ITU AI/ML in 5G Challenge Global Round in Japan

ITU-ML5G-PS-032-KDDI

On Failure Classification Based on GNN in IP Core Networks by NFV-Based Test Environment.

Nara Institute of Science and Technology

Takanori HARA and Kentaro FUJITA

Team name: naist-lsm

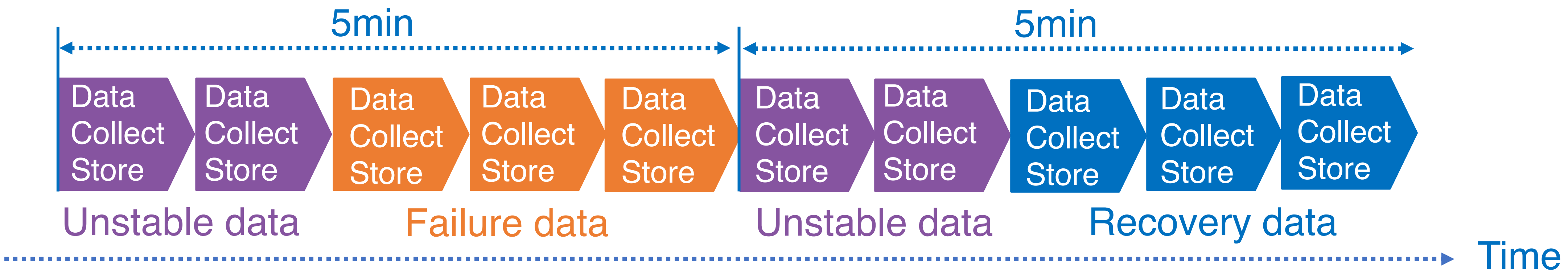
Problem Statement

- **With profileration of 5G mobile network, mobile operators have to continuously provide the stable and high-quality internet services**
- **To tackle the unexpected defect in the IP core network, machine learning based network operations can achieve to operate automatically and rapidly as well as to reduce operation expenditures**
- **The dataset at border gateway routers includes network status such as normal and a failure, mis-operation, and normal or abnormal labels.**
- **We create a model for detecting and/or classifying the network status of a failure utilizing the dataset and evaluate the performance using the proposed model.**

Dataset

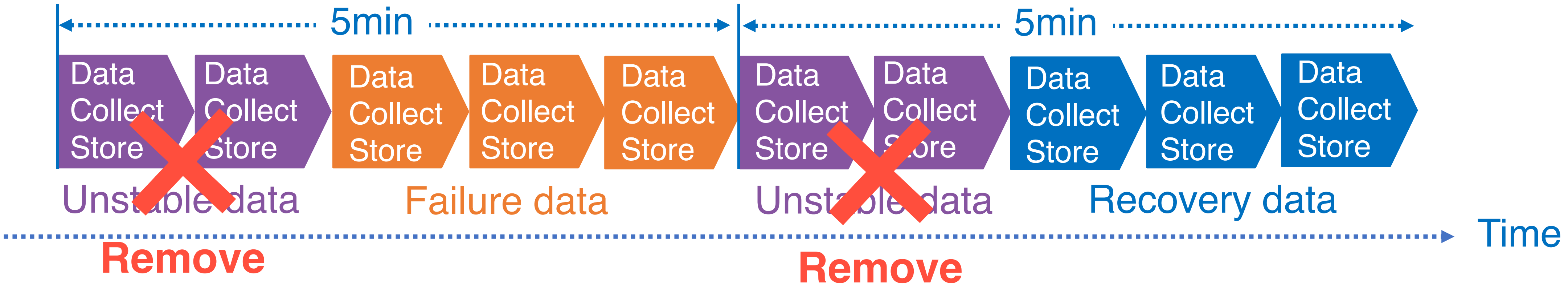
Category	Filename	Description
Label	Failure	Event date and event types
Data	Virtual infrastructure	Performance monitoring data sets on instances and virtual network functions gathered from OpenStack ceilometer
	Physical infrastructure	Performance monitoring data sets gathered from the physical server under OpenStack
	Network device	Performance monitoring information and BGP route information gathered from NEs under the virtual IP network

- **The dataset generator [Kawasaki+20] is used**
- **These dataset are partly unstable due to the data collection principles**

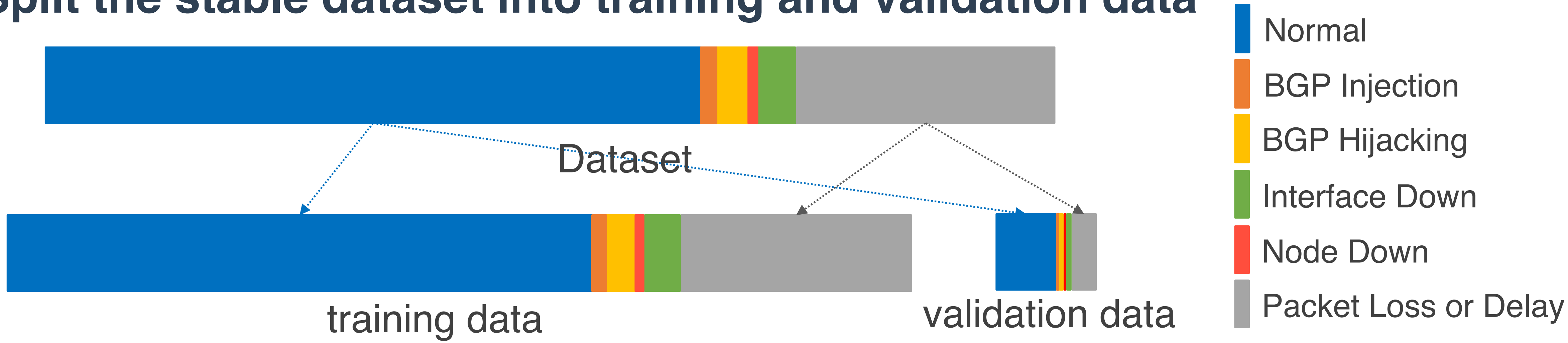


Data Preprocessing

- **Retrieve the stable dataset from the dataset**
  - The dataset includes unstable data due to the data collection principles



- **Split the stable dataset into training and validation data**



Related Work

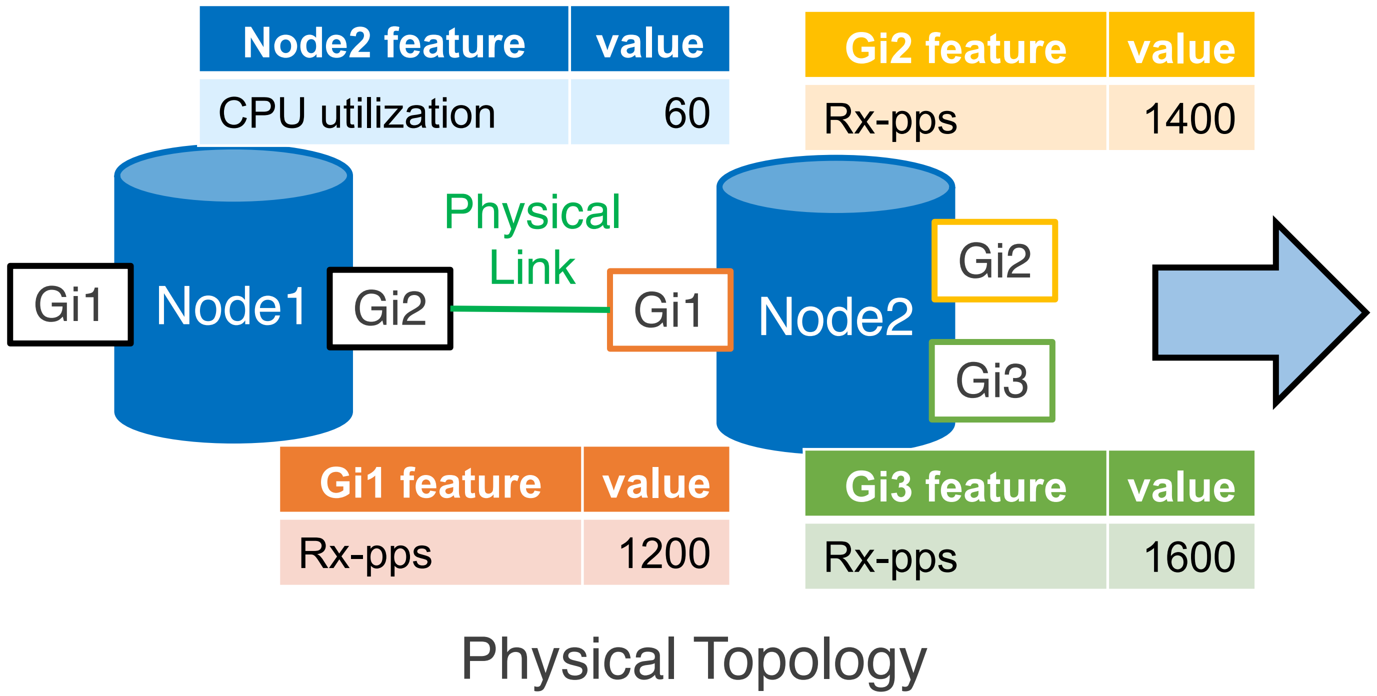
- **Network Fault Analysis**
  - Network fault classification using machine learning [Kawasaki+20]
  - Network traffic faults classification using clustering [Qader+17]
  - BGP-related failure classification/detection [Al-Musawi+17, Cho+19]
- **Graph Neural Networks (GNNs)**
  - NN-based ML which enables explicit topology embedding in learning model [T.N.Kipf+17, Geyer17]
  - Network traffic classification [Zheng+19]
  - Estimation of communication delay between node pairs [Suzuki+20]
  - Channel allocation for wireless LANs [Nakashima+20]

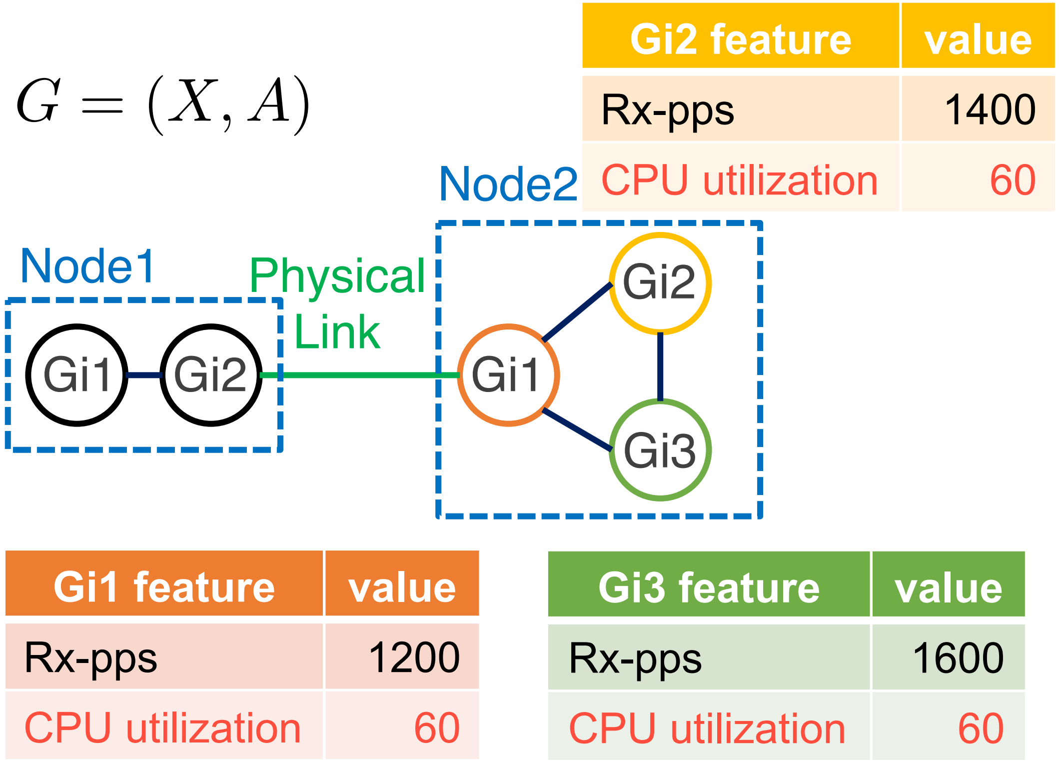
Motivation

- **Initial step toward the realization of the failure classification in IP core networks with the explicit topology embedding**
- **This project investigates**
  - Potential of the supervised graph classification with graph convlutional networks (GCNs) for detecting and classifying the network status
    - i.e., route information failures, single point failures, paket loss/delay
  - How the GCN contributes to the performance improvement compared with the other machine learning based schemes
    - XGBoost, Random forest, SVM, MLP

Graph Transpotation

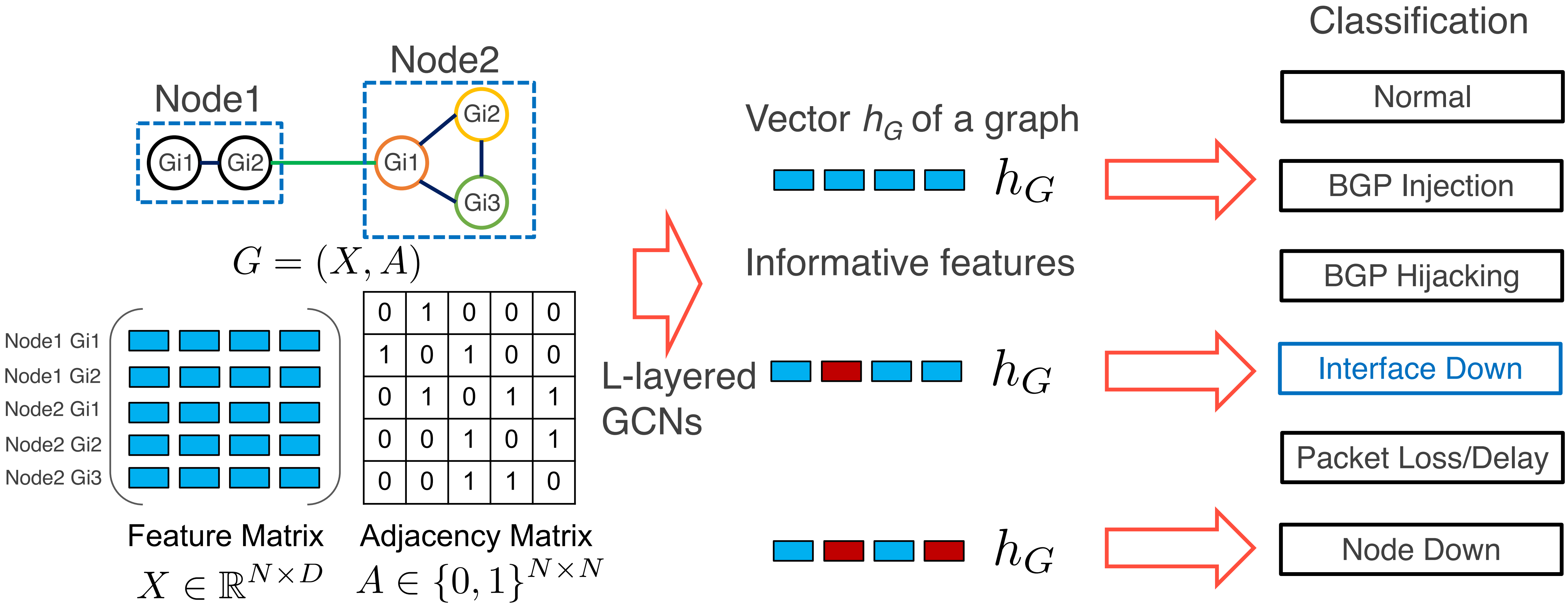
- **We transform the physical topology into the graph** $G = (X, A)$
  - where X denotes a feature matrix and A denotes an adjacency matrix
- **We use the seven types of node features**
  - CPU utilization, interface condition, tx/rx-pps, network incoming/outgoing packet rate, prefix activity





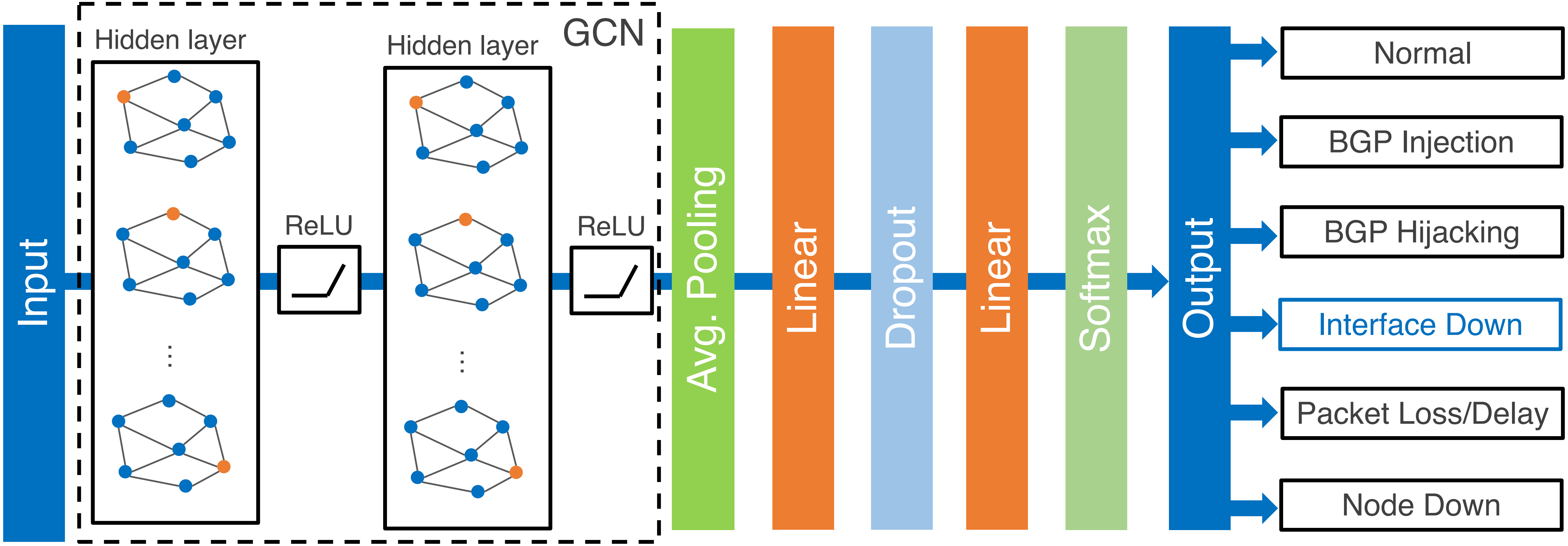
Supervised Graph Classification with GCN

- **Supervised Graph Classification with the GCN**
  - Predict the failure type from features of an entire graph



Our Model

- **This classifier finds six failure categories**
- **We use seven types of data as the inputs for our model**

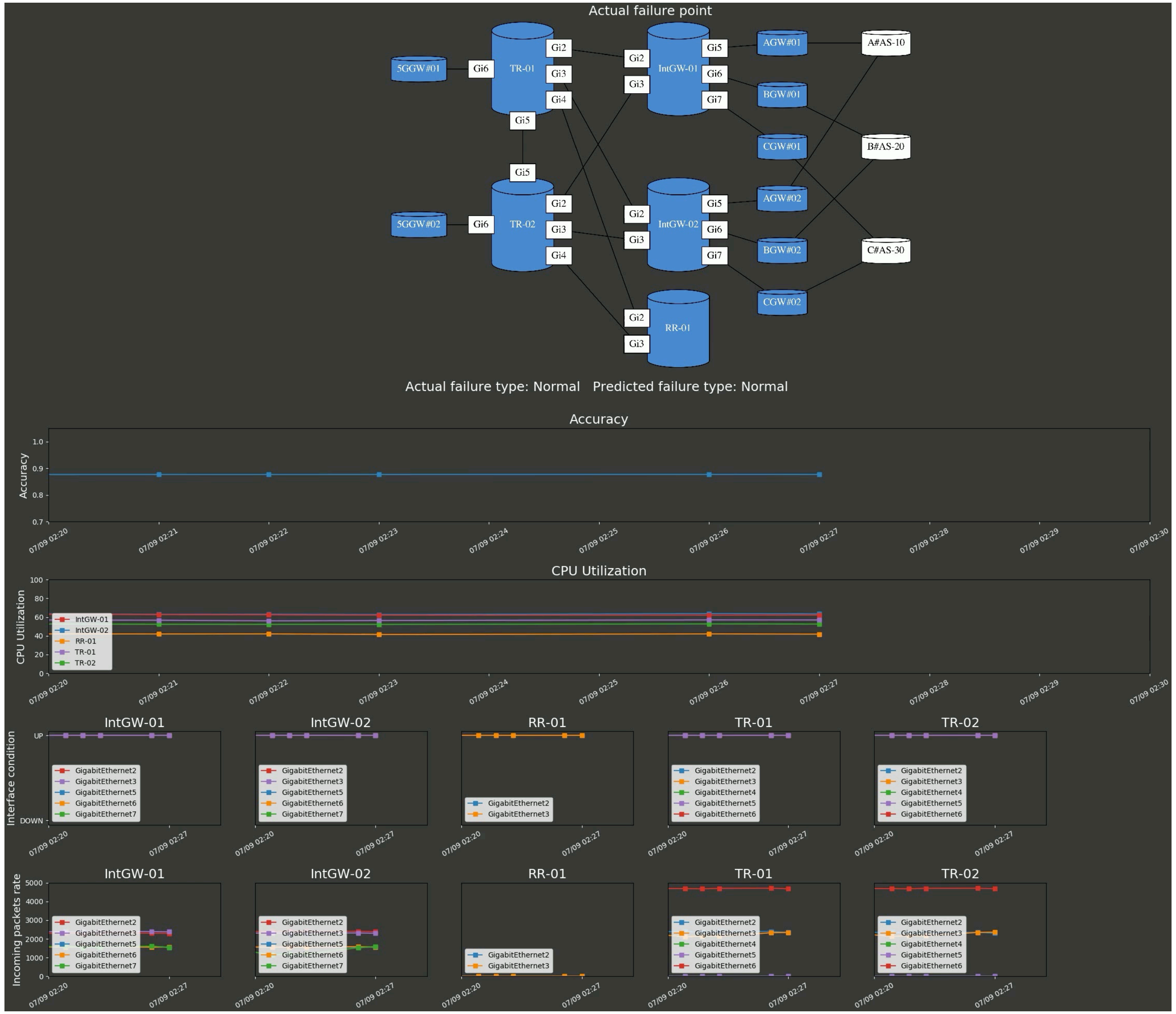


Performance Comparison

scheme	criteria	normal	BGP hijacking	BGP injection	node down	interface down	packet loss/delay	accuracy	inference time [ms]
		Failure type							
XGBoost	precision	0.93	0.68	1.00	0.92	0.88	0.75	0.89	20
	recall	0.91	0.99	0.99	1.00	0.85	0.70		
	f1-score	0.92	0.81	0.99	0.96	0.87	0.73		
RF	precision	0.90	0.70	1.00	1.00	0.91	0.75	0.87	8
	recall	0.91	0.99	0.99	1.00	0.95	0.64		
	f1-score	0.91	0.82	0.99	1.00	0.93	0.69		
SVM	precision	0.99	0.54	0.25	0.96	0.60	0.62	0.84	1319
	recall	0.82	1.00	0.95	1.00	0.96	0.86		
	f1-score	0.90	0.70	0.40	0.98	0.74	0.72		
GCN	precision	0.89	0.97	0.98	0.99	0.99	0.98	0.91	274
	recall	0.99	0.70	0.96	1.00	1.00	0.62		
	f1-score	0.94	0.82	0.97	1.00	1.00	0.76		
MLP	precision	0.90	0.97	0.97	1.00	0.96	0.66	0.87	17
	recall	0.91	0.71	0.92	0.99	0.96	0.73		
	f1-score	0.91	0.82	0.95	1.00	0.96	0.69		

- **The GCN becomes higher accuracy compared with other schemes**
- **The GCN contributes to the performance improvement for detecting packet loss/delay**
    - The dataset does not include the explicit information of the packet loss/delay
- **Allowable inference time: 274 [ms]**

Brief Demonstration



Conclusion

- **The supervised graph classification with the GCN**
  - Becomes higher accuracy compared with other schemes
  - Contributes to the performance improvement for detecting packet loss/delay
- **Future Work**
  - Accuracy improvement for BGP-related failures
    - To adopt good features, e.g., information on as-path
  - Heterogeneous graph
    - To consider not only physical topology but also logical one
  - Semi-supervised graph classification
    - Failure classification from the observation of small samples

Reference

- [Kawasaki+20] J. Kawasaki, G. Mouri, and Y. Suzuki, “Comparative Analysis of Network Fault Classification Using Machine Learning,” in Proc. of *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*. Budapest, Hungary: IEEE, Apr. 2020, pp. 1–6.
- [Qader+17] K. Qader, M. Adda, and M. Al-kasassbeh, “Comparative Analysis of Clustering Techniques in Network Traffic Faults Classification,” *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 5, no. 4, pp. 6551–6563, 2017.
- [Al-Musawi+17] B. Al-Musawi, P. Branch, and G. Armitage, “BGP Anomaly Detection Techniques: A Survey,” *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 377–396, Firstquarter 2017.
- [Cho+19] S. Cho, R. Fontugne, K. Cho, A. Dainotti, and P. Gill, “BGP Hijacking Classification,” in *Proc. of 2019 Network Traffic Measurement and Analysis Conference (TMA)*, Jun. 2019, pp. 25–32.
- [T.N.Kipf+17] T. N. Kipf and M. Welling, “Semi-Supervised Classification with Graph Convolutional Networks,” *arXiv:1609.02907 [cs, stat]*, Feb. 2017.
- [Geyer17] F. Geyer, “Performance Evaluation of Network Topologies using Graph-Based Deep Learning,” in *Proc. of the 11th EAI International Conference on Performance Evaluation Methodologies and Tools*, ser. VALUETOOLS 2017. New York, NY, USA: Association for Computing Machinery, Dec. 2017, pp. 20–27.
- [Zheng+19] J. Zheng and D. Li, “GCN-TC: Combining Trace Graph with Statistical Features for Network Traffic Classification,” in *Proc. of ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6.
- [Suzuki+20] T. Suzuki, Y. Yasuda, R. Nakamura, and H. Ohsaki, “On Estimating Communication Delays using Graph Convolutional Networks with Semi- Supervised Learning,” in *Proc. of 2020 International Conference on Information Networking (ICOIN)*. Barcelona, Spain: IEEE, Jan. 2020, pp. 481–486.
- [Nakashima+20] K. Nakashima, S. Kamiya, K. Ohtsu, K. Yamamoto, T. Nishio, and M. Morikura, “Deep Reinforcement Learning-Based Channel Allocation for Wireless LANs With Graph Convolutional Networks,” *IEEE Access*, vol. 8, pp. 31 823–31 834, 2020.