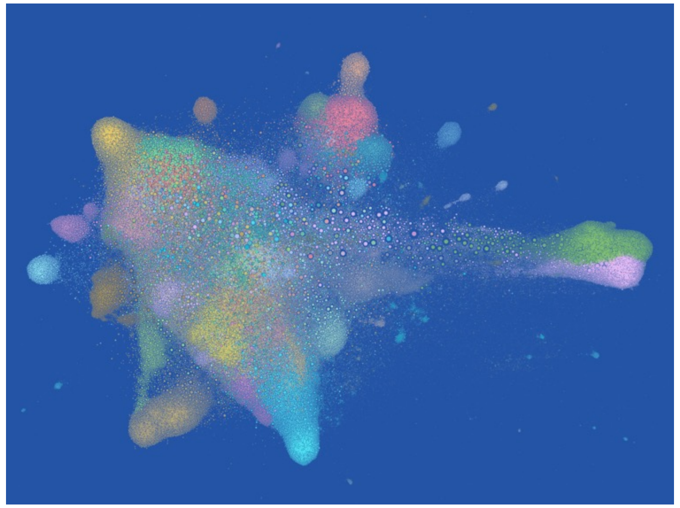
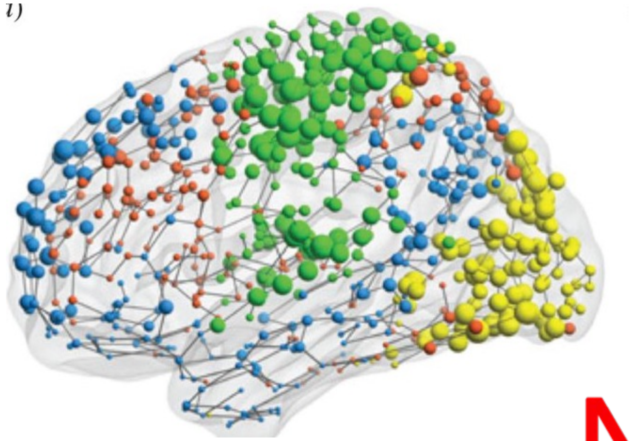


Learning on Graphs

Mark Coates
Department of Electrical and Computer Engineering
McGill University

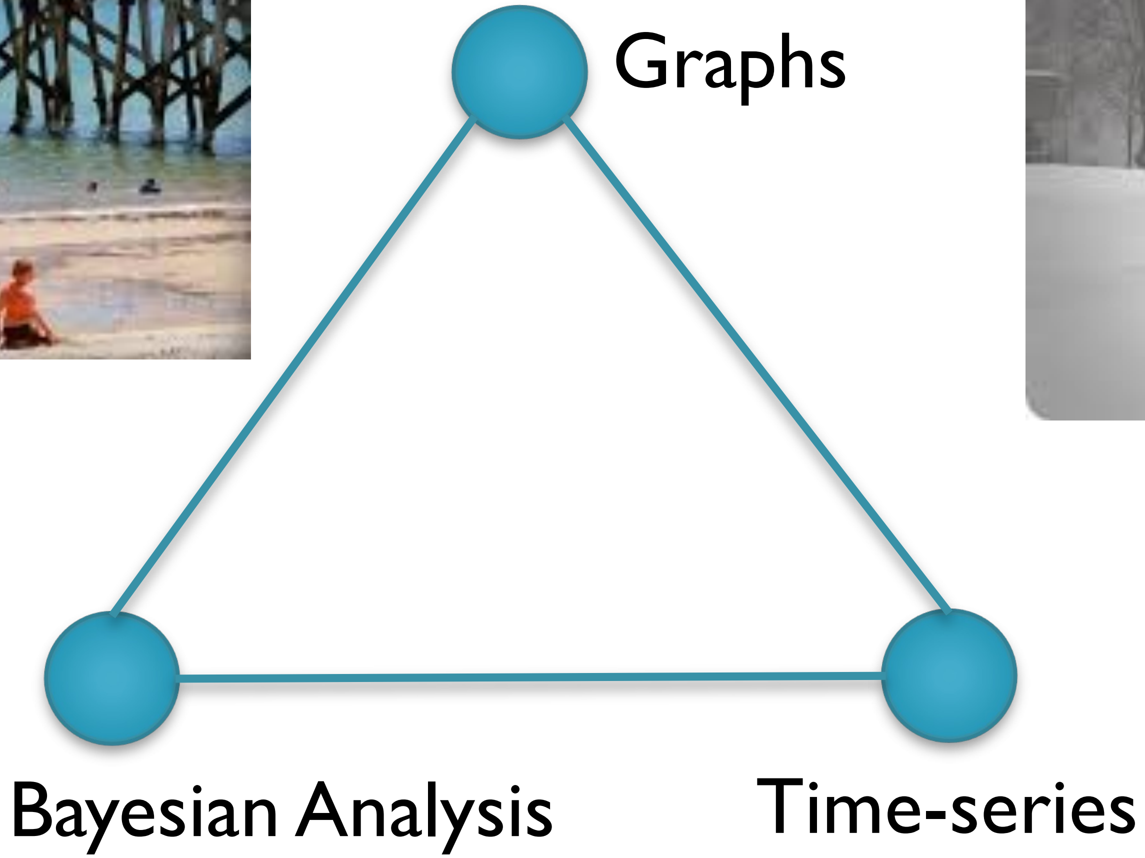


Figures: Vertés et al. 2014; Axel Bruns / QUT Digital Media Research Centre

McGill Networks Lab

McGill Networks Research Laboratory

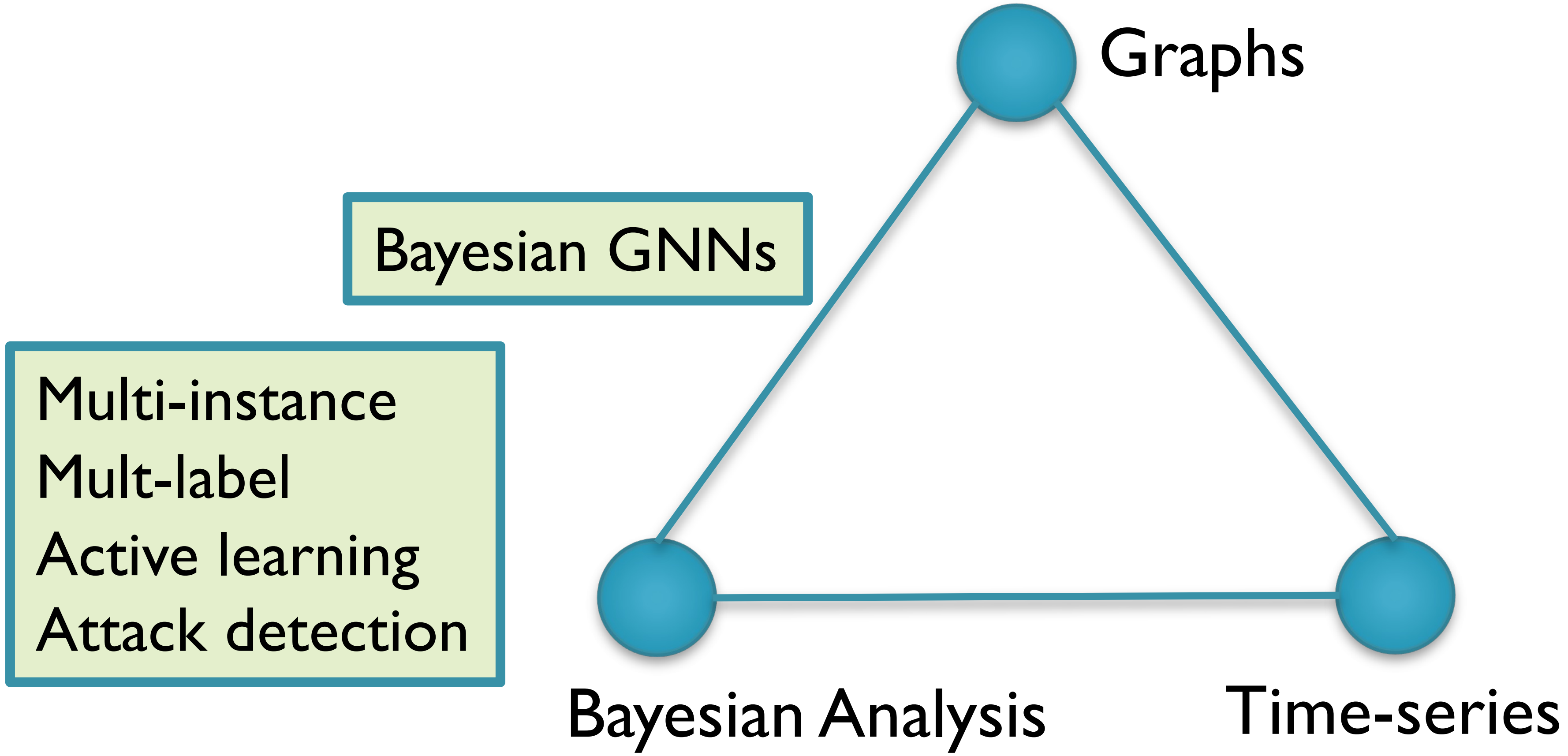
McGill



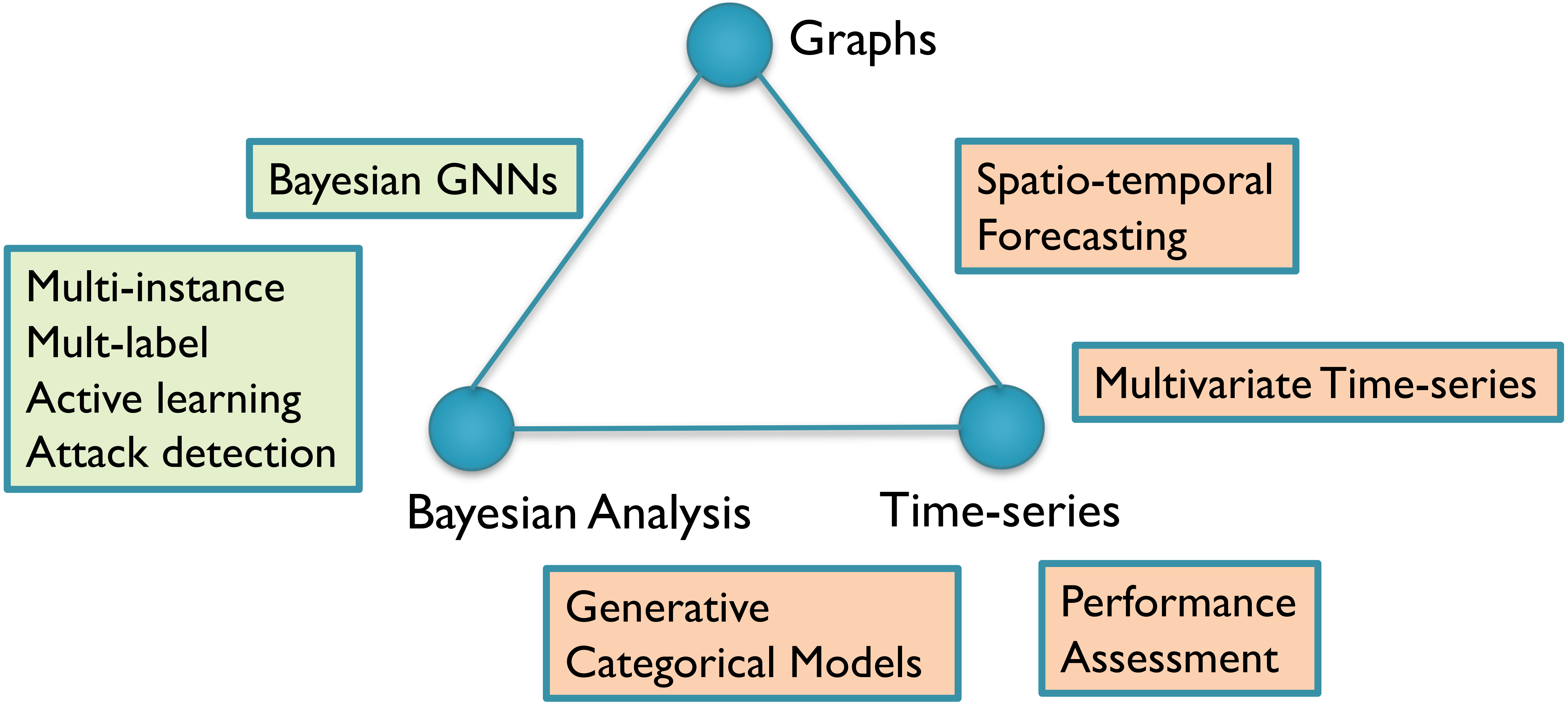
McGill Networks Lab

McGill Networks Research Laboratory

McGill



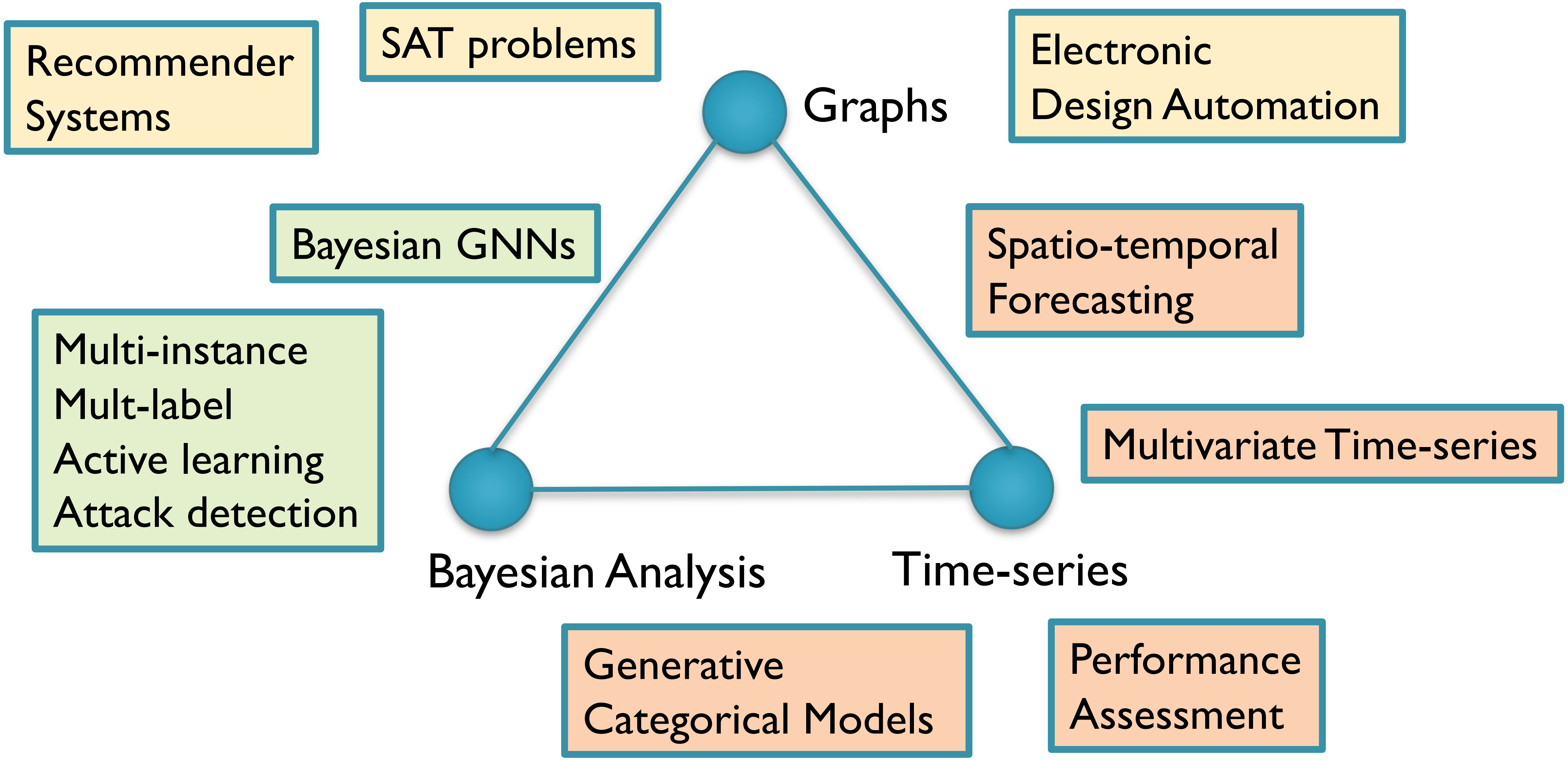
McGill Networks Lab



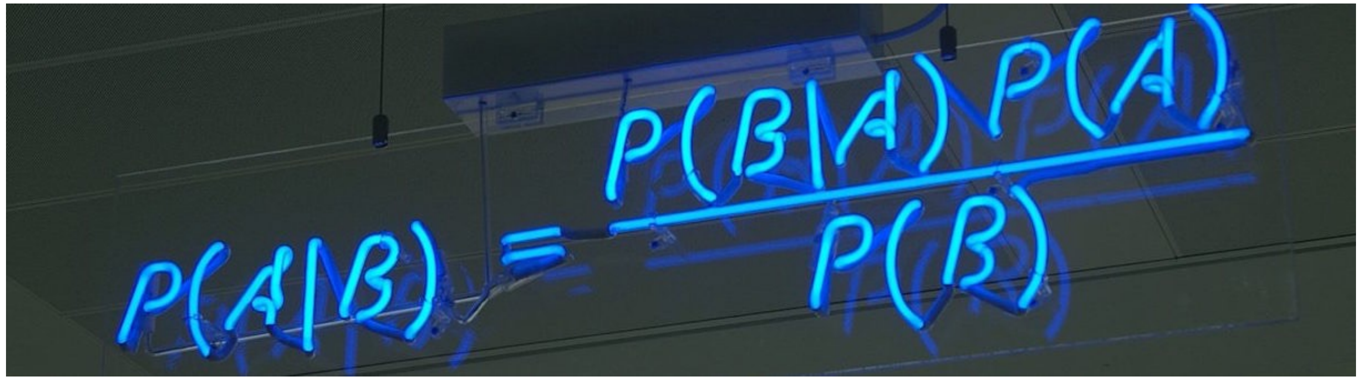
McGill Networks Lab

McGill Networks Research Laboratory

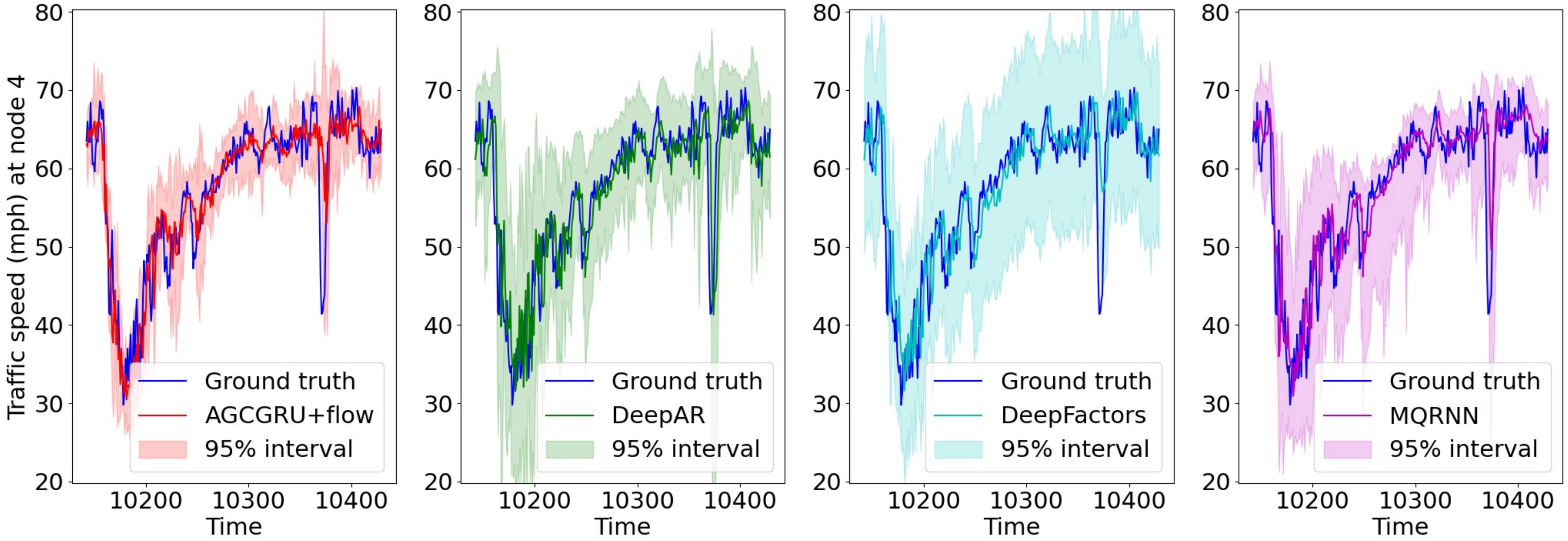
McGill



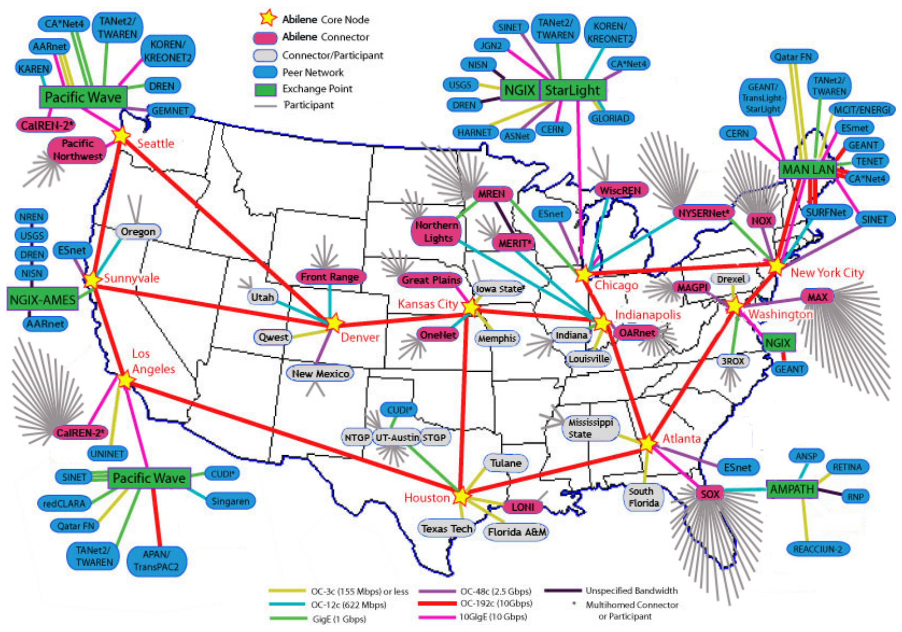
Why Bayes?



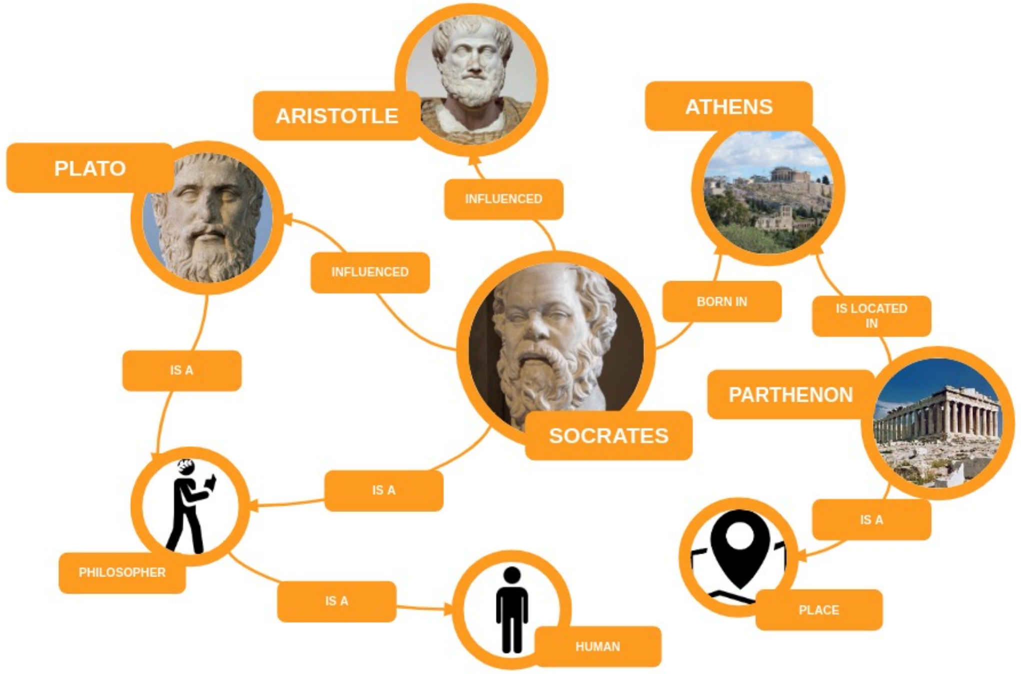
- Incorporate prior knowledge (learn from less data + active learning)
- Handle missing data (or noisy data)
- Confidence intervals → prediction intervals



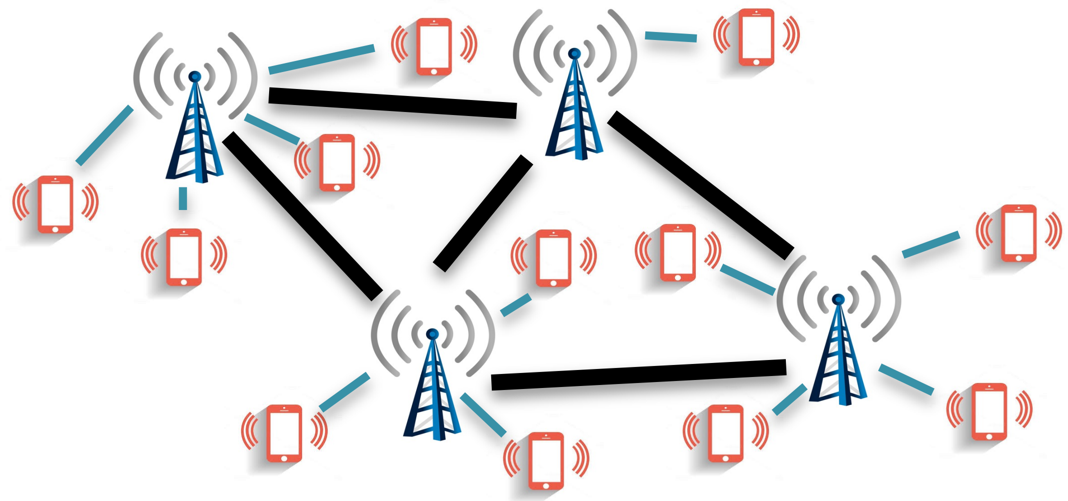
Why Graphs?

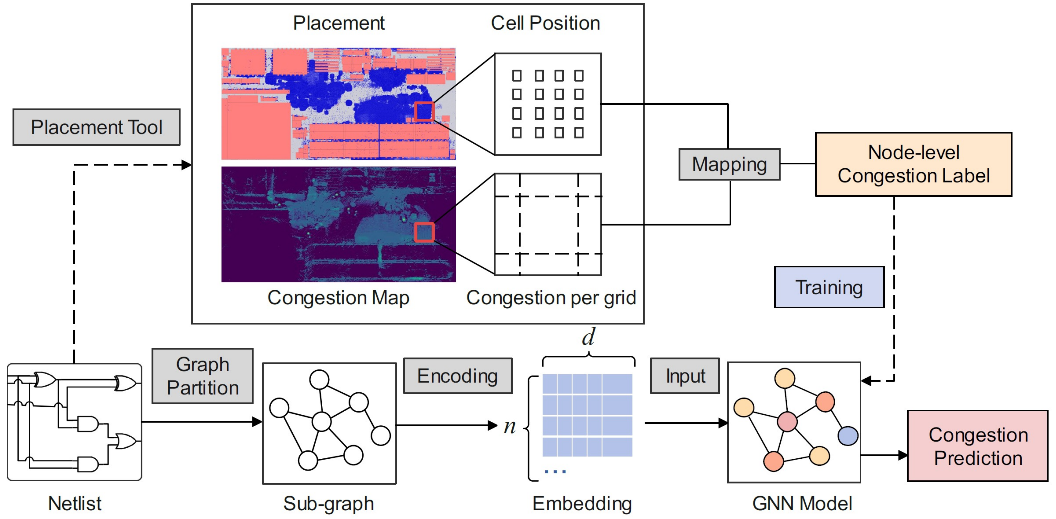


Seattle
Sunnyvale
Los Angeles
Denver
Kansas City
Houston
Chicago
Indianapolis
Atlanta
New York City
Washington
Pacific Wave
NGIX StarLight
MAN LAN



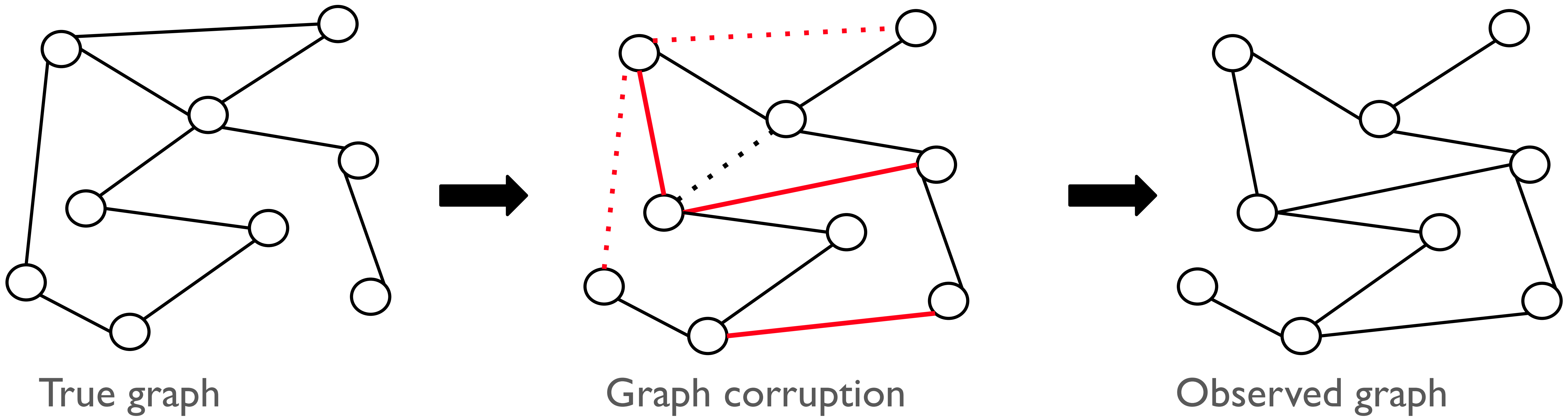
ARISTOTLE
PLATO
ATHENS
INFLUENCED
INFLUENCED
BORN IN
IS LOCATED IN
IS A
SOCRATES
PARTHENON
IS A
IS A
PHILOSOPHER
IS A
HUMAN
PLACE





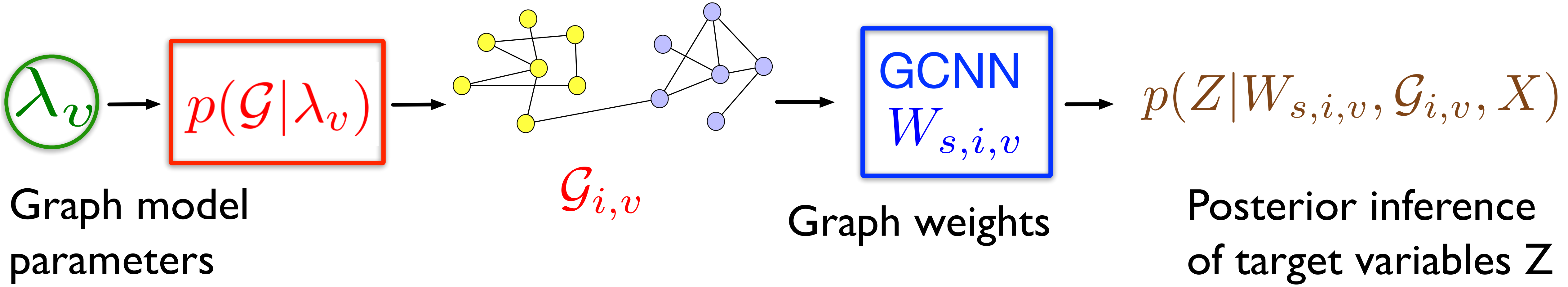
Placement
Cell Position
Placement Tool
Mapping
Node-level Congestion Label
Congestion Map
Congestion per grid
Training
Graph Partition
Encoding
Input
Netlist
Sub-graph
Embedding
GNN Model
Congestion Prediction

Bayesian Graph Neural Networks



- Employ Bayesian framework to account for uncertainty in the graph

Bayesian Graph Neural Networks



$$p(\mathbf{Z}|\mathbf{Y}_{\mathcal{L}},\mathbf{X},\mathcal{G}_{obs}) = \int p(\mathbf{Z}|W,\mathcal{G},\mathbf{X})p(W|\mathbf{Y}_{\mathcal{L}},\mathbf{X},\mathcal{G})p(\mathcal{G}|\lambda)p(\lambda|\mathcal{G}_{obs})\,dW\,d\mathcal{G}\,d\lambda\,,$$

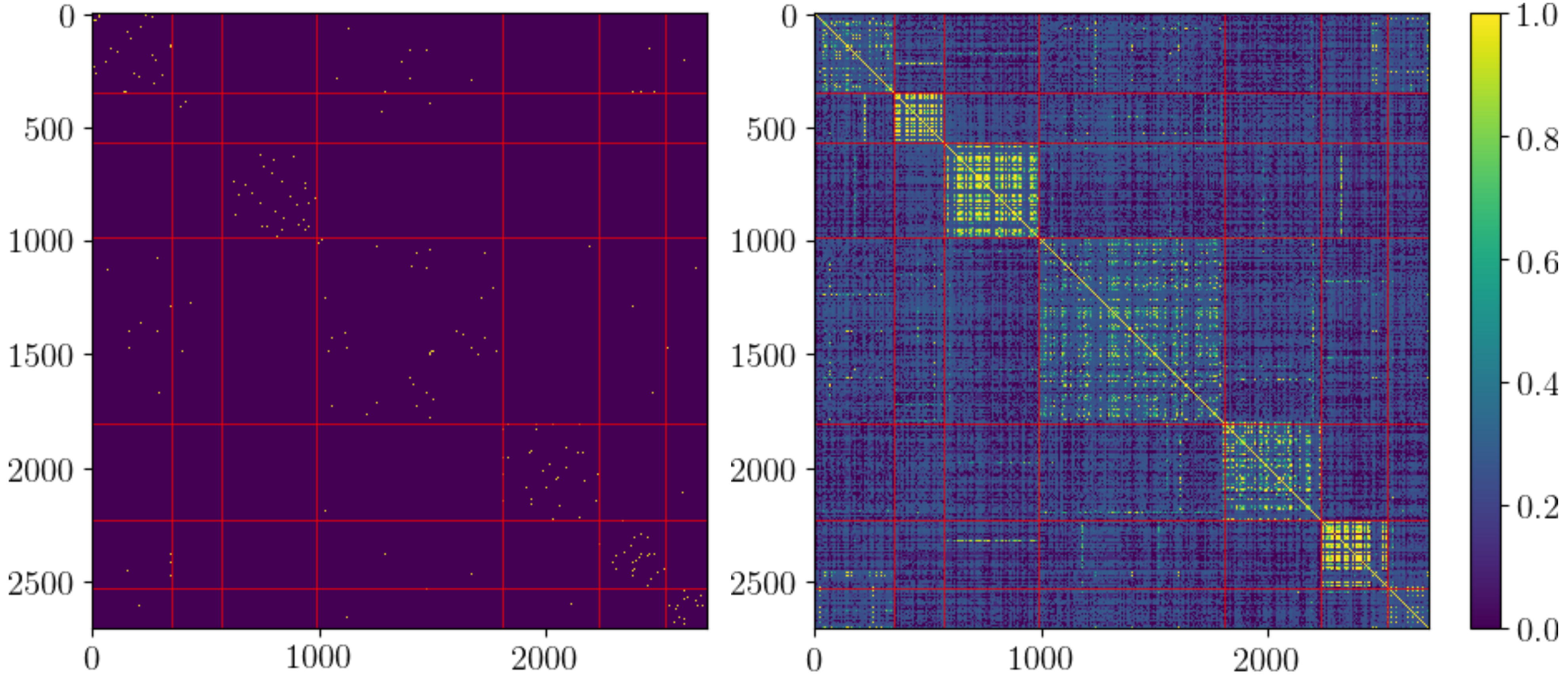
$$\approx \frac{1}{V}\sum_{v=1}^{V}\frac{1}{N_G S}\sum_{i=1}^{N_G}\sum_{s=1}^{S} p(\mathbf{Z}|W_{s,i,v},\mathcal{G}_{i,v},\mathbf{X})$$

$p(G|\lambda)$: Assortative mixed membership stochastic block model

Zhang, Pal, Coates, Üstebay, AAAI 2019

Bayesian Graph Neural Networks

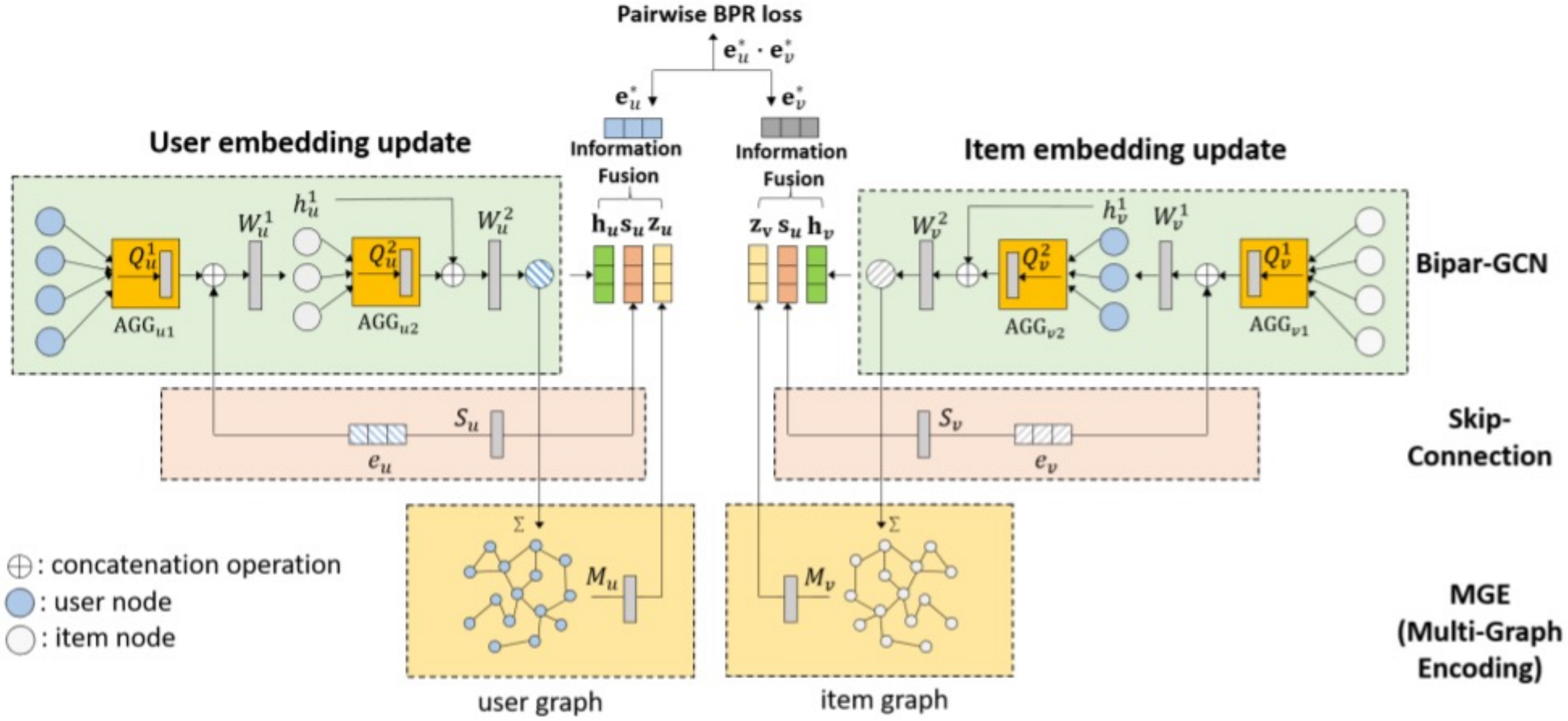
Maximum a posteriori estimate of adjacency matrix using non-parametric graph model



$A_{\mathcal{G}_{obs}}$ (Cora) $A_{\hat{\mathcal{G}}}$ **(MAP estimate)**

Pal et al., UAI 2020

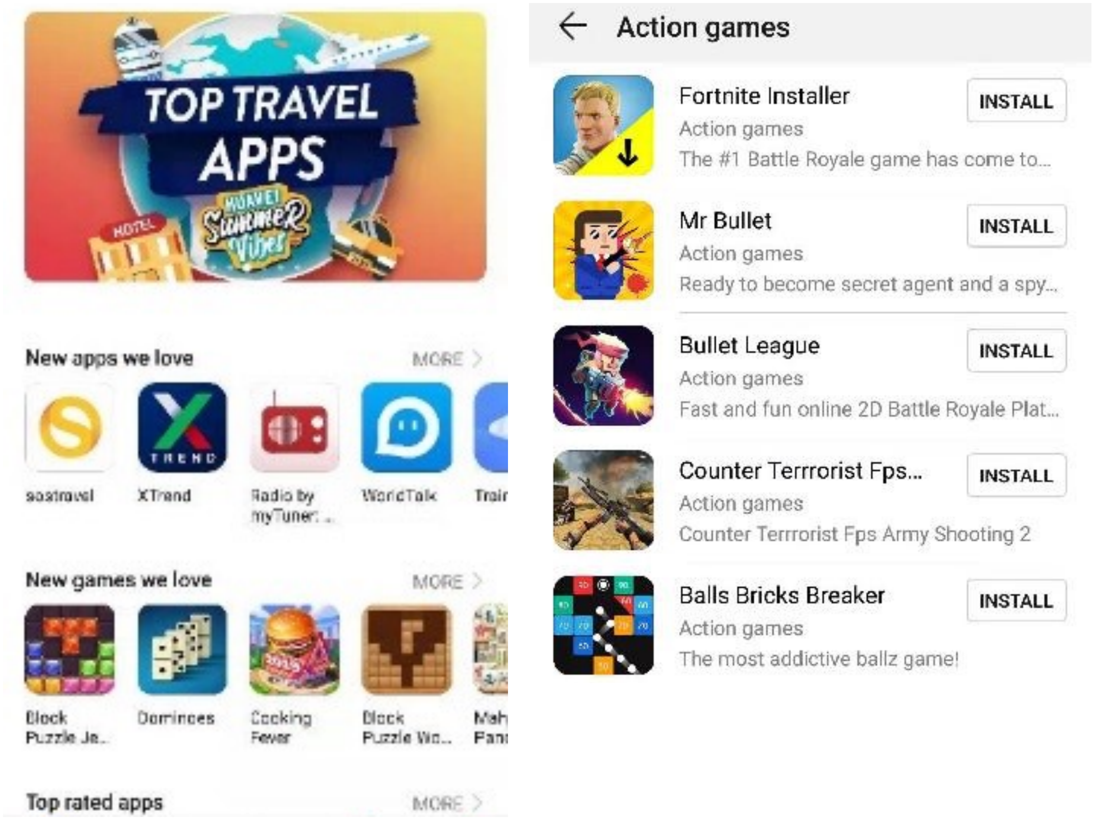
Graph-based Recommender Systems

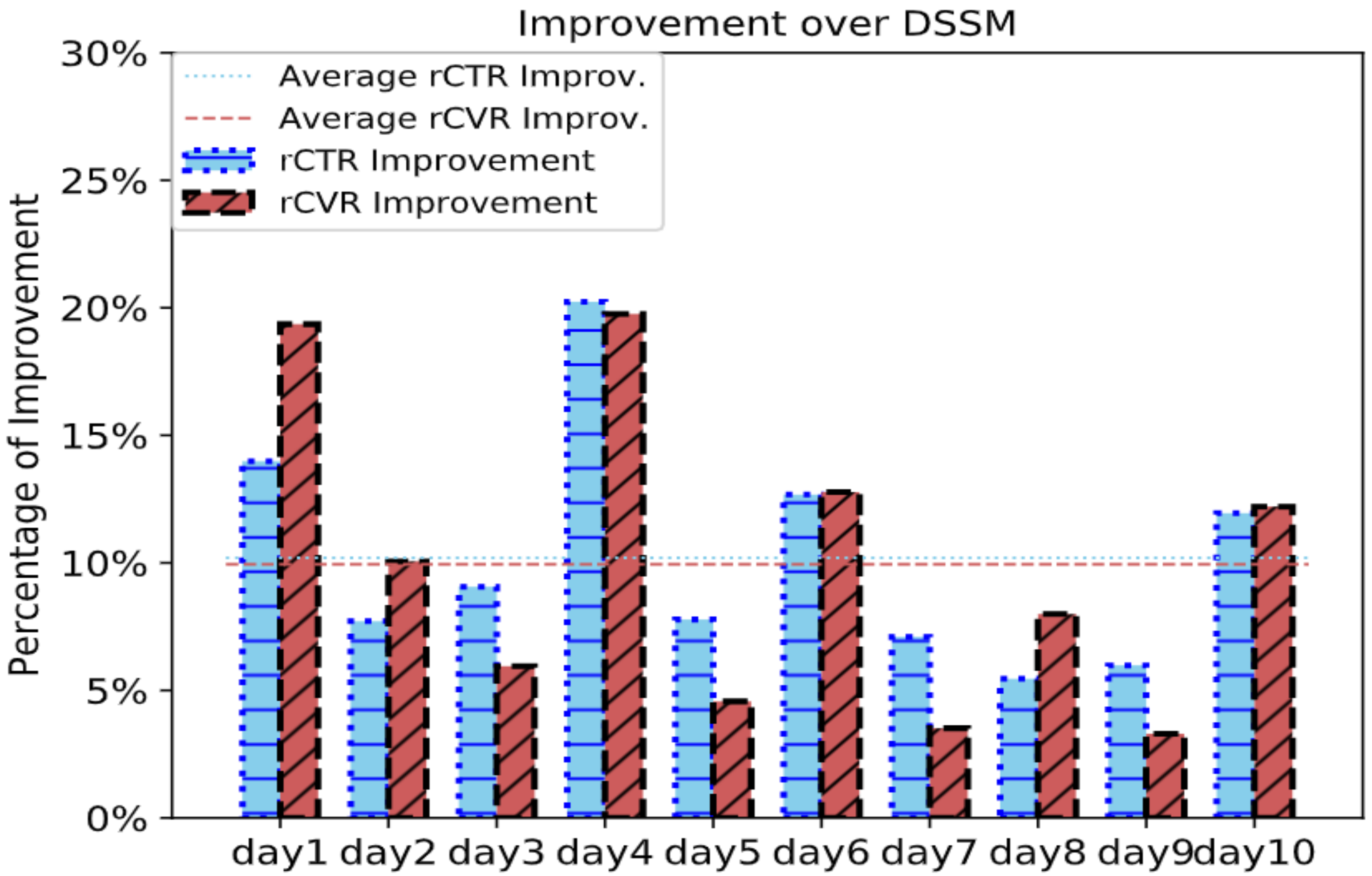


Sun et al. SIGIR 2020, KDD 2021

Graph-based Recommender Systems

>500 million users, >200,000 apps



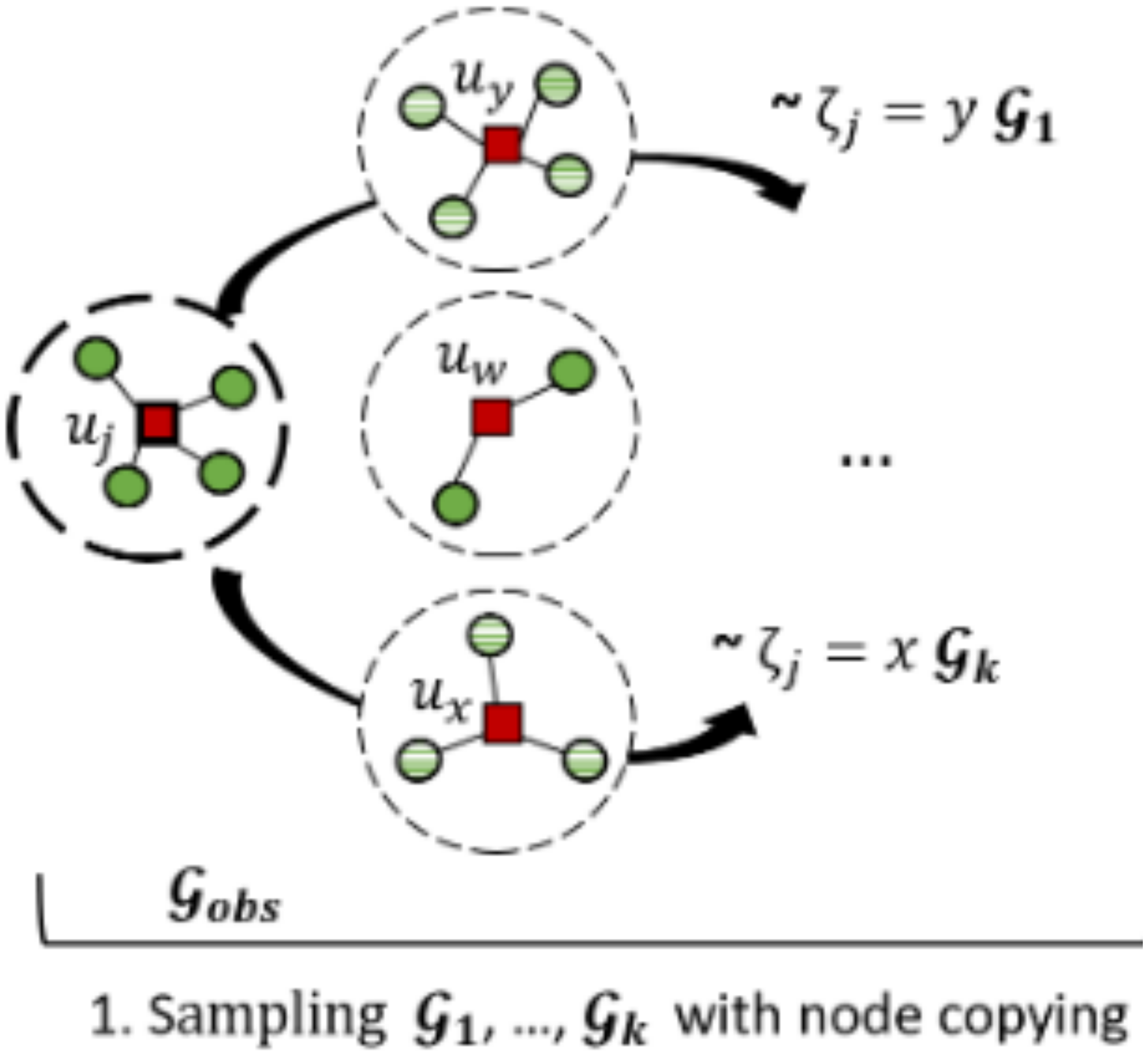


Online A/B Test in Ten Days in App store Rec.

On average 10+% improvement over CTR/CVR

Sun et al. SIGIR 2020, KDD 2020

Bayesian Approach



- Graph is extremely sparse
- Diversity is a problem

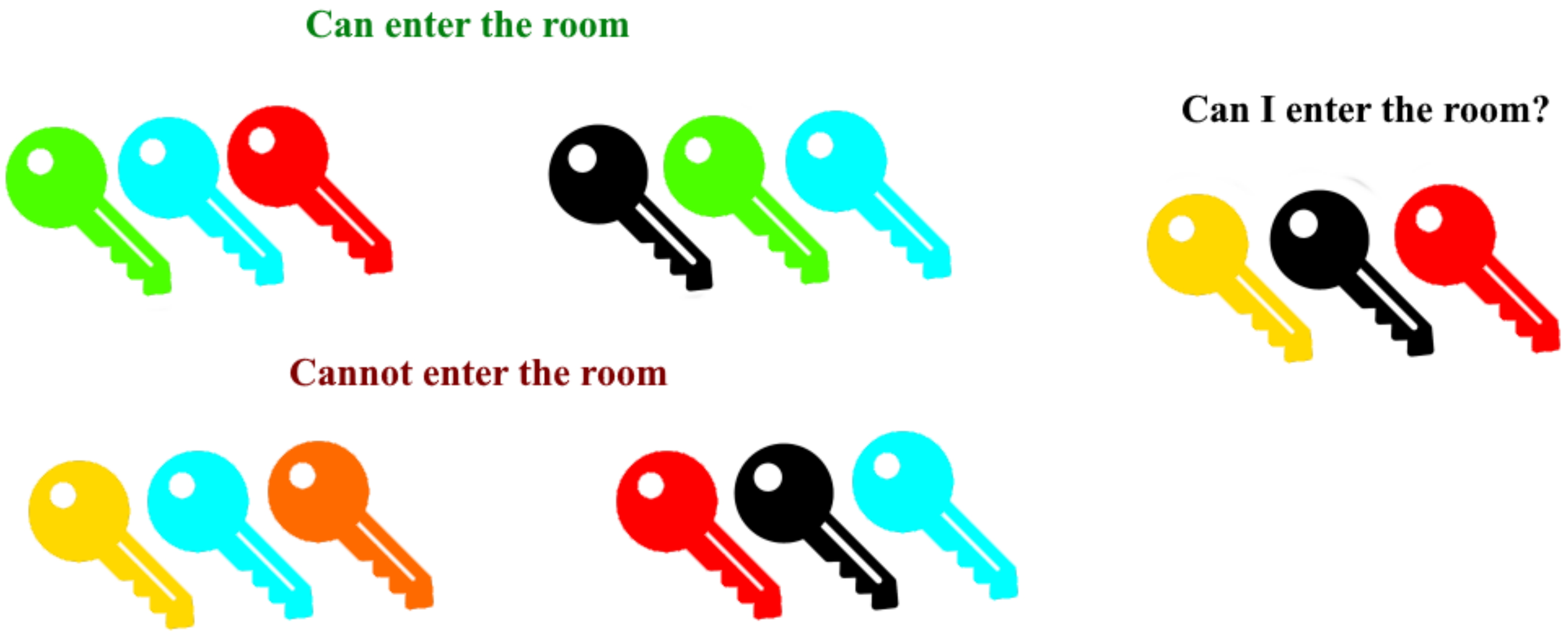
- Use Bayesian GCNs to promote information diffusion and diversity

Sun et al. SIGIR 2020, KDD 2020

Multiple Instance Learning

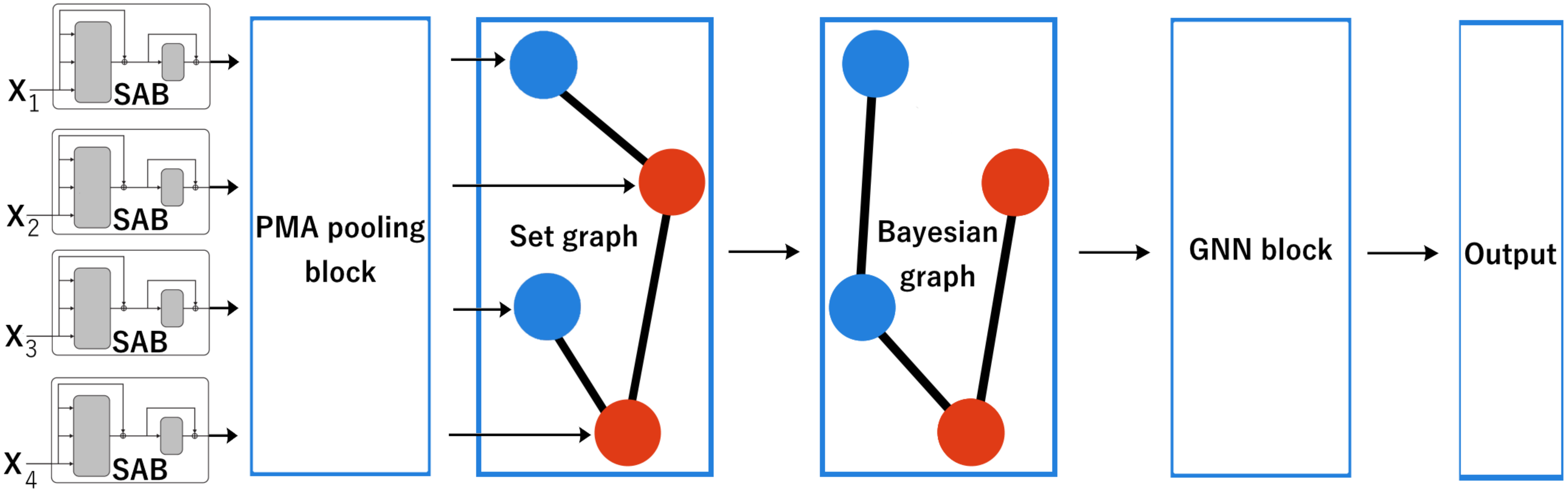
Assign labels to bags (sets of instances), rather than individual instances

- Instance-based approaches:
  - Label instances then pool
- Bag-space approaches
  - Learn mapping from bag descriptor to label



Pal, Valkanas, Regol, Coates AAAI 2022

Proposed Architecture



Goal: approximate posterior of unknown labels conditioned on training labels and bag features.

Key innovation: Use a graph to model relationships between the bags

MIL Experiments: Election result prediction

Given demographic data from US census per county and some voting results, can we predict how the rest of the country will vote?

Dataset source:

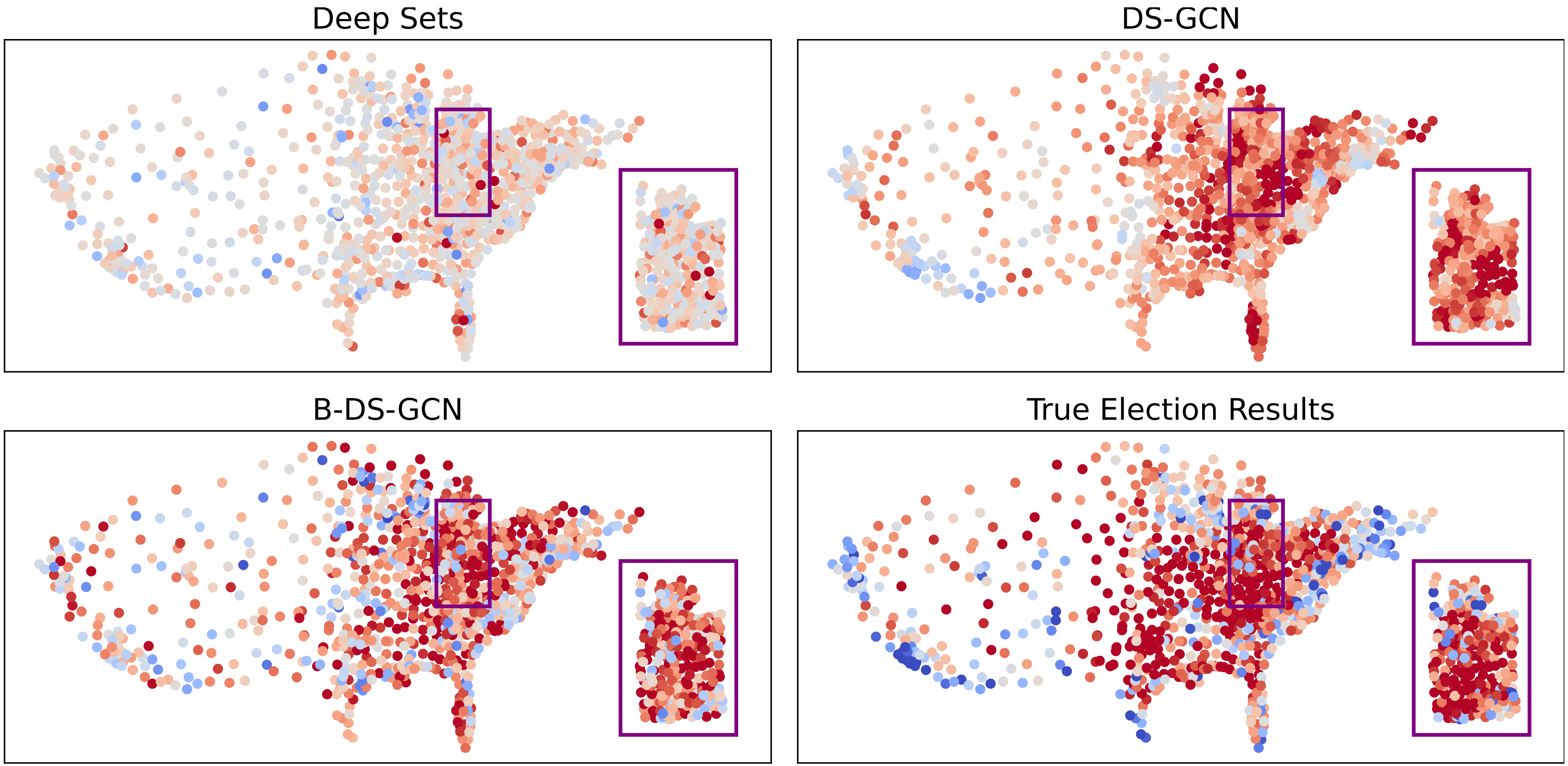
Flaxman et al. "Understanding the 2016 US Presidential Election using ecological inference and distribution regression with census microdata", arXiv 2016

Instances: sampled voters for each county

Features: census data

Bags: counties

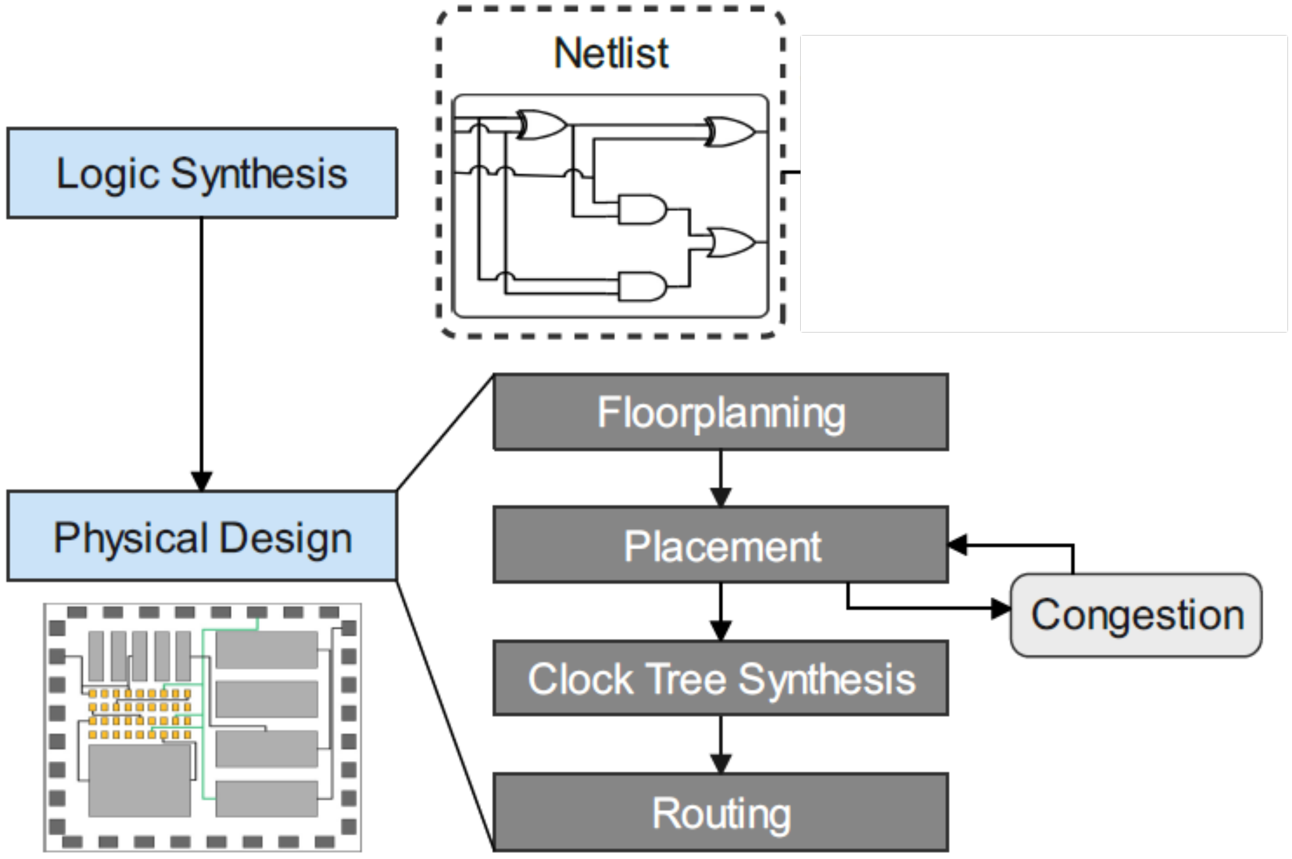
MIL Experiments: Election result prediction



Electronic Design Automation (EDA) Workflow

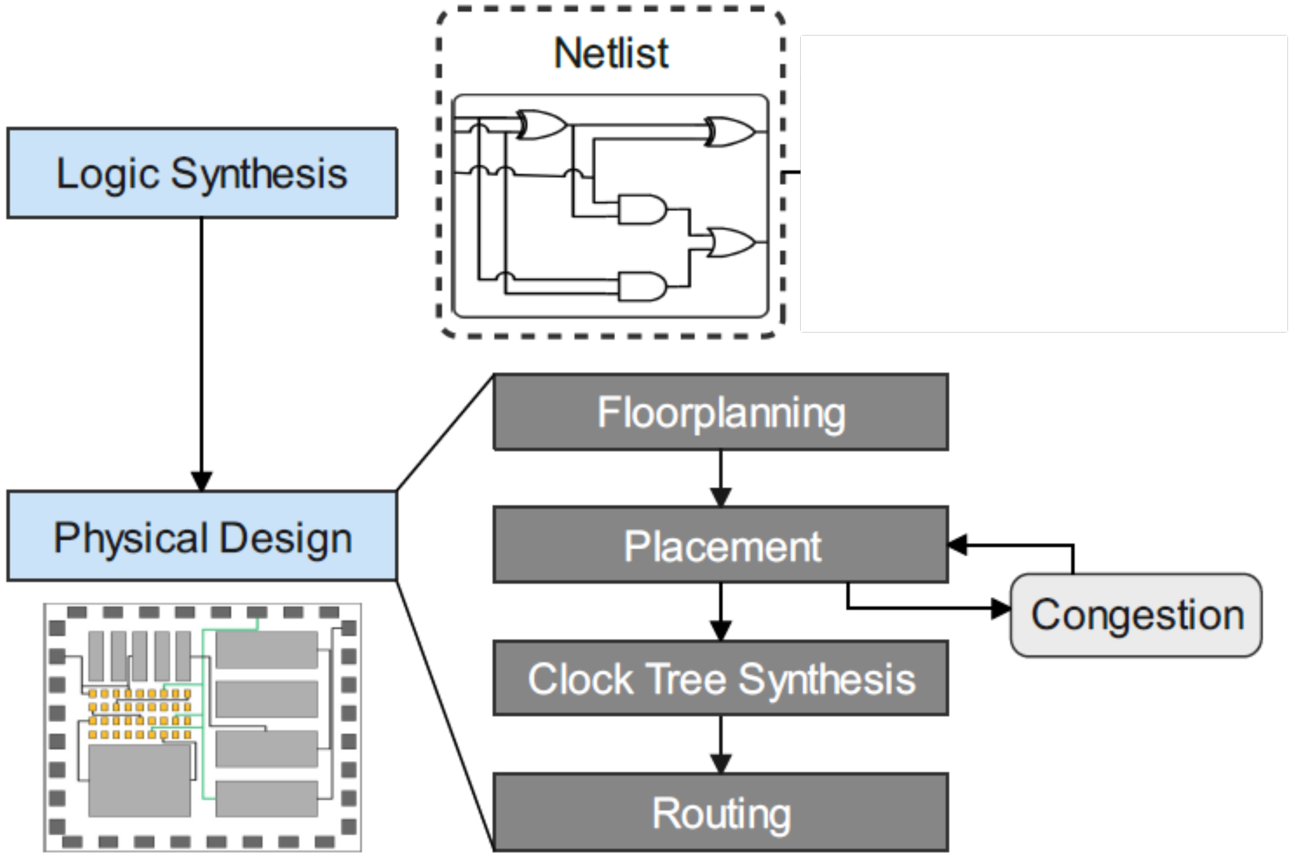
- Register Transfer Level (RTL) design: VDHL/ Verilog
- models a synchronous digital circuit in terms of
  - flow of signals between hardware registers
  - logical operations performed on signals

- Convert to physical layout through logic synthesis & physical design



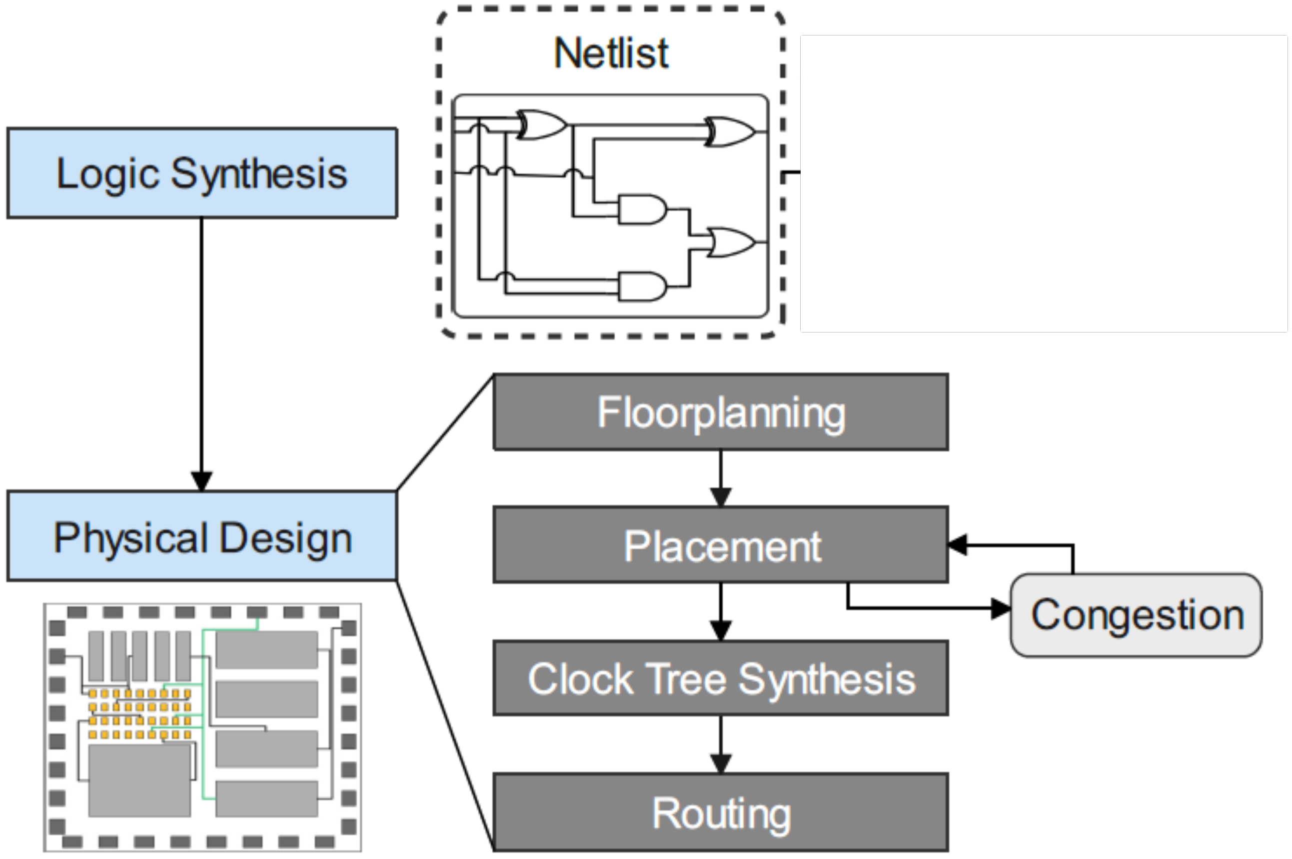
Electronic Design Automation (EDA) Workflow

- Logic synthesis
  - Convert to a netlist: contains interconnection information of all circuit elements
  - Cells: groups of transistors & interconnects that provide a Boolean logic function
- Physical design
  - All circuit elements placed on circuit boards & connected by wires



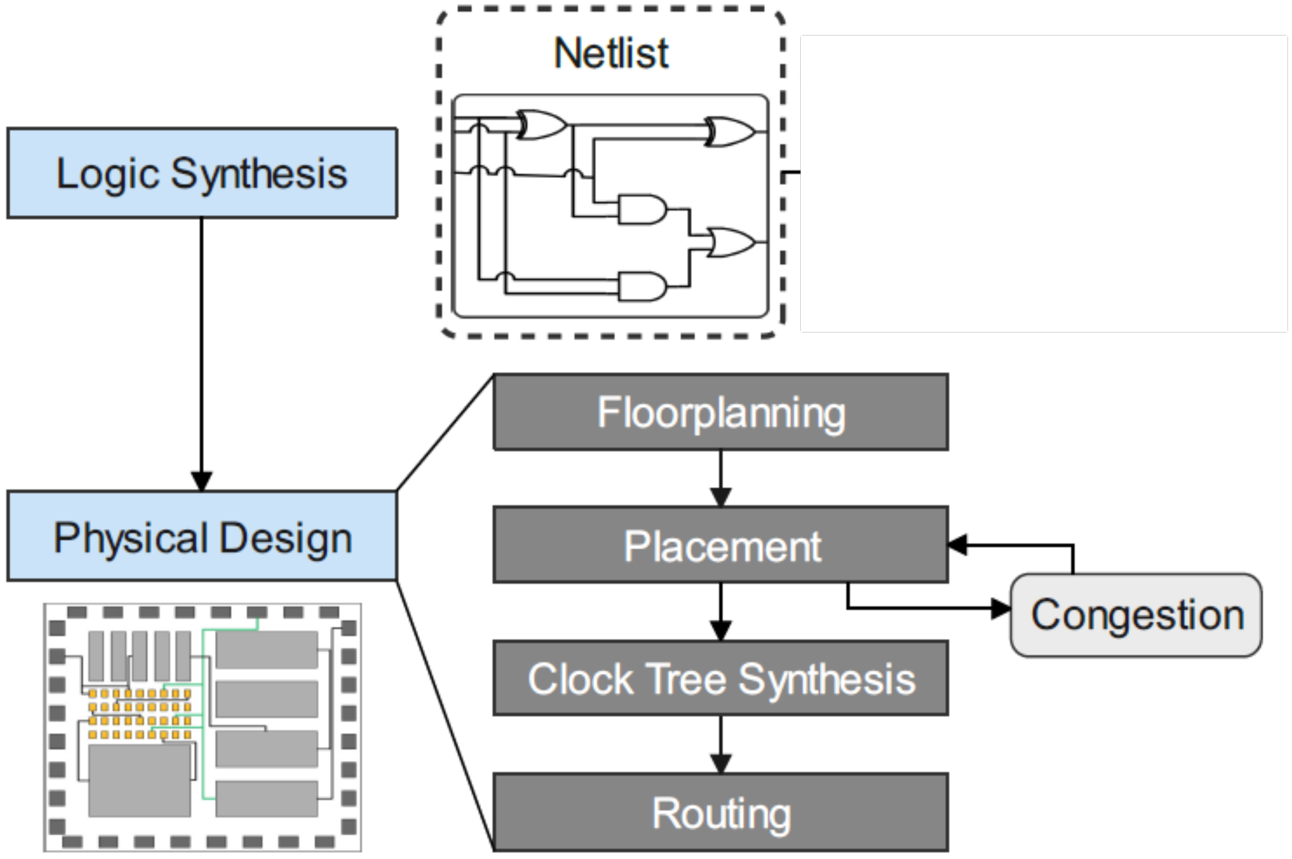
Routing Congestion

- Routing congestion: important metric that reflects the quality of the chip design
- Most EDA tools: congestion predicted AFTER cell placement
- Used as a feedback signal to optimize placement solution



Routing Congestion

- Routing congestion: important metric that reflects the quality of the chip design.
- Most EDA tools: congestion predicted AFTER cell placement
- Used as a feedback signal to optimize placement solution



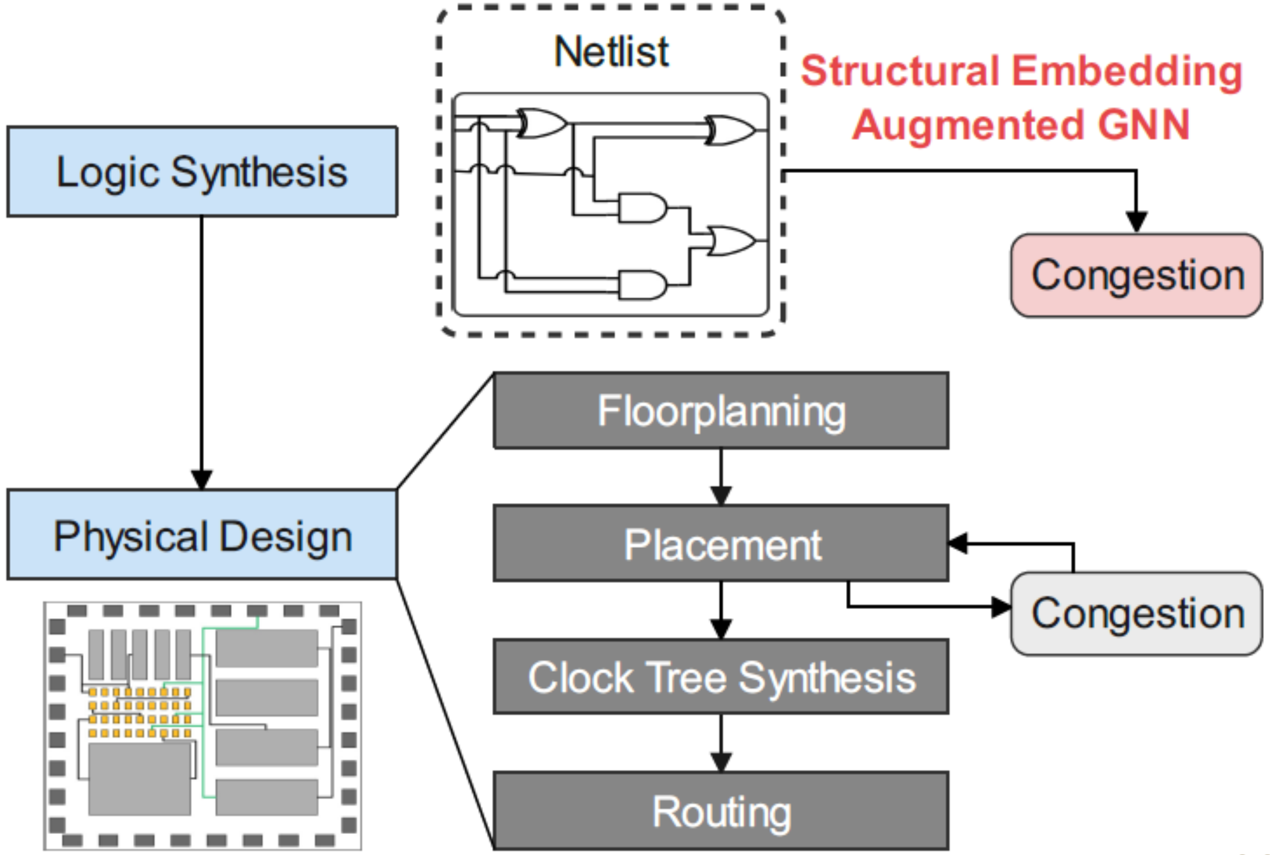
Problems:

- Large scale circuits - placement iteration is computationally expensive
- Some congestion caused by poor logic structures cannot be fixed by placement

Routing Congestion

- Goal: Estimate logic-induced congestion at logic synthesis stage
- Provide quick feedback and shorten design cycles.

- Map to node regression task
  - Train on one set of netlists (graphs)
  - Predict on another set of netlists

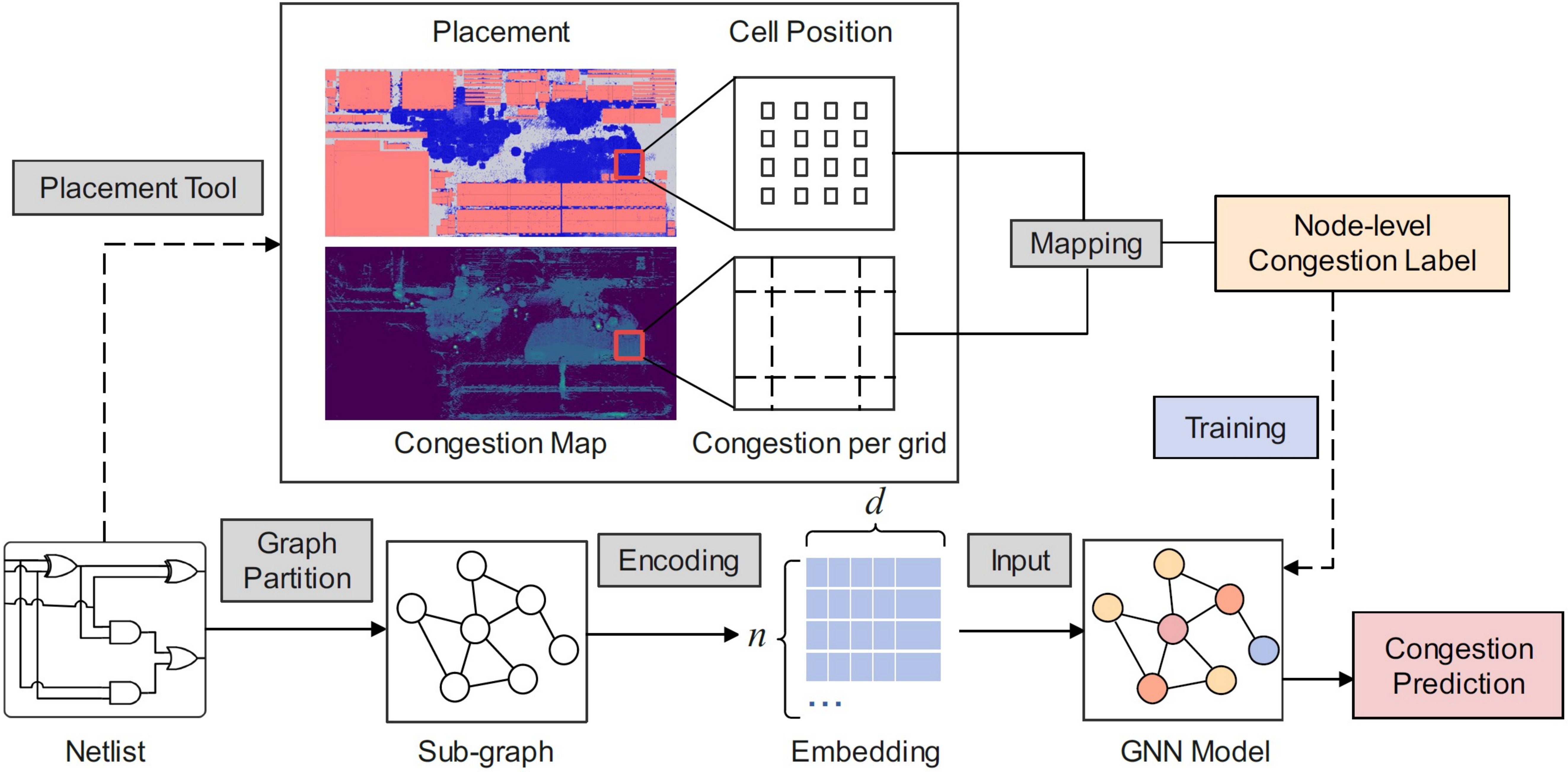


Challenge: embedding consistency

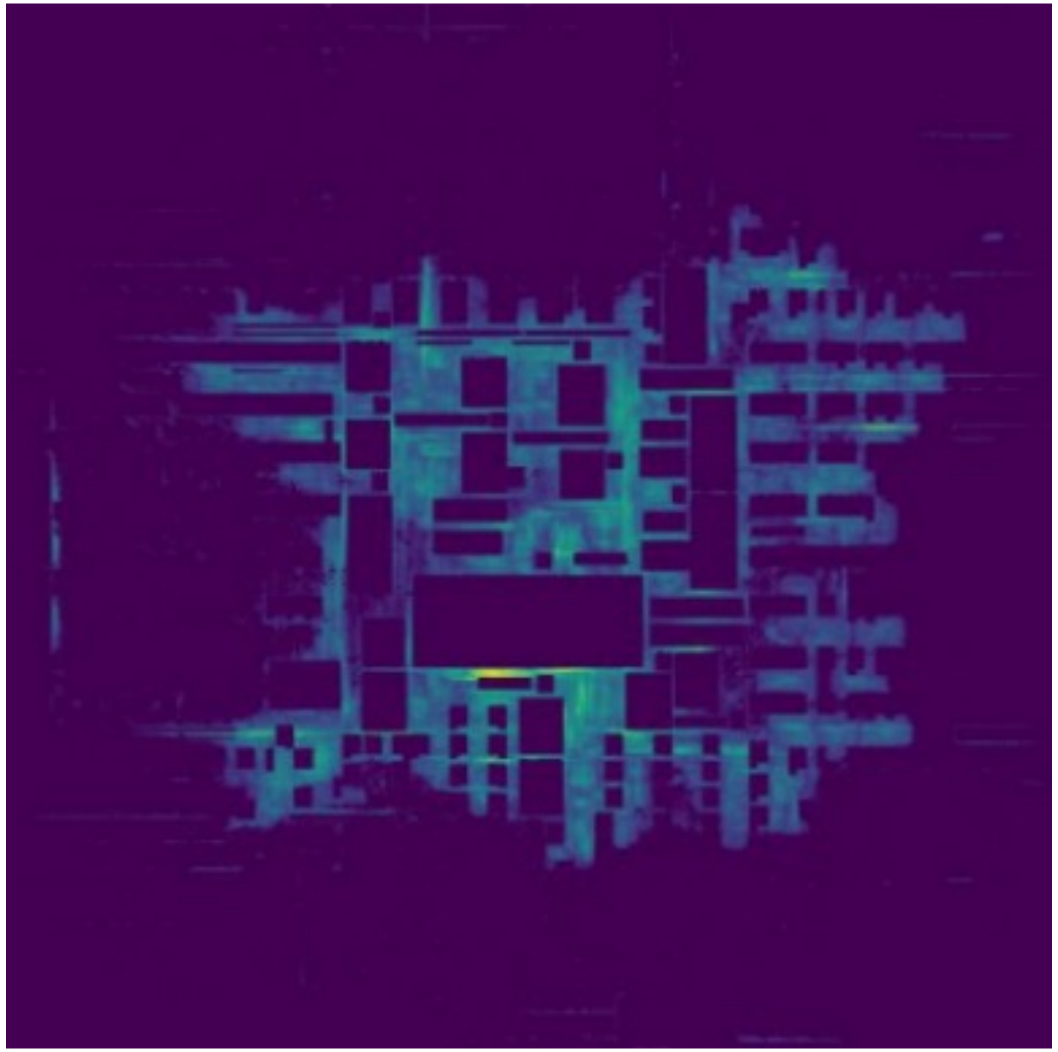
- Goal is to learn a structural embedding for each node
- Embeddings learned on one graph cannot be directly used in another distinct graph
- Need to perform alignment – inconsistent performance
- Our approach: factorization of Pointwise Mutual Information matrices XX^T
- PMT is rotation invariant ➡ no alignment necessary
- Obtain PMT via an "infinite" version of DeepWalk

Ghose et al., ICCAD 2021

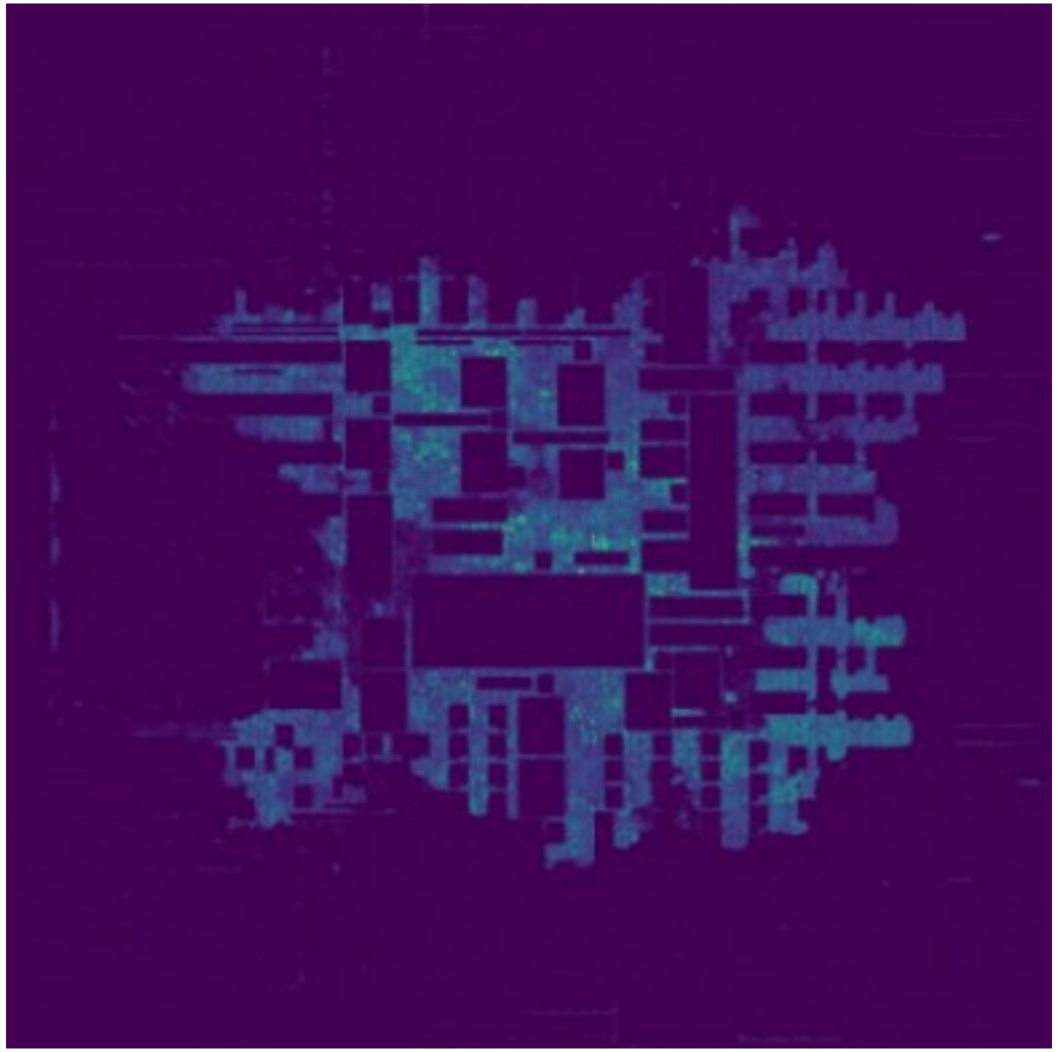
Method – Training and inference



Results – Prediction



Ground truth congestion map.



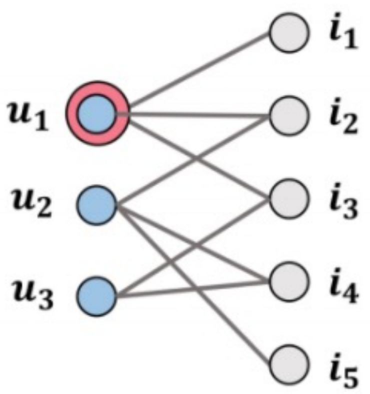
Predicted congestion map

Results – Prediction Accuracy

Methods	Lower level congestion					
	Pearson		Spearman		Kendall	
	Node	Grid	Node	Grid	Node	Grid
Adhesion metric	0.09	0.16	0.06	0.20	0.06	0.14
Neighbourhood metric	0.02	0.04	0.18	0.27	0.13	0.18
GTL metric	0.02	0.01	0.14	0.23	0.10	0.16
CongestionNet	0.26	0.35	0.27	0.33	0.19	0.24
Embedding-enhanced GNN (ours)	0.31	0.43	0.34	0.44	0.25	0.31

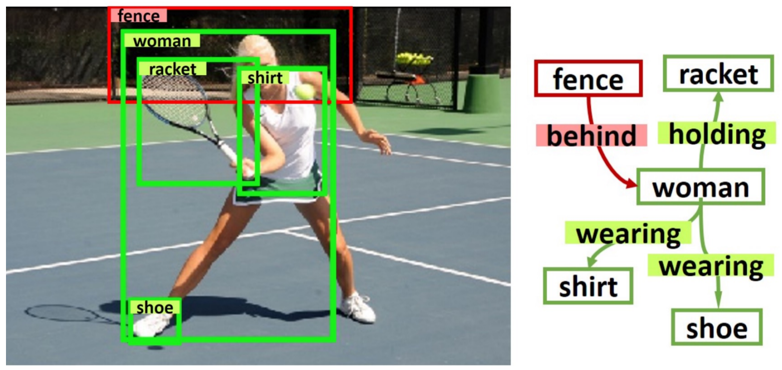
Open questions and areas of exploration

- Scalability – particularly for Bayesian or quasi-Bayesian approaches
- Continual, multi-task and streaming learning

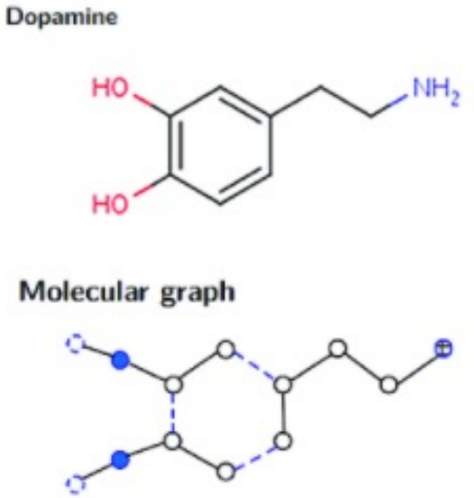


User-Item Interaction Graph

Node classification



Screen graph extraction



Molecule classification

Open questions and areas of exploration

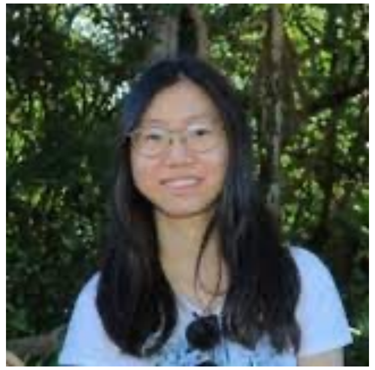
- Large language models – which graph tasks can be reformulated? Cost/benefit trade-off?
- SAT problems – graph representations of the problem and software
- Categorical sequence and graph generative models and evaluation

Florence Regol

Soumya Pal

Antonios Valkanas

Yingxue Zhang



Jianing Sun