
Signal Processing for Machine Learning

Carrson C. Fung

Associate Professor

Signal Processing for Machine Learning and Applications Group (SPMLAP)

Communication Electronics and Signal Processing Lab (CommLab)

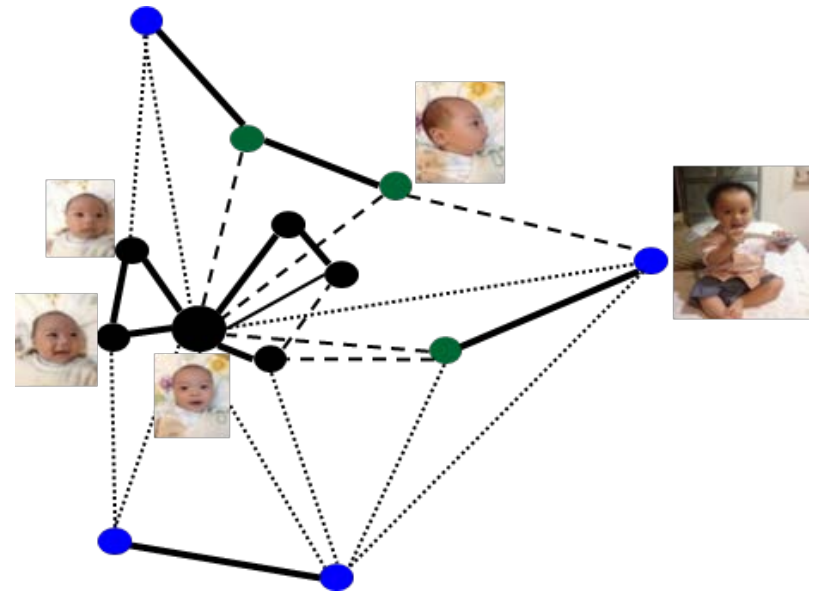
Institute of Electronics

National Chiao Tung University



SPMLAP Group

- Research focuses on
 - Self-supervised federated **and** distributed learning
 - **Graph** signal processing for graph learning **and** graph neural network
 - **6G**: Model-based DNN design for intelligent reflective surface (IRS)
- Summer internship **abroad** for Ph.D. candidates are strongly encouraged (possible for outstanding M.S. students)
 - M.S. and 1st-year Ph.D. students encouraged to apply for the industrial Ph.D. program (教育部產學博計畫)
- Group members
 - 1 Ph.D., 7 M.S., 1 U.G.
- Possible to get jobs with skills you learned in my group
 - **Google** (Taipei and Mountain View), **Qualcomm** (San Diego), **Amobee** (Hsinchu), **Realtek** (Hsinchu), **Umbo Computer Vision**, Netapp (Los Angeles)



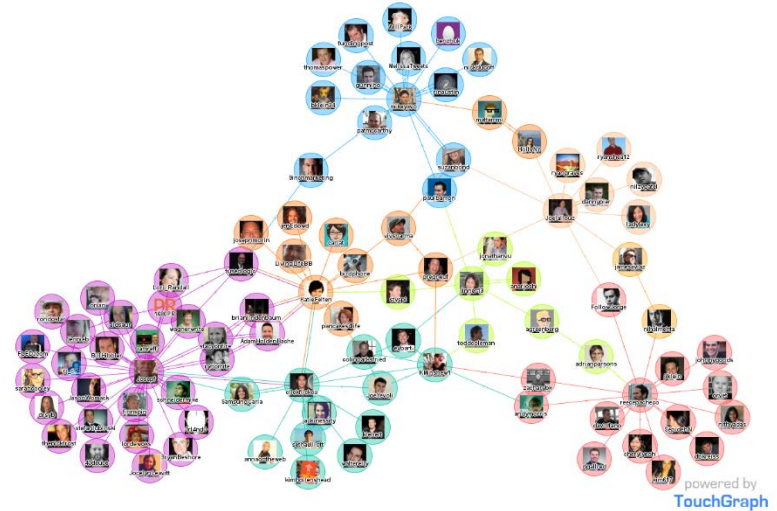
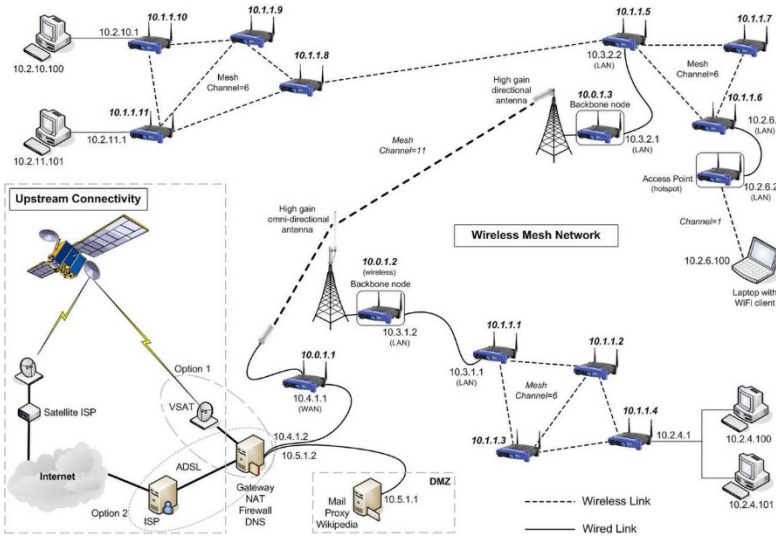
Networked Data



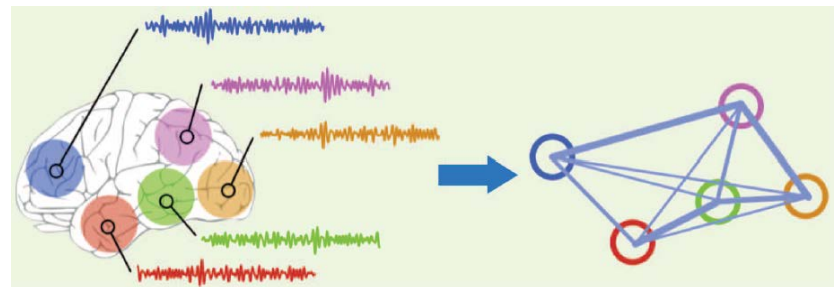
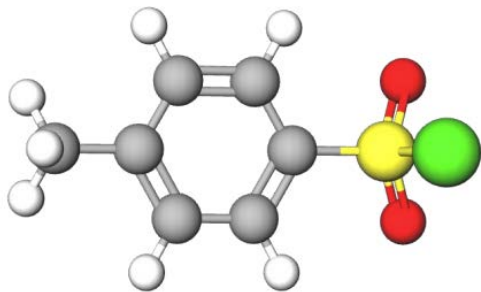
But the distance of each pixel is the same

→ Processing can be done using traditionally signal processing techniques (convolution, filtering, transform)

Networked non-Euclidean Data



3D Molecular Graph



Why Learn the Non-Euclidean Distance?

Node classification problem

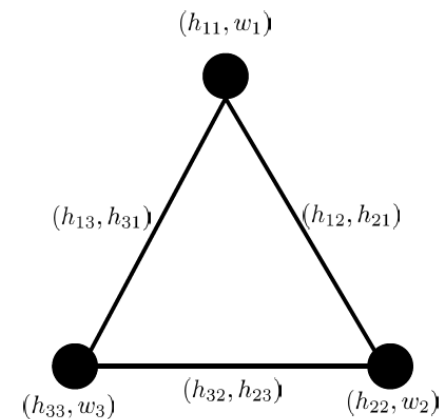
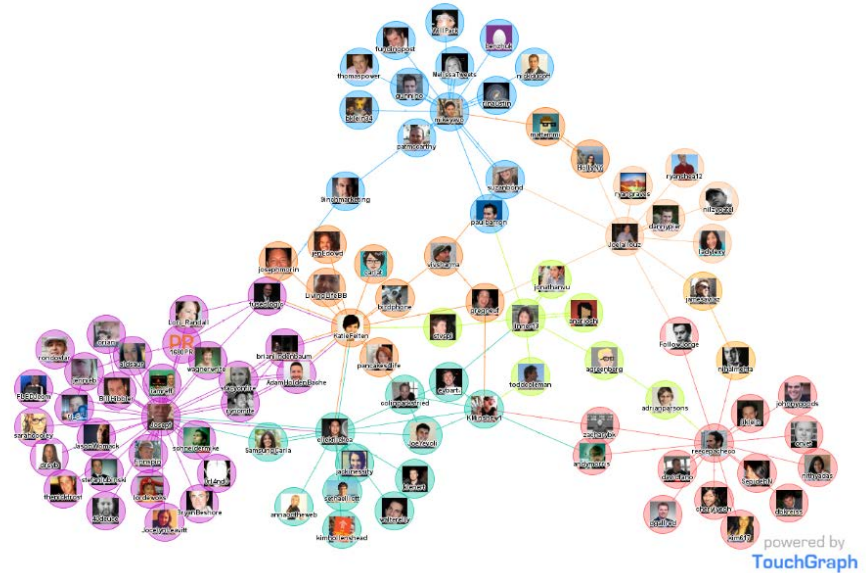
- **Applications:** Community discovery (e.g. Netflix, Pinterest) and offer targeted recommendations to different groups (prediction)

Graph classification problem

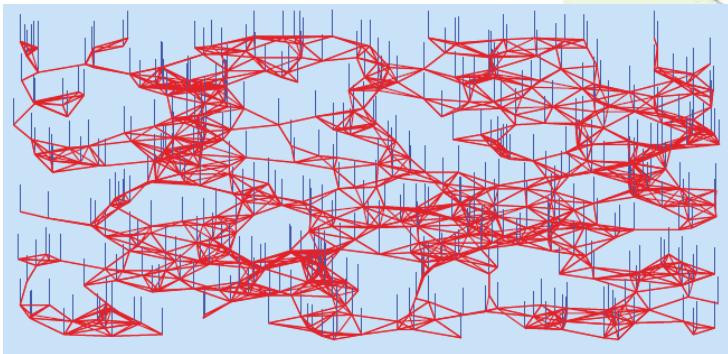
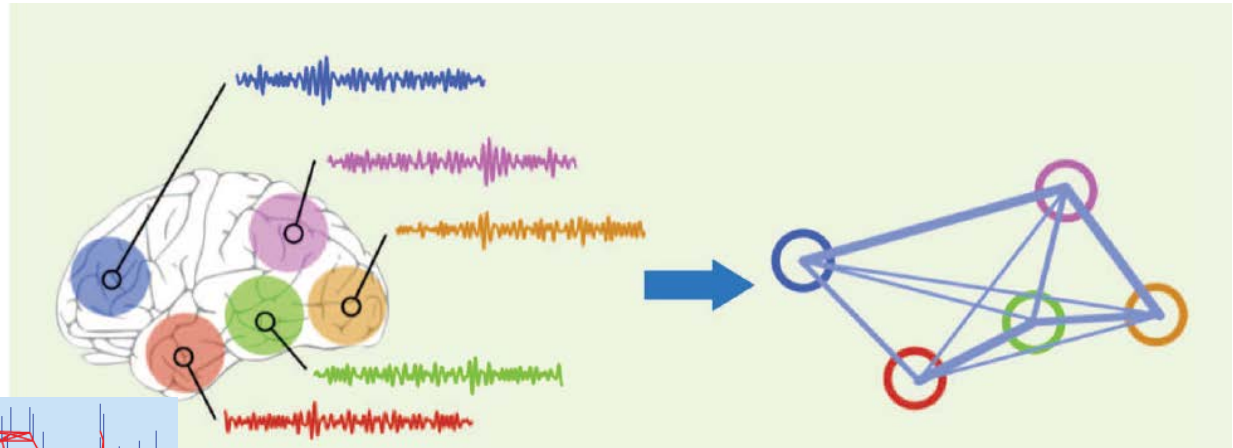
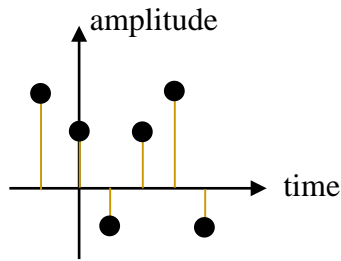
- **Application:** Compare brain graphs across different subjects that have labels (e.g. Alzheimer's disease) may identify if the subject without label may have Alzheimer's

Node regression

- **Application:** Building an interference graph and identifying the power needed for transmission in a multi-transmitter and multi-receiver environment

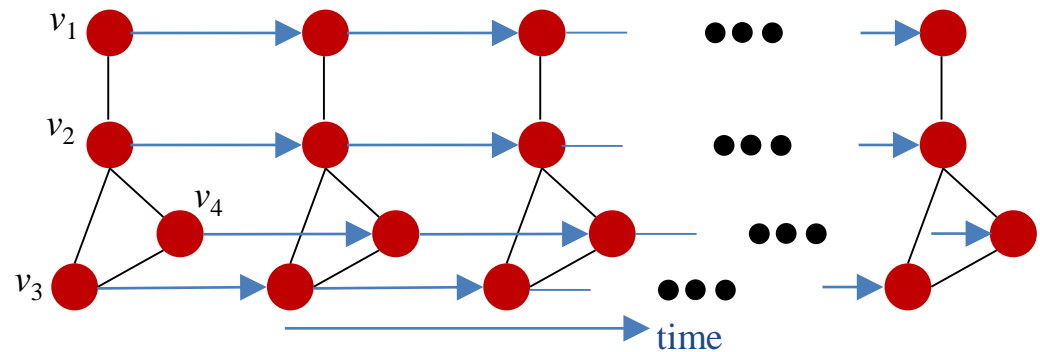


Online Graph Learning (Graph Tracking)

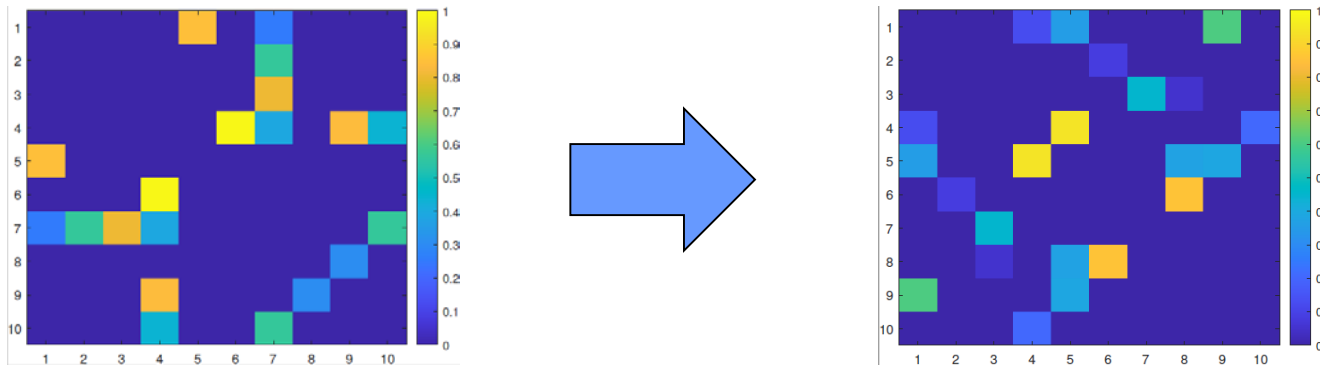


- Learn the **connectome** of the brain **over time**: map of the neural connections in the brain
 - **Structural** – white matter connection
 - **Functional** – statistical interdependencies between physiological time series from different brain regions
 - **Effective connectivity** – shows cause and effect of one neural element on another

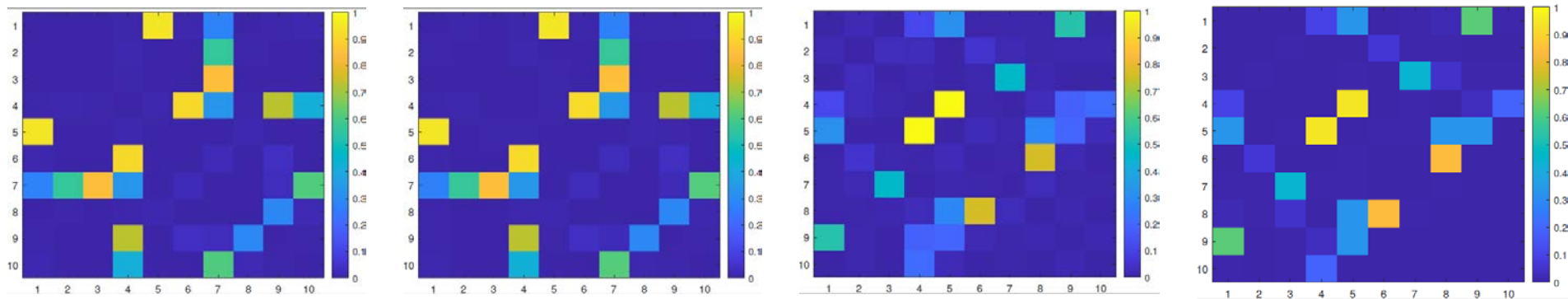
Online graph learning (graph tracking)



Some Graph Tracking Results

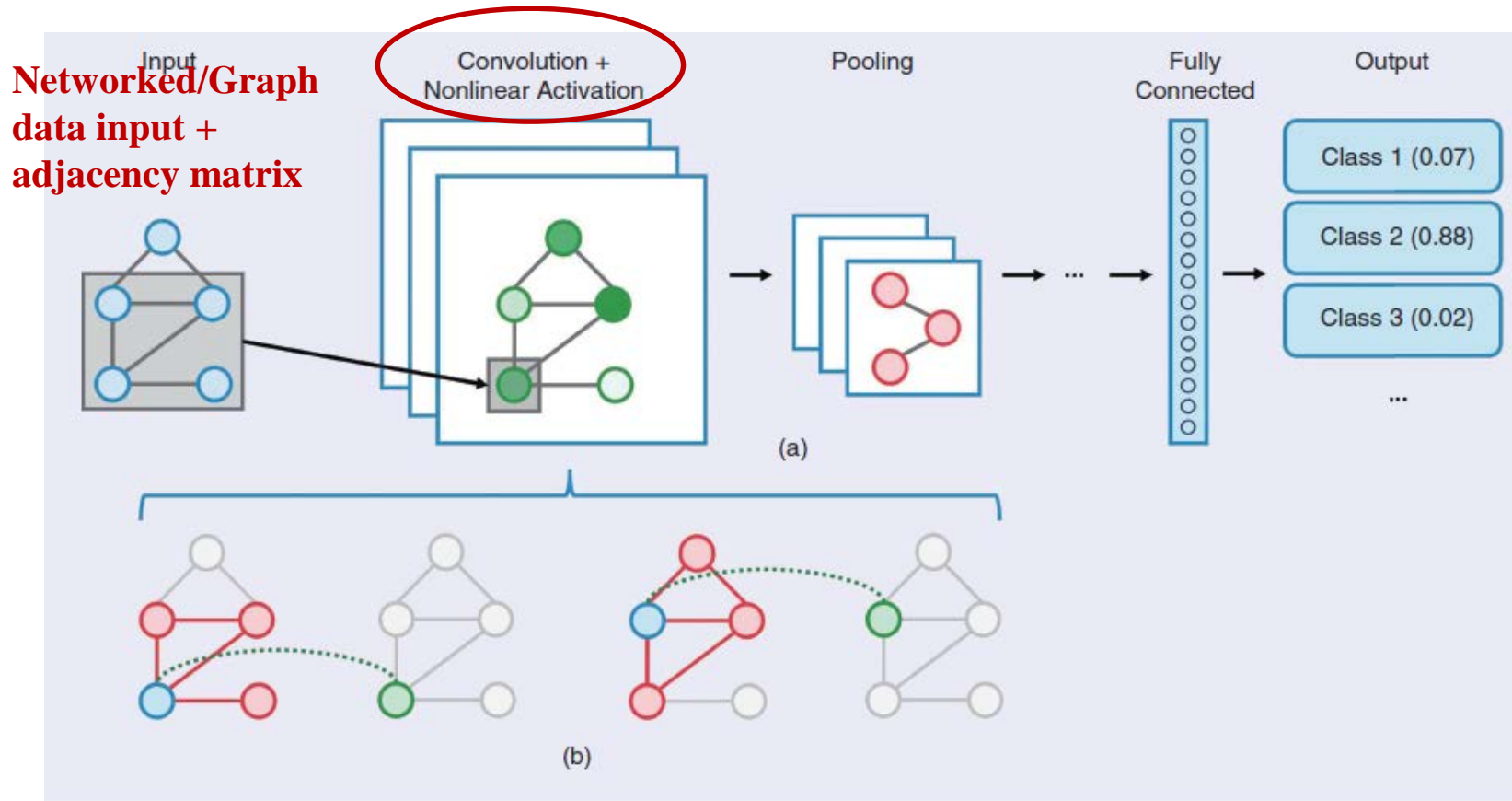


Ground truth: 1st (left) graph, 2nd (right) graph. Graph transitioned at $n_0 = 14,000$ sample.



PN-IEKF: (left to right): $n = 13001, 14001, 15001$ and 29001 sample.

Graph Neural Network



Graph classification

- Brain disease classification, e.g. Alzheimer's, Attention Deficit Hyperactive Disorder (ADHD)

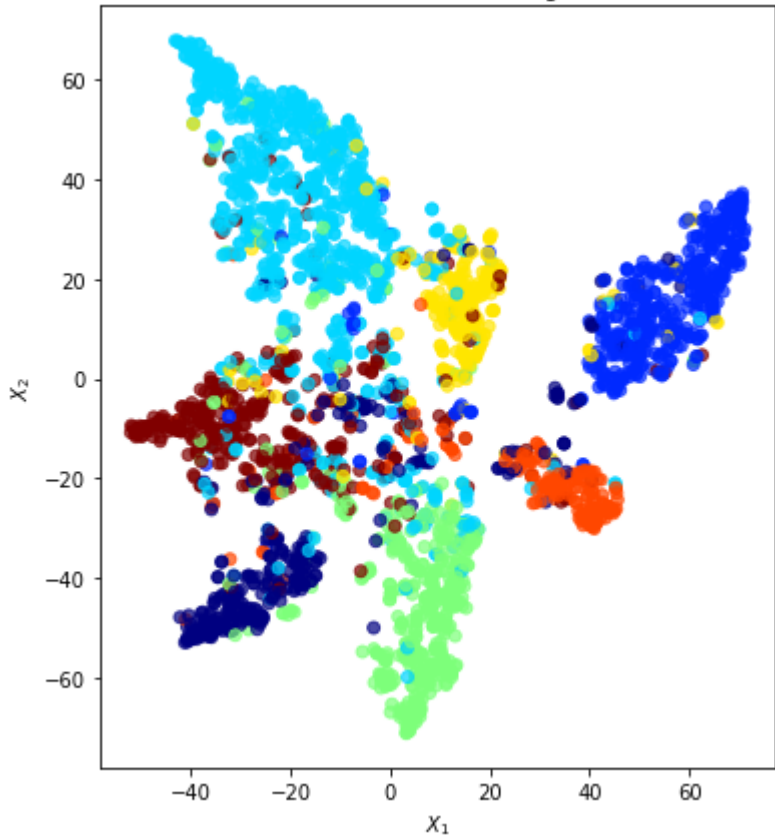
How do we predict certain patient will have Alzheimer's or ADHD?

How should we take into account dynamic graph?

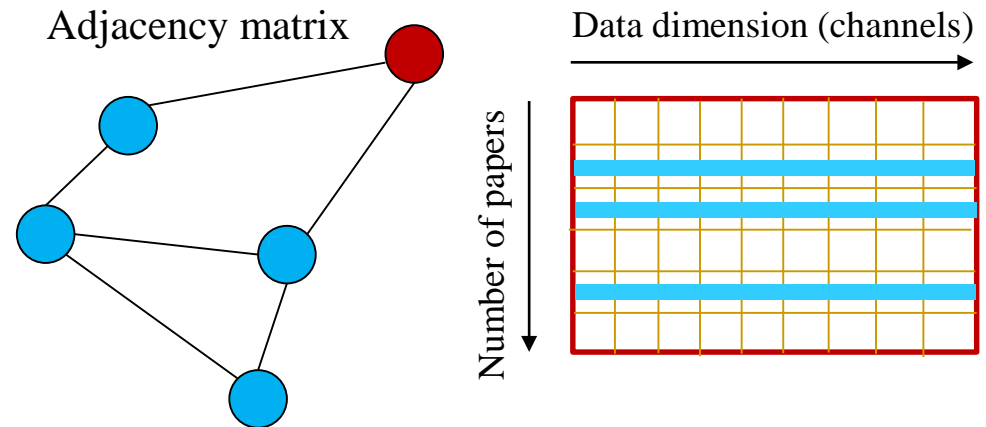


Graph Convolutional Neural (GCN) Network: Cora Dataset

TSNE visualization of GCN embeddings for cora dataset

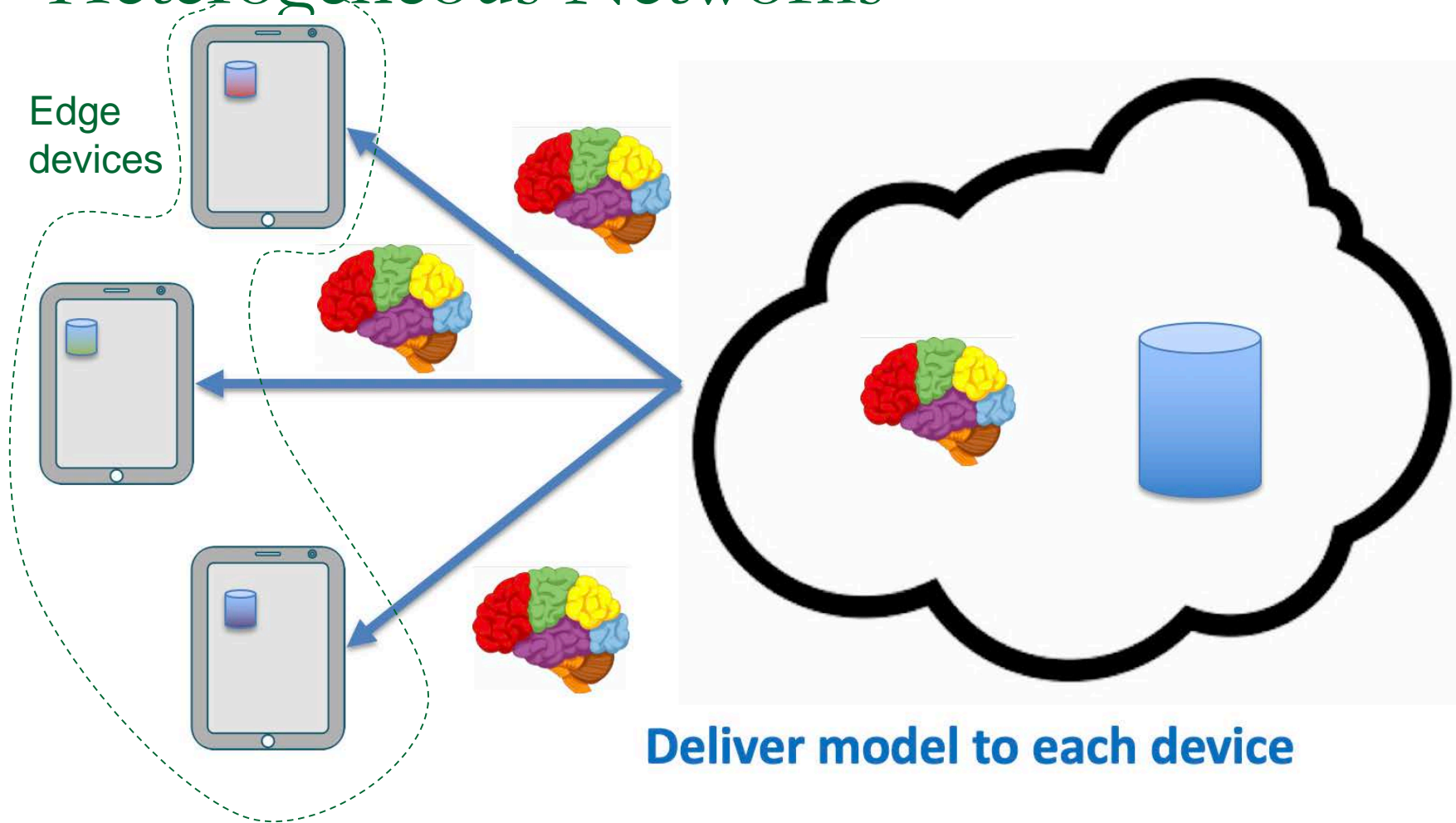


- List of publications in 7 categories (7 classes)
- Graph signal dimension (channels): 1433 (keywords), $\{0,1\}^{1433}$
- Number of papers (vertices): 2708



- Edge weights (adjacency matrix) are known
- Determine labels from 7 classes

Federated and Distributed Learning in Heterogeneous Networks



Federated and Distributed Learning over Graph (Heterogeneous Networks)

- Heterogeneous networks
 - Statistical (data) heterogeneity
 - How will data imbalance and non-IID distributed data affect learning?
 - System heterogeneity
 - How will stragglers affect the learning outcome?
- How do we perform “secure” communications during training?
- How changes in the underlying connections (graph) affect training?
 - E.g. In model training over wireless networks, how the learning strategy should adapt to bad channels?
- What if each device (or group of devices) is training a different (personalized) models?
 - In a vehicular (wireless) network, network of cars can detect (and may classify) different objects near them while working together

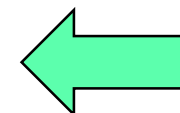


3D mmWave Radar

The screenshot displays the DriveRecorder3 software interface, which is used for recording and analyzing vehicle data. The main window shows a top-down view of a road with two red cars. Yellow lines represent the radar's field of view. Several cyan circles are scattered around the road, indicating detected objects. The interface is divided into several panels:

- Controller:** Shows the current recording status, including the file source (20160929_0300.dr2) and video source (20160929_0300.avi).
- HW Monitor:** Displays hardware status, including a warning "No CAN-HW found!" and the CAN1 channel (1000k).
- TargetList I & II:** Lists detected targets, with "Umrz2GID0" visible in TargetList I.
- TargetDraw & CSV Export:** Provides options for drawing targets and exporting data to CSV.
- Status I & II:** Shows tracking data for two different targets, including device mode, source device, source diagnosis, time, number of objects, and number of messages.
- Video:** Displays a live video feed of a person walking in a courtyard.
- CANDataGrid:** Shows a table of CAN data messages.

#Msgs	Ident	Len	Data bytes [7...0]	CAN Nr	CANCard-Time
99	1F5	8	1E.FF.FF.CE.00.00.00.1	1	1805783210
196	400	7	00.00.00.00.00.00.01	1	1805752948
2778	401	8	87.8C.1C.1C.27.C8.D1.1	1	1805766407
2778	402	7	04.00.00.02.1A.00.1D	1	1805766635
196	410	7	00.00.00.00.00.00.01	1	1805767080
3564	411	8	80.2C.10.14.29.60.FB.1	1	1805782793



What skills are required/learned to be successful?

- Good in mathematics and programming
 - Optimization, graph theory (graph signal processing), statistics, Matlab+Python/Julia
- Willingness and courage to explore and learn new (cross-disciplinary) subjects
- Ingenuity
- Be vocal, especially with your adviser

THEN MY GROUP IS FOR YOU!!!

Stop by and talk to me (ED 639)!

c.fung@ieee.org

<https://mcube.lab.nycu.edu.tw/~cfung>

or Google “Carrson Fung”



Self-Supervised Learning (SSL)

