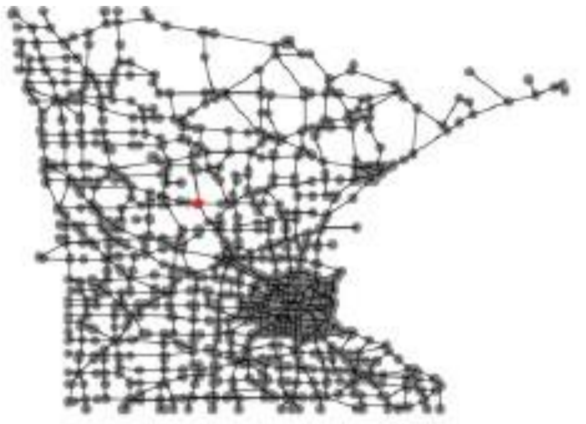
Recent advances in learning with graphs

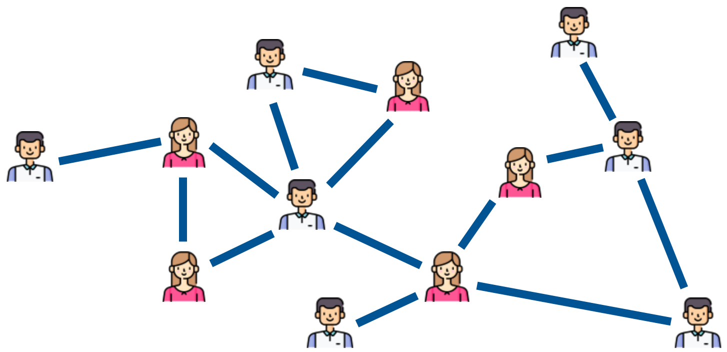
Xiaowen Dong

Department of Engineering Science
University of Oxford

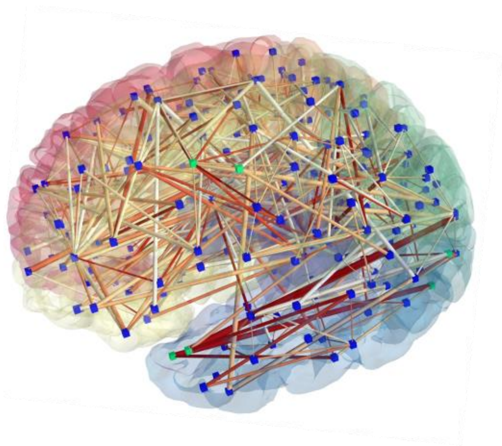
Networks are pervasive



traffic network



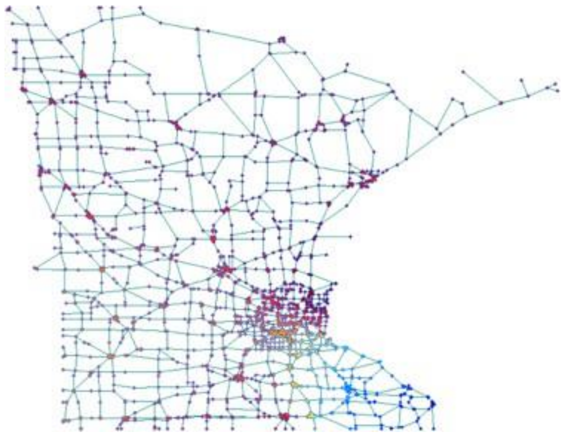
social network



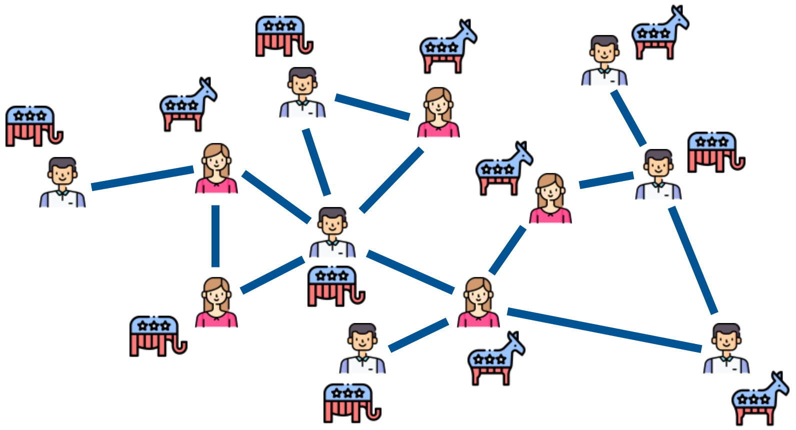
brain network

networks are mathematically represented by graphs

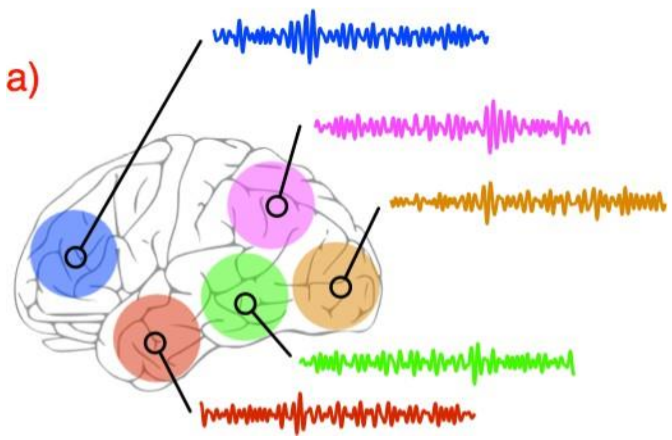
Data collected in networks are pervasive



congestion in road
junctions



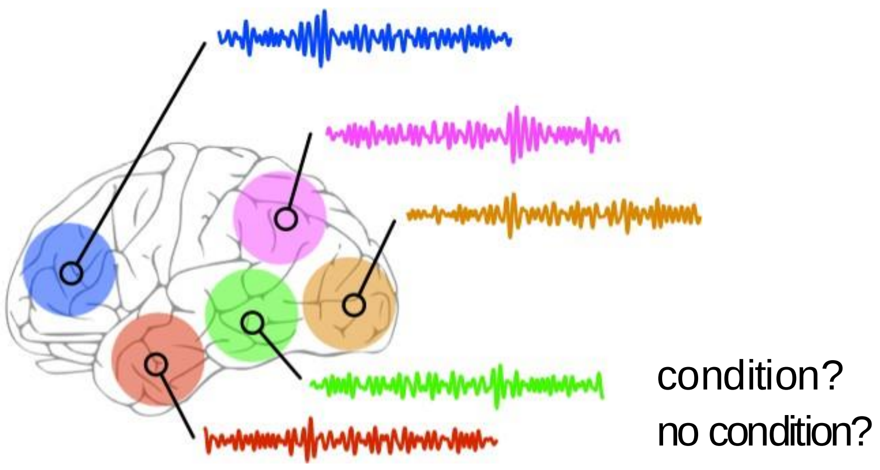
preferences of
individuals



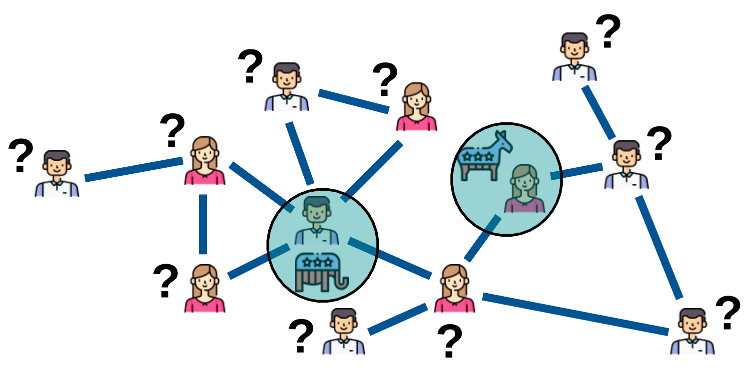
activities in brain
regions

from graphs to graph-structured data

Learning with graph-structured data



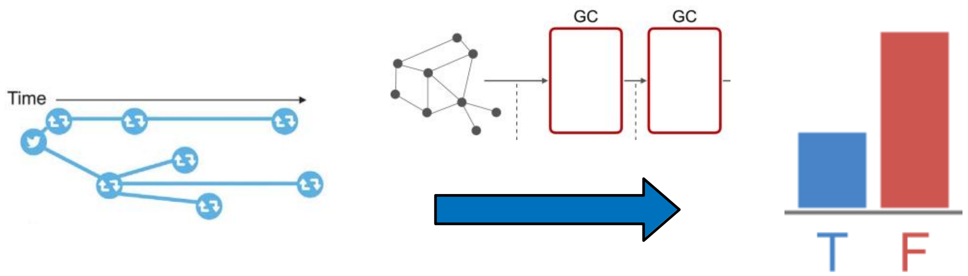
graph-level classification (supervised)



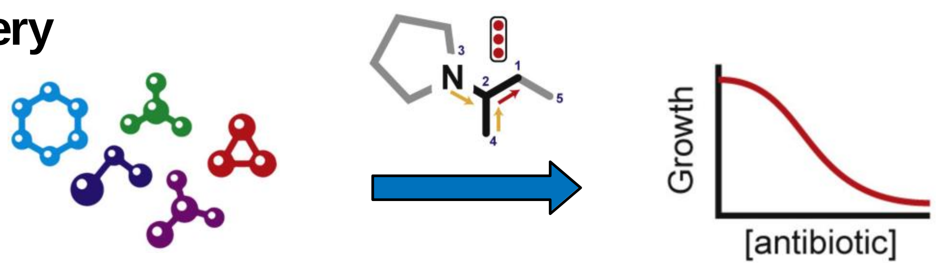
node-level classification (semi-supervised)

Exciting possibilities enabled by graph ML

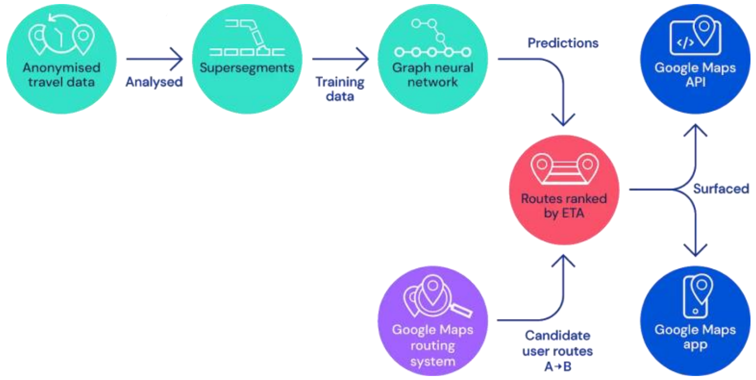
fake news detection

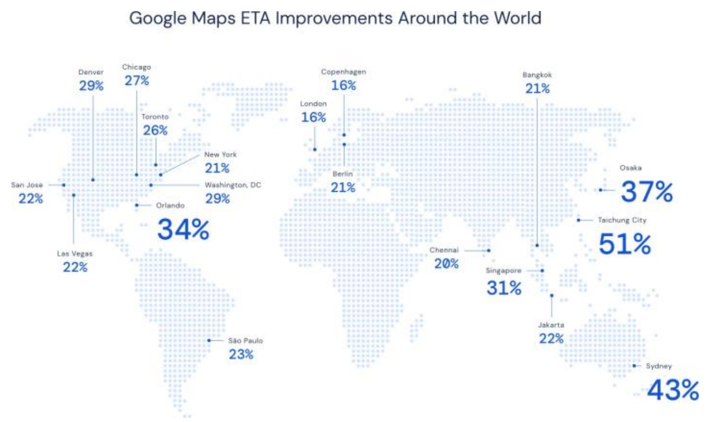


drug discovery



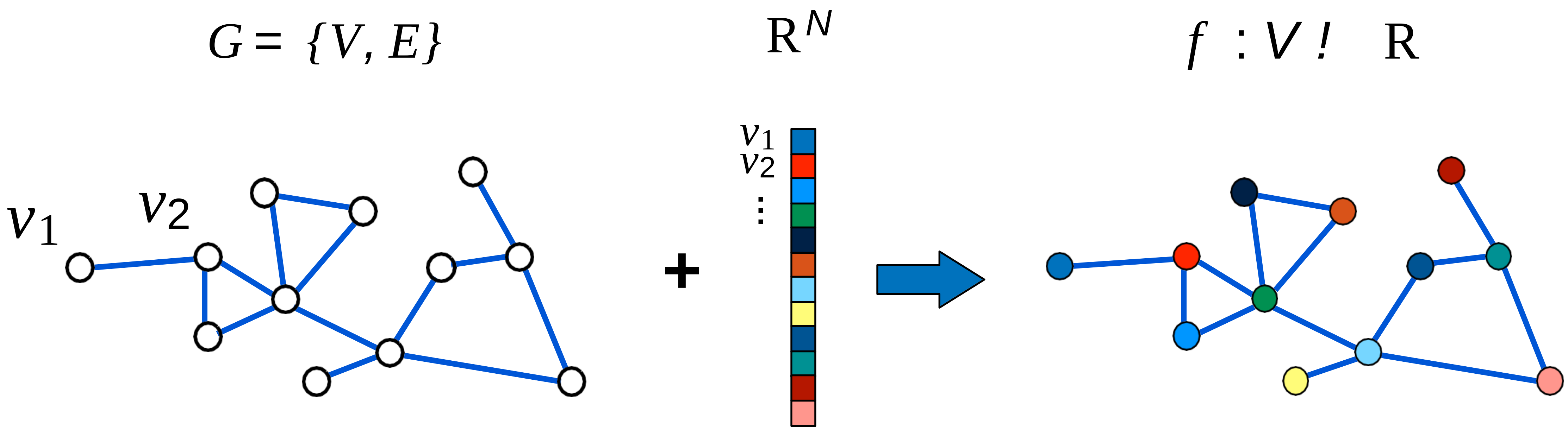
traffic prediction





Graph signal processing

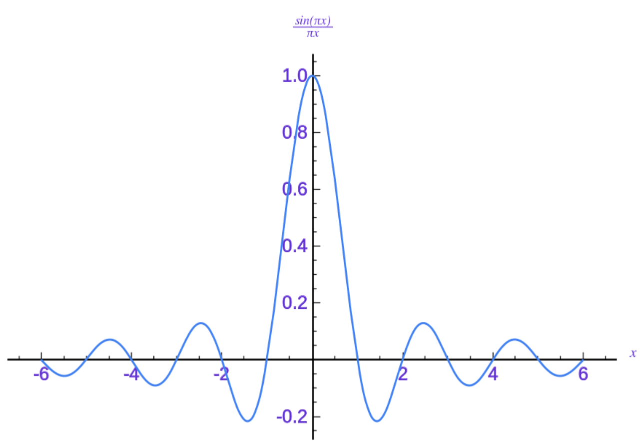
- Graph-structured data can be represented by graph signals



takes into account both structure (edges) and data (values at nodes)

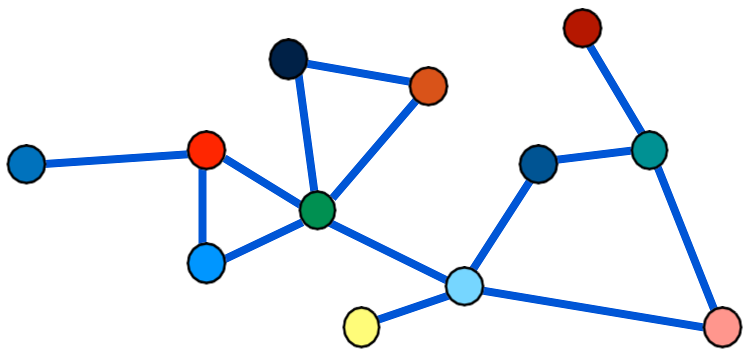
Graph signal processing

1D signal



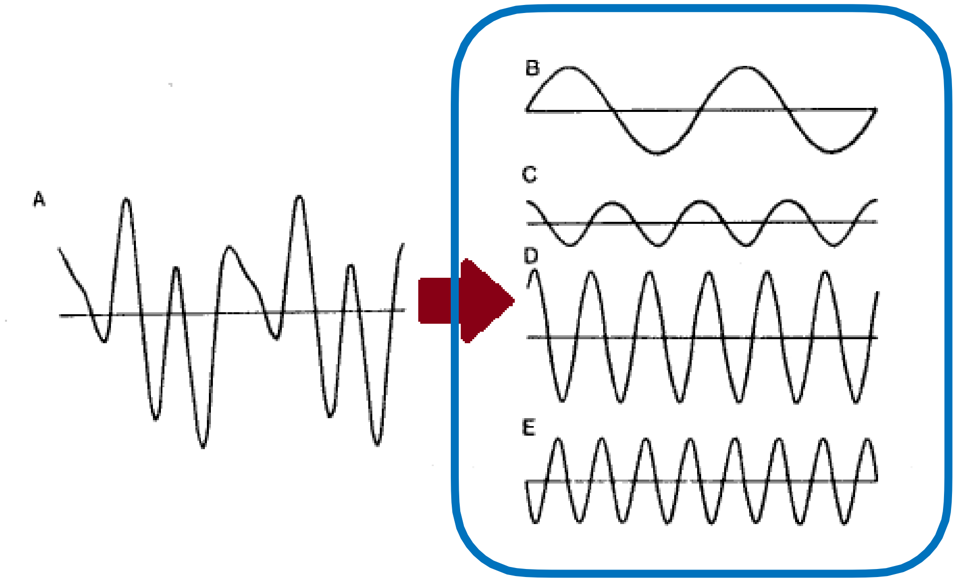
2D signal

$f : V \to \mathbb{R}$



how to generalise classical signal processing tools on irregular domains such as graphs?

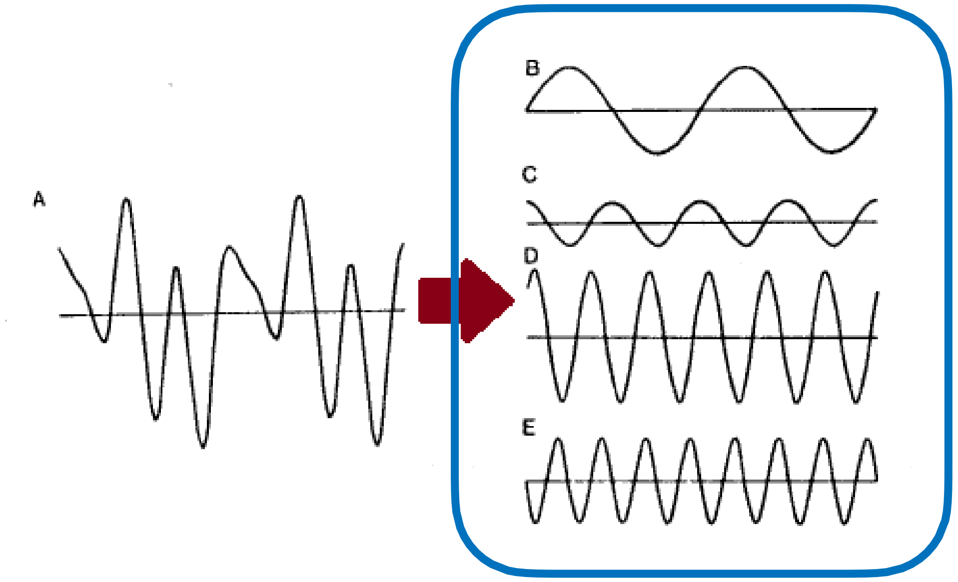
Graph signal processing



classical signal processing

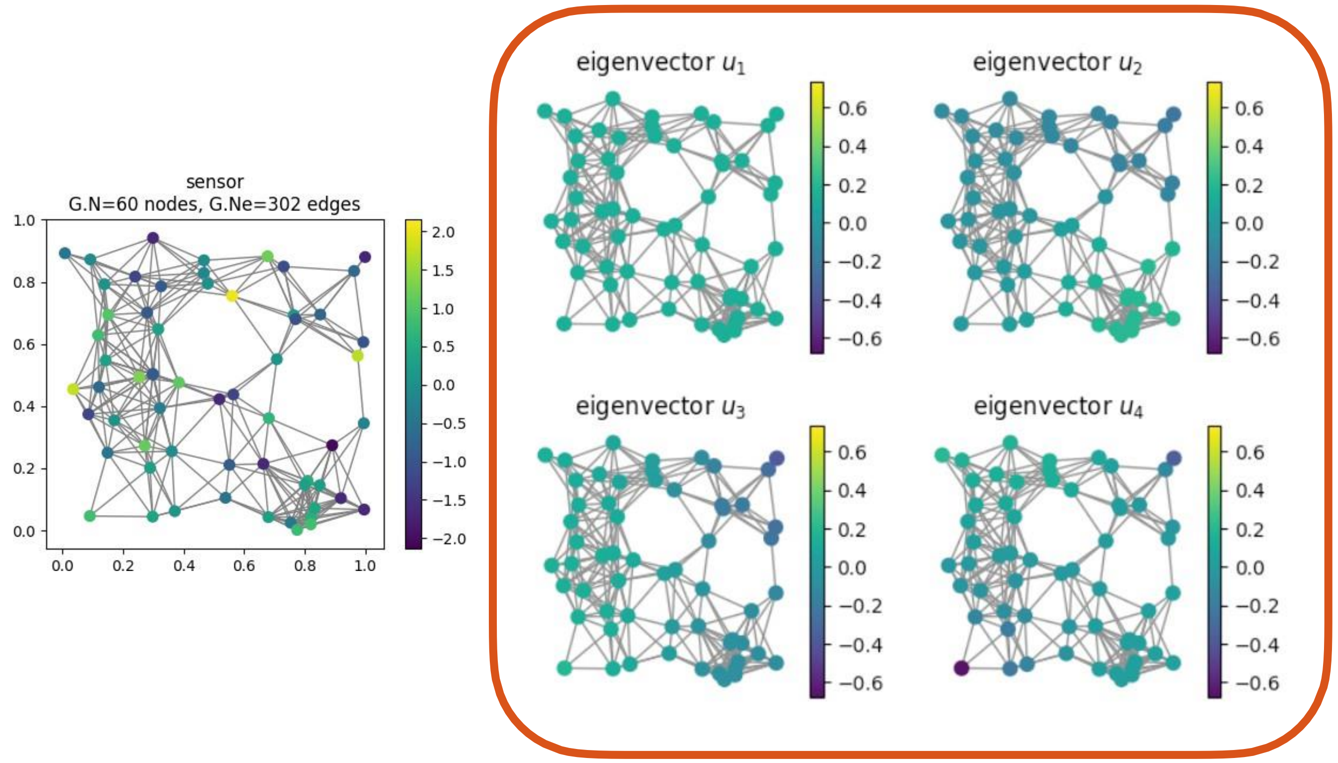
- complex exponentials provide "building blocks" of 1D signal (different oscillations or frequencies)
- leads to **Fourier transform**
- enables frequency filtering (equivalent to convolution)

Graph signal processing



classical signal processing

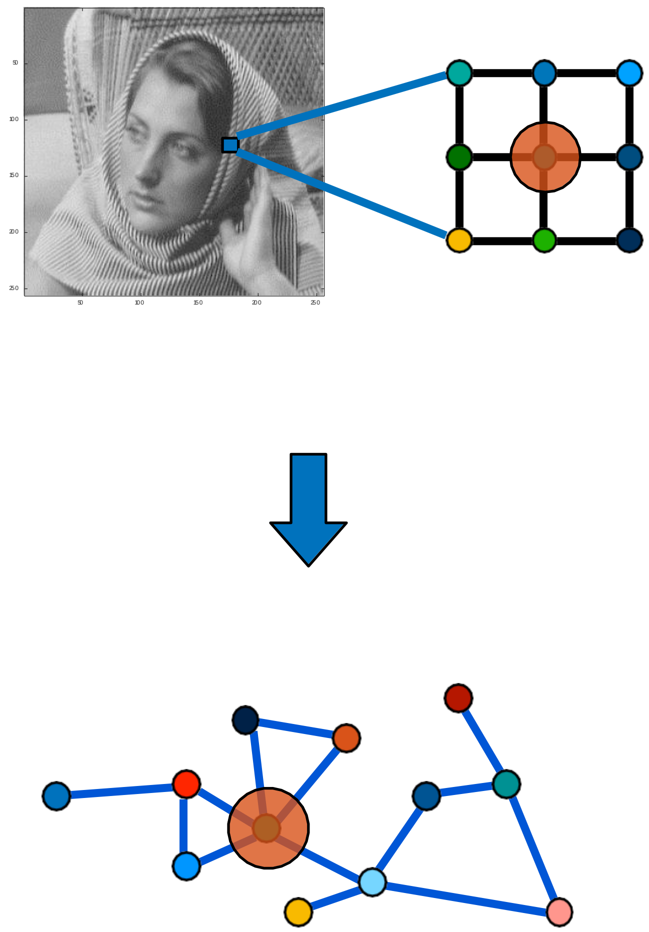
- complex exponentials provide "building blocks" of 1D signal (different oscillations or frequencies)
- leads to **Fourier transform**
- enables frequency filtering (equivalent to convolution)



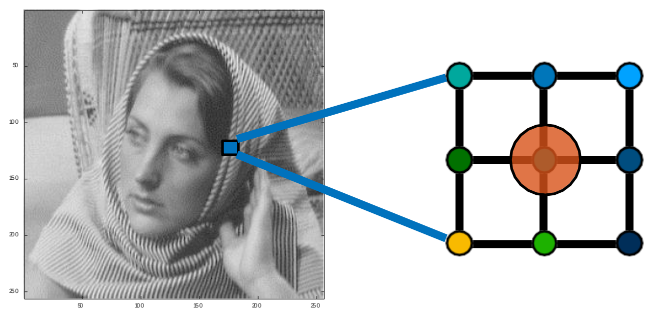
graph signal processing

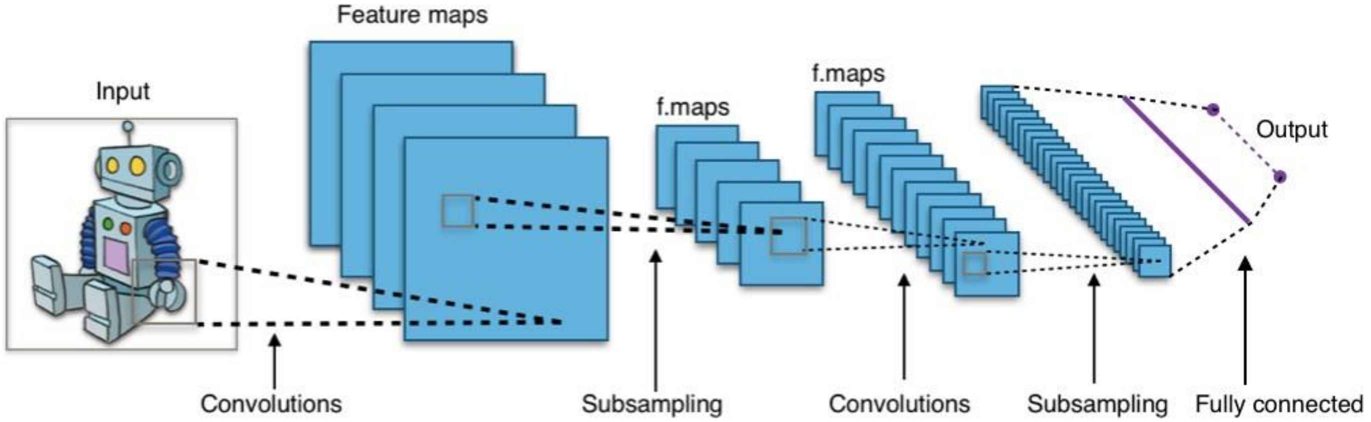
- Laplacian eigenvectors provide "building blocks" of graph signal (different oscillation or frequencies)
- leads to **graph Fourier transform**
- enables **convolution** and **filtering** on graphs

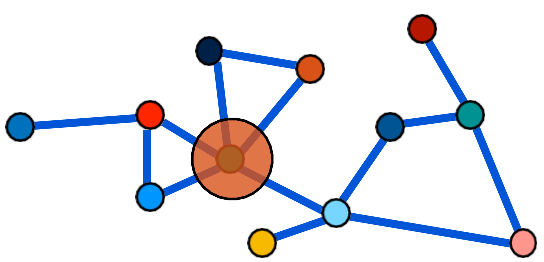
Convolutional neural networks on graphs

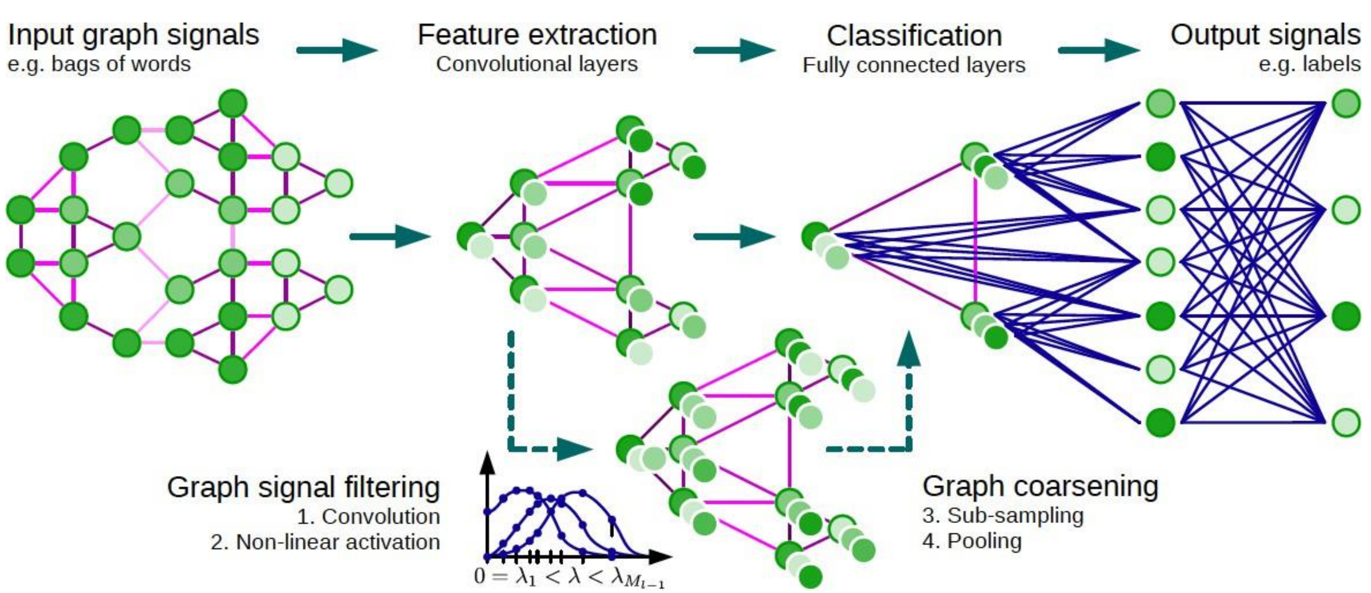


Convolutional neural networks on graphs

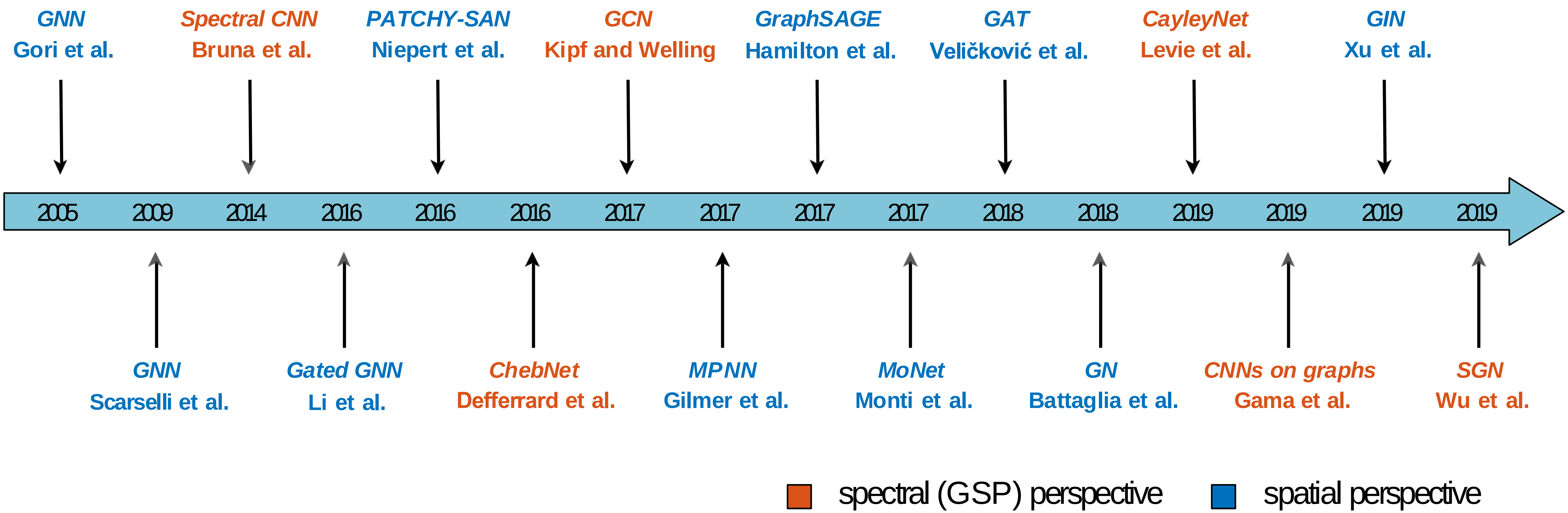




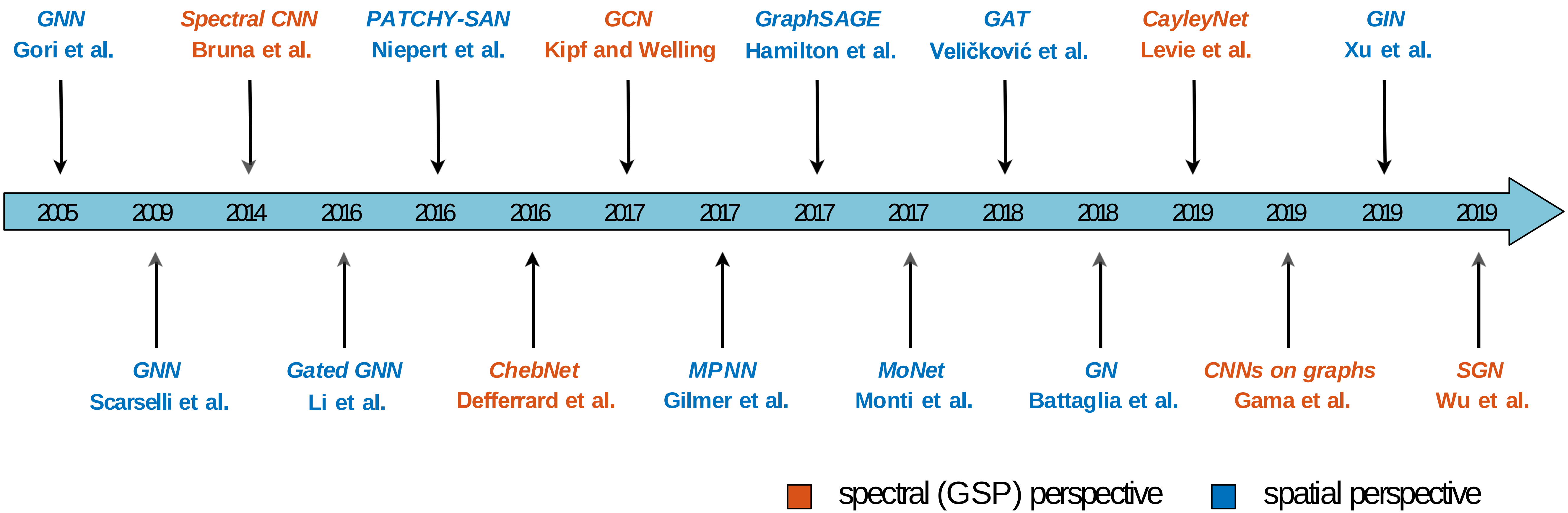




(More generally) Graph neural networks

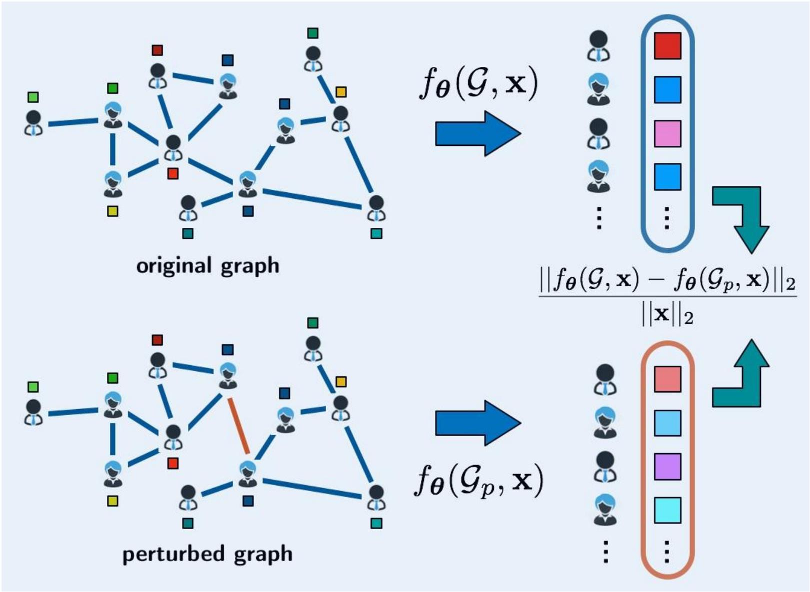


(More generally) Graph neural networks

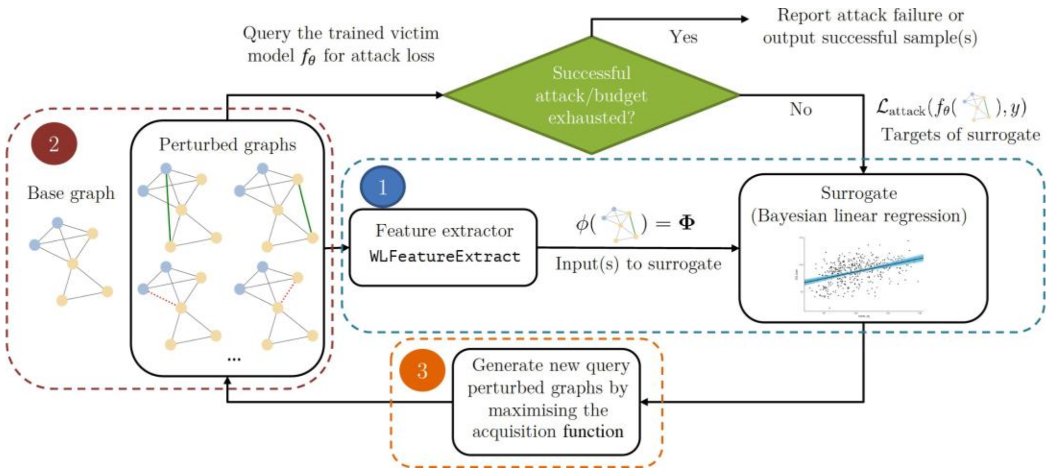


more recently: graph transformers and LLM-powered models

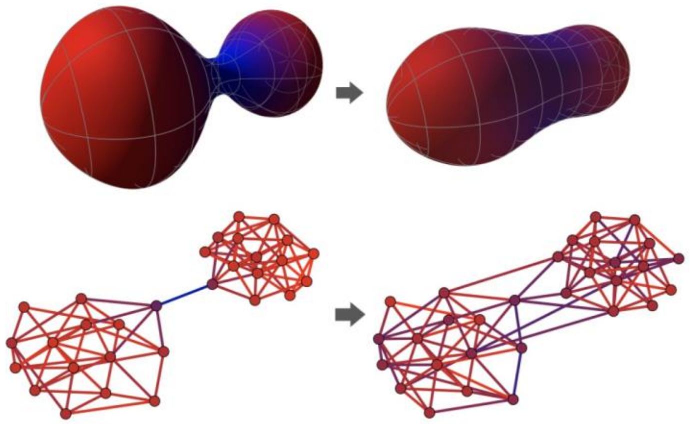
Our research - Theoretical investigation



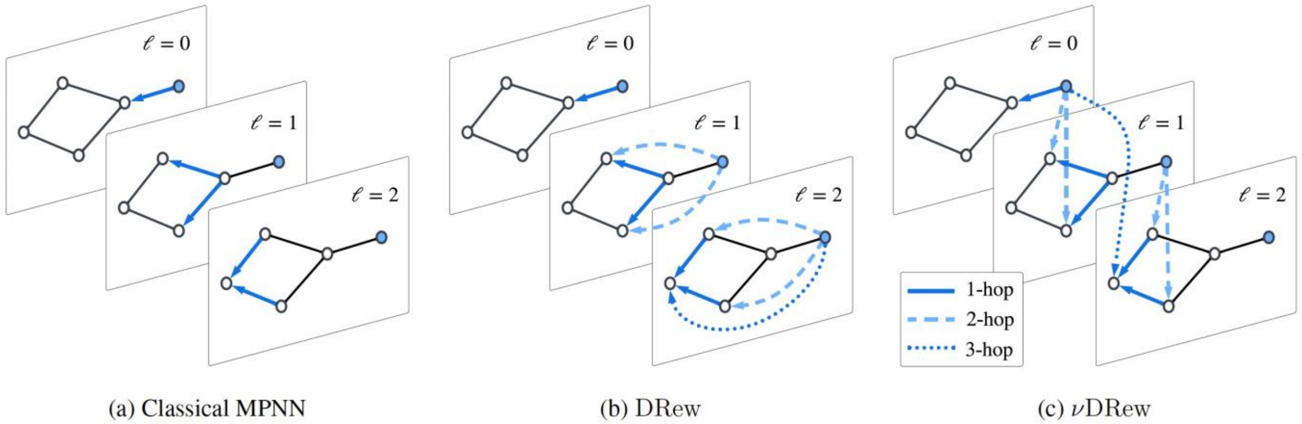
filter stability



adversarial attacks

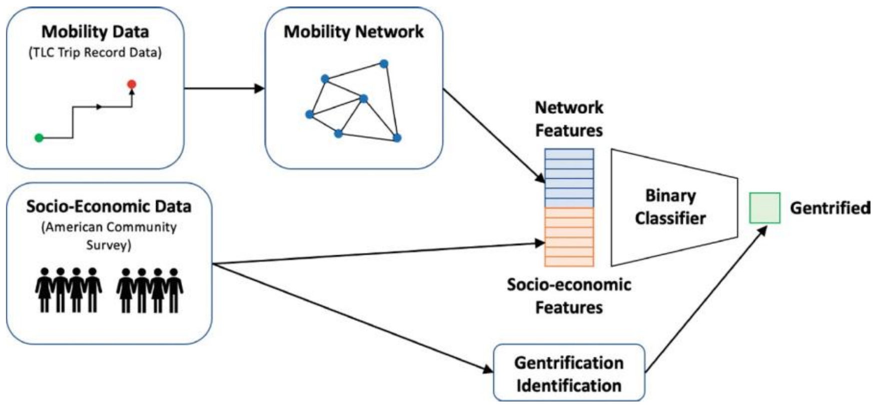


understanding/mitigating over-squashing

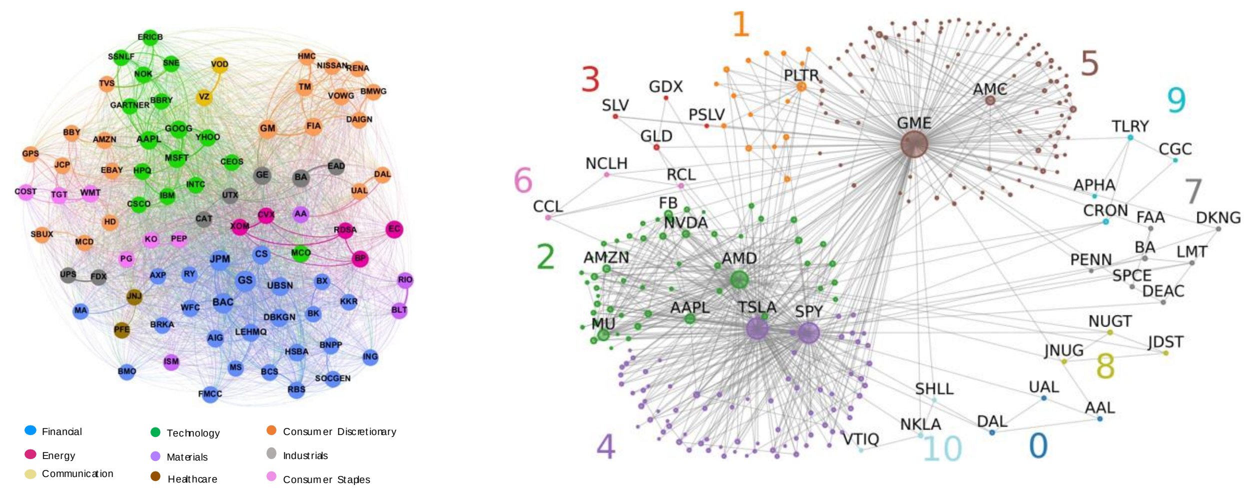


dynamic message passing with delay

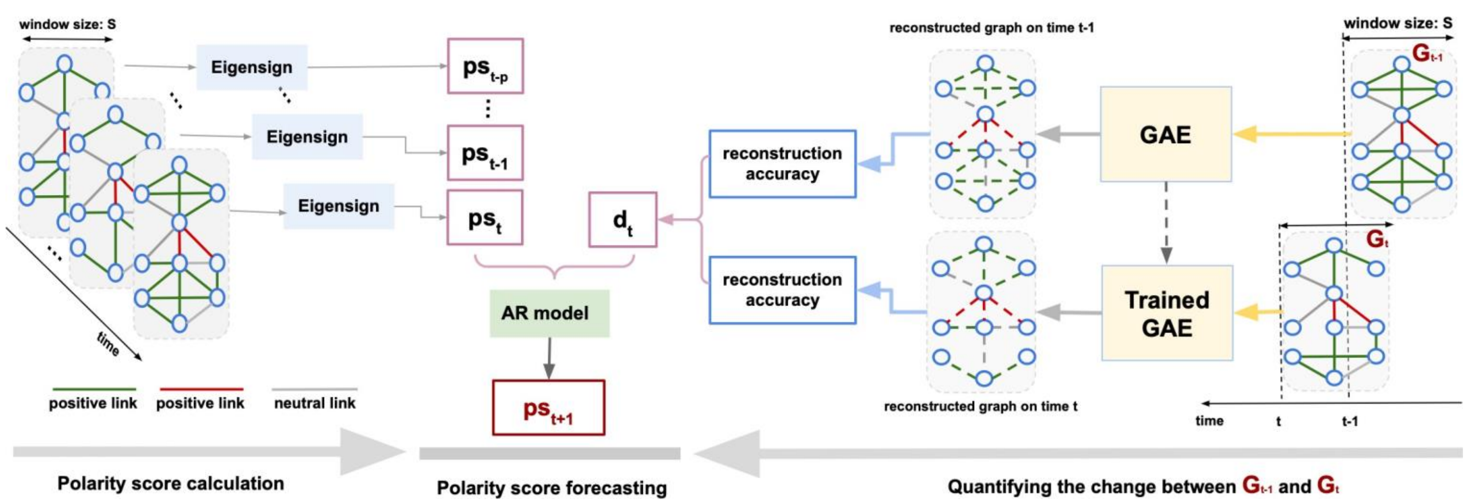
Our research - Applications



urban gentrification



financial market analysis



social network polarisation

Why we need JADE

- Graph-structured datasets are typically very large
  - millions of nodes and billions of edges
  - rich features associated with nodes/edges
- Graph ML models can be computationally costly to train
  - on large graphs (even with linear complexity)
  - for graph transformers (in theory quadratic complexity)
- Graph ML tasks take diverse forms
  - graph-level tasks
  - node/edge-level tasks
  - dynamic/online settings

Thank you!

contact: xdong@robots.ox.ac.uk