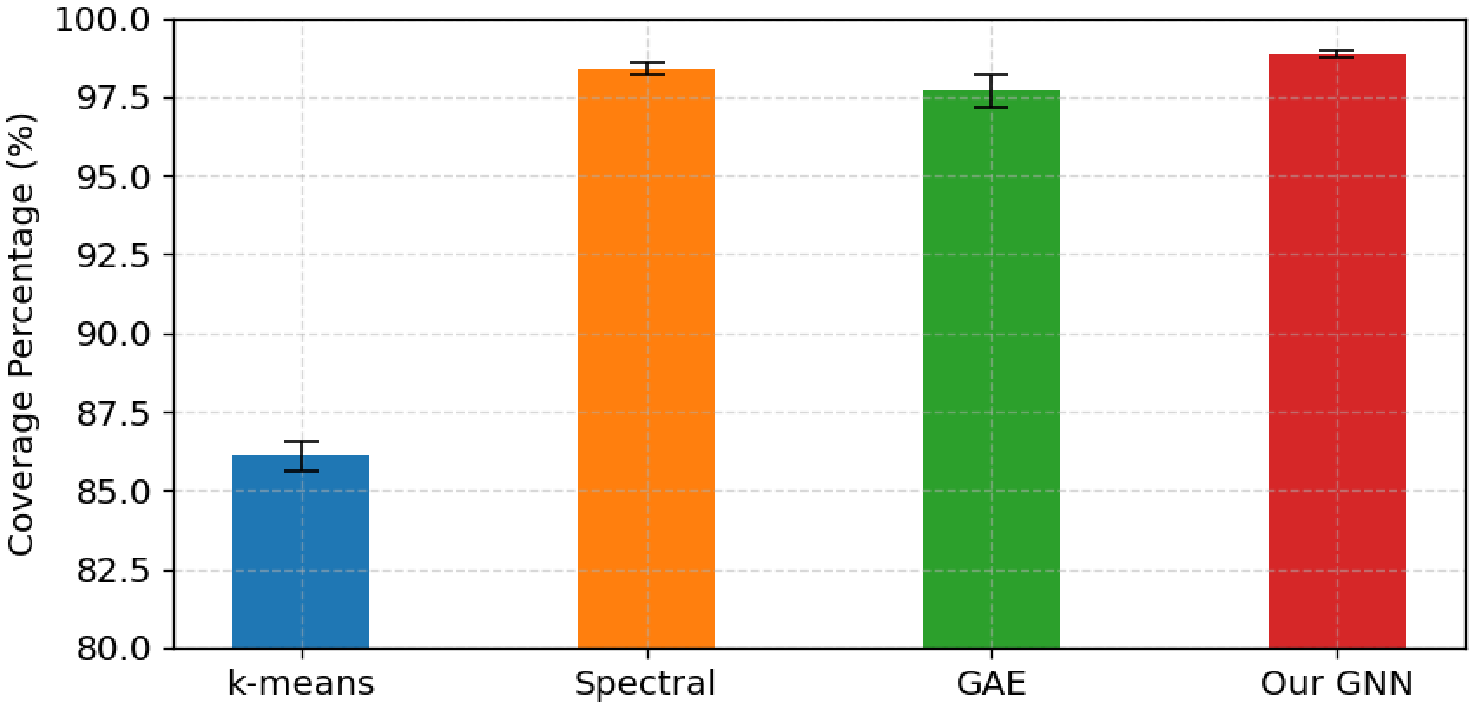
- Average cluster lifetimes of GNN-based method is 12.069±0.037s with confidence 95%. Compared with baseline algorithms, it has the longest average cluster lifetime



Coverage Percentage Evaluation

- Cluster efficiency of GNN-based algorithms achieve 98.927±0.111% with confidence 95%

$$CP = \frac{N - N_{Iso}}{N}$$



Outline

- Motivation and Background
- GNN-based Clustering Algorithm
- Evaluation Results
- Conclusions and Future Works

Conclusions and Future Works

- High performance GNN-based VANET clustering on open-source highD traffic dataset
  - Average cluster lifetime (12.069±0.037s)
  - Coverage percentage (98.927±0.111%)
- Future works
  - Study other traffic scenarios like urban environment
  - Simulation of Urban Mobility (SUMO) for long-term performance

Acknowledgment

- This research is supported by NSF IRNC Grant No. 2029295, and Grove School of Engineering, The City University of New York, City College.

Thank You!

THE CITY UNIVERSITY OF NEW YORK