Half-Time Seminar: Graph-Based Machine Learning for Spatio-Temporal Data

Joel Oskarsson, joel.oskarsson@liu.se

Division of Statistics and Machine Learning
Department of Computer and Information Science
Linköping University

Discussion Leader: Zheng Zhao, Uppsala University

LINKÖPING UNIVERSITY

Discussed Papers

[Paper 1] Joel Oskarsson, Per Sidén, and Fredrik Lindsten. "Scalable Deep Gaussian Markov Random Fields for General Graphs". In: *Proceedings of the 39th International Conference on Machine Learning*. PMLR. 2022

[Paper 2] Joel Oskarsson, Per Sidén, and Fredrik Lindsten. "Temporal Graph Neural Networks for Irregular Data". In: *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*. PMLR. 2023

[Paper 3] Theodor Westny, Joel Oskarsson, Björn Olofsson, and Erik Frisk. "MTP-GO: Graph-Based Probabilistic Multi-Agent Trajectory Prediction with Neural ODEs". In: *arXiv preprint* (2023)

Outline

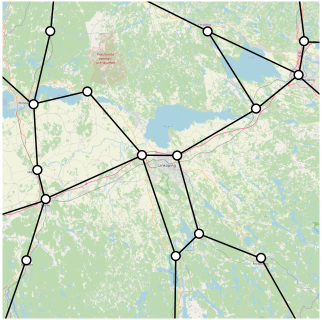
- Introduction
  - Spatio-temporal modeling
  - Graph-based machine learning
- Deep Gaussian Markov Random Fields (DGMRFs) [Paper 1]
- Temporal Graph Neural Networks (TGNNs) [Paper 2], [Paper 3]
- Summary and future research

Introduction

LINKÖPING UNIVERSITY

Modeling Spatio-Temporal Data

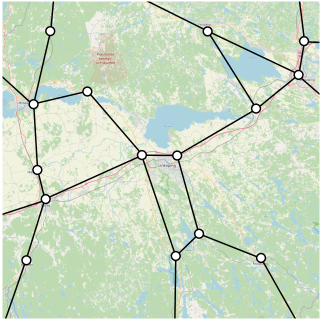
Spatial Graphs



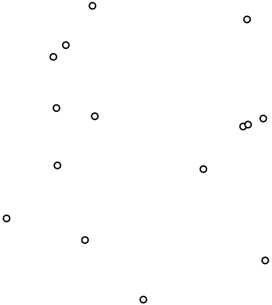
Existing spatial networks[1]

[1] © OpenStreetMap Contributors

Spatial Graphs



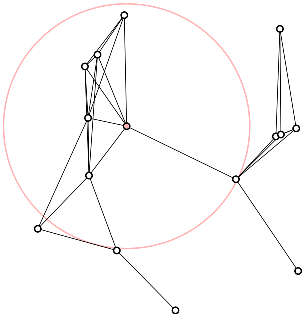
Existing spatial networks[1]



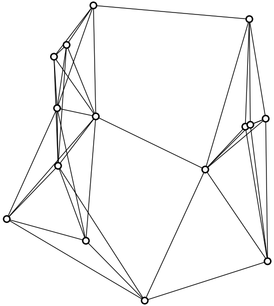
Sets of spatial points

[1] © OpenStreetMap Contributors

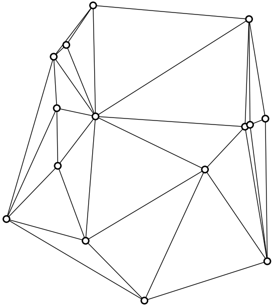
Constructing Spatial Graphs



Radius graph



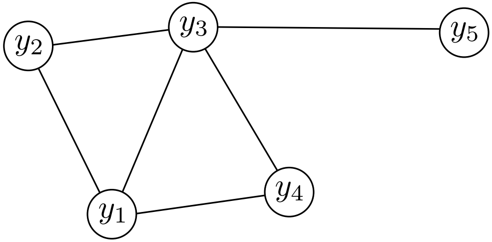
k-nearest neighbors graph
[Paper 2]



Delaunay triangulation
[Paper 1]

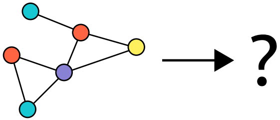
Data on Graphs

- Graph $\mathcal{G} = (V, E)$
  - $|V| = N$ nodes
  - Neigborhood $n(i)$
  - Node degree $d_i = |n(i)|$
- Observation of graph $\boldsymbol{y} \in \mathbb{R}^N$
  - y_i is observation of node i
- Node, edge and graph features



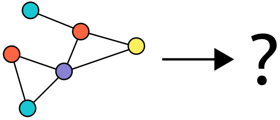
Machine Learning Problems for Graph-Structured Data

Machine Learning Problems for Graph-Structured Data

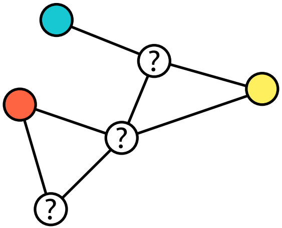


Graph-level prediction

Machine Learning Problems for Graph-Structured Data

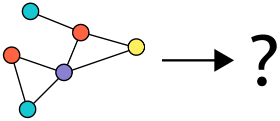


Graph-level prediction

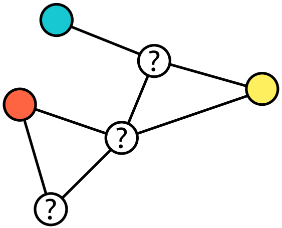


Node-level prediction

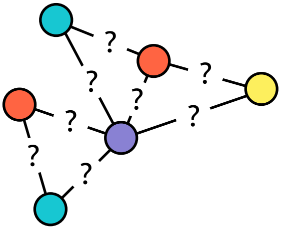
Machine Learning Problems for Graph-Structured Data



Graph-level prediction



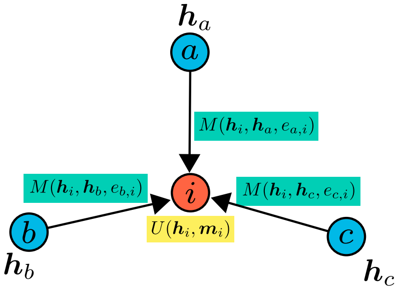
Node-level prediction



Structure learning

Graph Neural Networks (GNNs)

- Message Passing Neural Networks (MPNNs)[1]
- $\boldsymbol{h}_i^{(l)}$ is representation of node i after layer l



$n(i) = \{a, b, c\}$

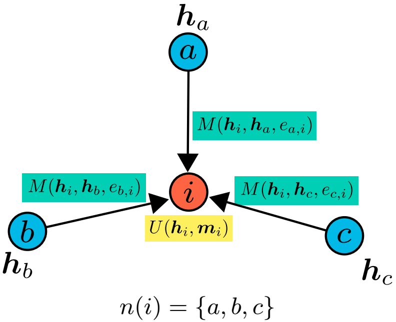
[1] Justin Gilmer et al. "Neural Message Passing for Quantum Chemistry". In: *Proceedings of the 34th International Conference on Machine Learning*. PMLR. 2017

Graph Neural Networks (GNNs)

- Message Passing Neural Networks (MPNNs)[1]
- $\boldsymbol{h}_i^{(l)}$ is representation of node i after layer l

$$\boldsymbol{m}_i^{(l+1)} = \sum_{j \in n(i)} M\left(\boldsymbol{h}_i^{(l)}, \boldsymbol{h}_j^{(l)}, e_{j,i}\right)$$

$$\boldsymbol{h}_i^{(l+1)} = U\left(\boldsymbol{h}_i^{(l)}, \boldsymbol{m}_i^{(l+1)}\right)$$



[1] Justin Gilmer et al. "Neural Message Passing for Quantum Chemistry". In: *Proceedings of the 34th International Conference on Machine Learning*. PMLR. 2017

GNNs: GraphConv Example

- MPNN:

$$\boldsymbol{m}_i^{(l+1)} = \sum_{j \in n(i)} M\left(\boldsymbol{h}_i^{(l)}, \boldsymbol{h}_j^{(l)}, e_{j,i}\right)$$

$$\boldsymbol{h}_i^{(l+1)} = U\left(\boldsymbol{h}_i^{(l)}, \boldsymbol{m}_i^{(l+1)}\right)$$

- MPNN example: GraphConv[1]

$$\boldsymbol{m}_i^{(l+1)} = \sum_{j \in n(i)} \frac{e_{j,i}}{d_i} W_M \boldsymbol{h}_j^{(l)}$$

$$\boldsymbol{h}_i^{(l+1)} = W_U \boldsymbol{h}_i^{(l)} + \boldsymbol{m}_i^{(l+1)} + \boldsymbol{b}_U$$

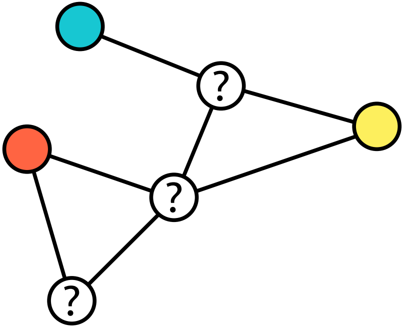
  – Used in [Paper 2] and [Paper 3]

[1] Christopher Morris et al. "Weisfeiler and leman go neural: Higher-order graph neural networks". In: *Proceedings of the AAAI conference on artificial intelligence*. 2019

Deep Gaussian Markov Random Fields

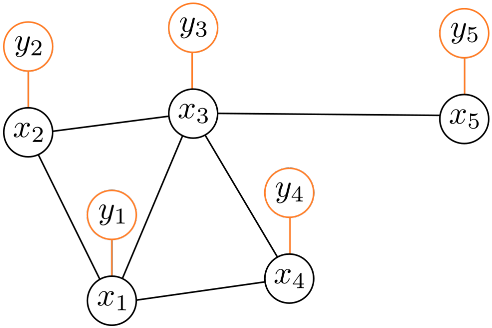
LINKÖPING UNIVERSITY

Setting of [Paper 1]



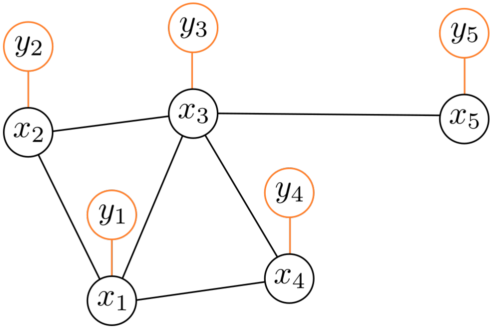
- Node prediction
- Single, fully known graph
- Subset of nodes observed

Probabilistic Model



$$p(\boldsymbol{x}, \boldsymbol{y}) = p(\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x})$$
$$= p(\boldsymbol{x}) \prod_{i=1}^{N} p(y_i|x_i)$$

Probabilistic Model



$$p(\boldsymbol{x}, \boldsymbol{y}) = p(\boldsymbol{x})p(\boldsymbol{y}|\boldsymbol{x})$$
$$= p(\boldsymbol{x}) \prod_{i=1}^{N} p(y_i|x_i)$$

- Gaussian prior: $p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, Q^{-1})$
- Likelihood: $p(y_i|x_i) = \mathcal{N}(y_i|x_i, \sigma^2)$

- Posterior of interest: $p(\boldsymbol{x}|\boldsymbol{y}_{\text{Obs.}})$

Deep Gaussian Markov Random Field (DGMRF)[1] Prior

- Define Gaussian prior using affine map g

$$\boldsymbol{z} = g(\boldsymbol{x}) = G\boldsymbol{x} + \boldsymbol{b}, \quad \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, I) \qquad \Rightarrow \quad \boldsymbol{x} \sim \mathcal{N}\left(\boldsymbol{\mu}, (G^{\intercal}G)^{-1}\right)$$

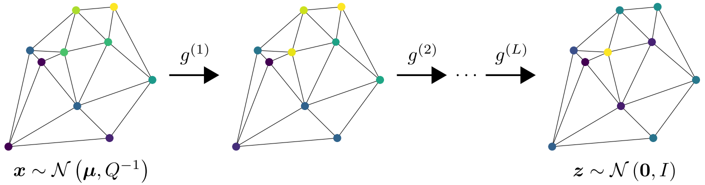
[1] Per Sidén and Fredrik Lindsten. "Deep Gaussian Markov Random Fields". In: *Proceedings of the 37th International Conference on Machine Learning*. PMLR. 2020

Deep Gaussian Markov Random Field (DGMRF)[1] Prior

- Define Gaussian prior using affine map g

$$\boldsymbol{z} = g(\boldsymbol{x}) = G\boldsymbol{x} + \boldsymbol{b}, \quad \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, I) \qquad \Rightarrow \qquad \boldsymbol{x} \sim \mathcal{N}\left(\boldsymbol{\mu}, (G^\intercal G)^{-1}\right)$$

- Layered structure $g(\boldsymbol{x}) = g^{(L)} \circ g^{(L-1)} \circ \cdots \circ g^{(1)}(\boldsymbol{x})$

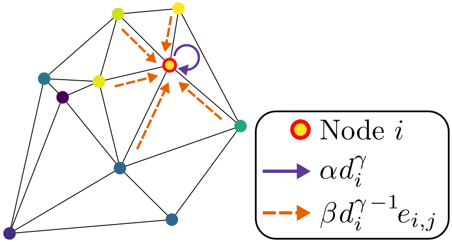


[1] Per Sidén and Fredrik Lindsten. "Deep Gaussian Markov Random Fields". In: *Proceedings of the 37th International Conference on Machine Learning*. PMLR. 2020

A DGMRF Layer for Graphs[1]

$$h_i^{(l+1)} = g^{(l+1)}\left(\boldsymbol{h}^{(l)}\right)_i$$
$$= b + \alpha d_i^{\gamma} h_i^{(l)} + \sum_{j \in n(i)} \beta d_i^{\gamma-1} e_{j,i} h_j^{(l)}$$

- Trainable parameters α, β, γ and b



[1] [Paper 1] Joel Oskarsson, Per Sidén, and Fredrik Lindsten. “Scalable Deep Gaussian Markov Random Fields for General Graphs”. In: *Proceedings of the 39th International Conference on Machine Learning.* PMLR. 2022

DGMRF Layer as a MPNN

- DGMRF layer for graphs

$$m_i^{(l+1)} = \sum_{j \in n(i)} \beta d_i^{\gamma - 1} e_{j,i} h_j^{(l)}$$

$$h_i^{(l+1)} = b + \alpha d_i^{\gamma} h_i^{(l)} + m_i^{(l+1)}$$

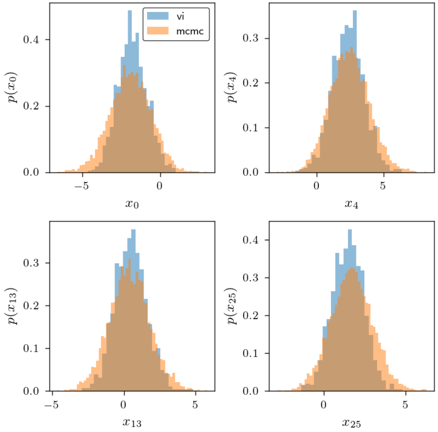
- MPNN

$$\boldsymbol{m}_i^{(l+1)} = \sum_{j \in n(i)} M\left(\boldsymbol{h}_i^{(l)}, \boldsymbol{h}_j^{(l)}, e_{j,i}\right)$$

$$\boldsymbol{h}_i^{(l+1)} = U\left(\boldsymbol{h}_i^{(l)}, \boldsymbol{m}_i^{(l+1)}\right)$$

Training and Inference

- Training
  - Variational objective, ELBO maximization
- Posterior Inference for $p(\boldsymbol{x}|\boldsymbol{y}_{\text{Obs.}})$
  - Conjugate prior, analytical form



Results: California Housing Data

- House values in 20 640 blocks in California[1]
- 50% of nodes observed
- Additional node features

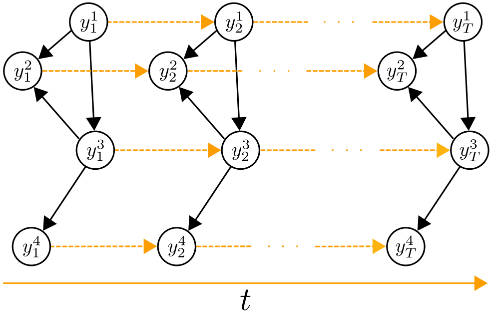
Model	RMSE	CRPS
Bayes LR	8.872	4.834
MLP	7.094	4.525
GCN	6.837	4.273
GAT	6.788	4.348
SVGP	7.287	3.930
DGMRF, $L = 1$	5.894	3.078
DGMRF, $L = 2$	5.810	3.041
DGMRF, $L = 3$	**5.804**	**3.039**

[1] R. Kelley Pace and Ronald Barry. "Sparse spatial autoregressions". In: *Statistics & Probability Letters* 33 (1997)

Temporal Graph Neural Networks

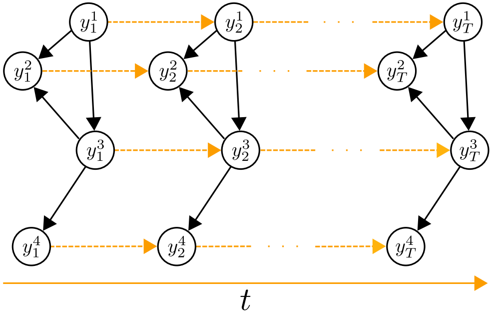
LINKÖPING UNIVERSITY

Graph-Structured Time Series



- Temporal Graph Neural Networks (TGNNs)
  - GNN over graph
  - Sequence model over time

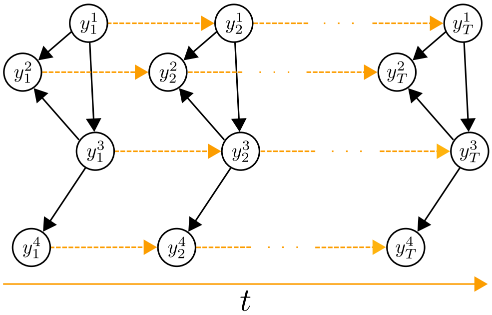
Graph-Structured Time Series



- Temporal Graph Neural Networks (TGNNs)
  - GNN over graph
  - Sequence model over time
- Gated Recurrent Unit (GRU)

$$\boldsymbol{h}_t = \mathrm{GRU}(\boldsymbol{h}_{t-1}, \boldsymbol{y}_t)$$

Graph-Structured Time Series



- Temporal Graph Neural Networks (TGNNs)
  - GNN over graph
  - Sequence model over time
- Gated Recurrent Unit (GRU)

$$\boldsymbol{h}_t = \text{GRU}(\boldsymbol{h}_{t-1}, \boldsymbol{y}_t)$$

  - Graph-GRU[1]:

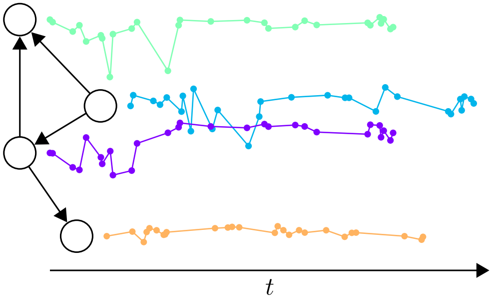
$$\boldsymbol{h}_t^i = \text{GRU}\left(\left\{\left(\boldsymbol{h}_{t-1}^j, y_t^j\right)\right\}_{j \in \{i\} \cup n(i)}\right)$$

  - Used in [Paper 2] and [Paper 3]

[1] Xujiang Zhao, Feng Chen, and Jin-Hee Cho. "Deep learning for predicting dynamic uncertain opinions in network data". In: *IEEE International Conference on Big Data*. 2018

[Paper 2]: TGNNs for Irregular Data[1]

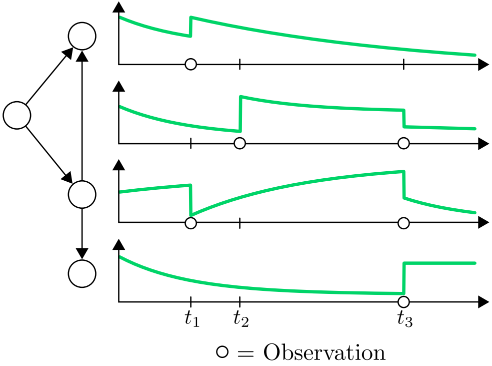
- Forecasting problems
- Irregular observations
  - Irregular time steps
  - Observing subset of nodes



[1] [Paper 2] Joel Oskarsson, Per Sidén, and Fredrik Lindsten. "Temporal Graph Neural Networks for Irregular Data". In: *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*. PMLR. 2023

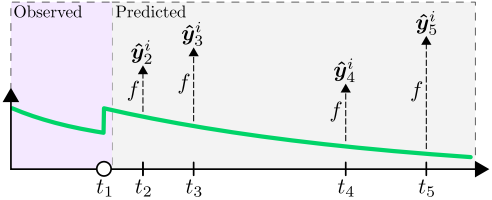
Time-Continuous Latent States

- Time-continuous latent state in each node
- Dynamics in-between observations
  - Linear ODE
- State update at observation
  - Graph-GRU



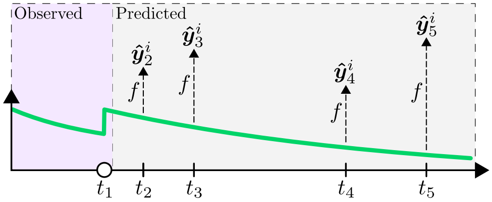
Prediction and Results

- Forecasting by rolling out latent states



Prediction and Results

- Forecasting by rolling out latent states

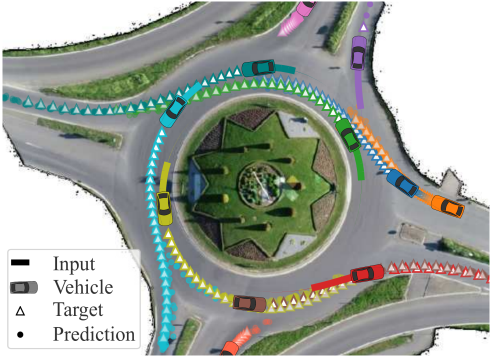


- Evaluation
  - Traffic data (PEMS-BAY)
  - Climate data (USHCN)

Model	PEMS-BAY	USHCN
Predict Previous	26.32	16.88
GRU-D	8.79	8.03
Transformer	12.05	7.36
LG-ODE	27.00	-
TGNN4I	**7.10**	**6.72**

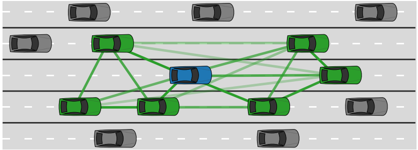
Multi-Agent Trajectory Prediction

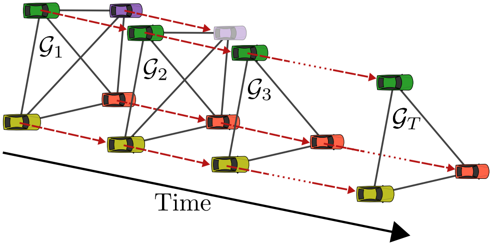
- The trajectory prediction problem
  - Given historical positions
  - Predict future trajectory
- Multi-agent trajectory prediction



[Paper 3]: TGNNs for Multi-Agent Trajectory Prediction[1]

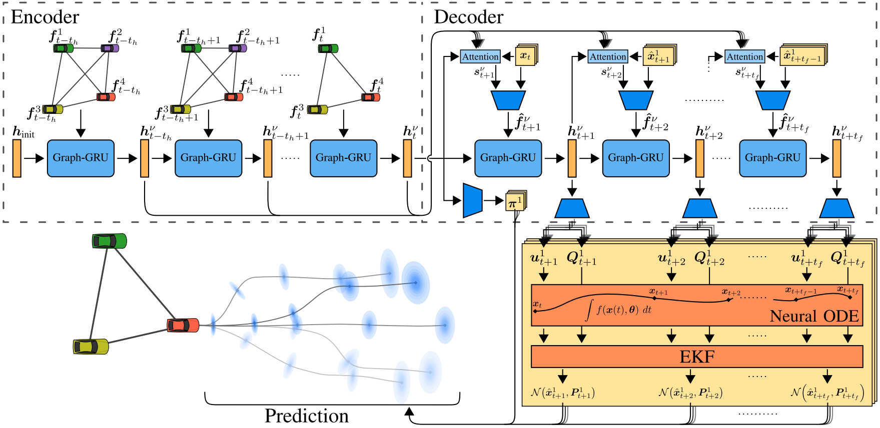
- Agent interactions
- Traffic situation as graph



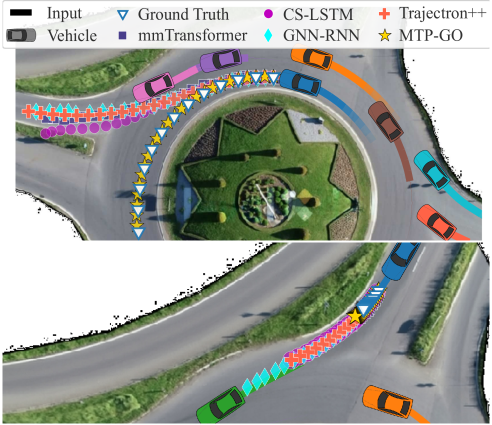


[1] [Paper 3] Theodor Westny, Joel Oskarsson, Björn Olofsson, and Erik Frisk. "MTP-GO: Graph-Based Probabilistic Multi-Agent Trajectory Prediction with Neural ODEs". In: *arXiv preprint* (2023)

The MTP-GO Model



Results: Roundabout Scenario



Model	ADE	FDE	ANLL	FNLL
Constant Acc.	4.83	16.2	—	—
Constant Vel.	6.49	17.1	—	—
Seq2Seq	1.46	3.66	—	—
S-LSTM	1.20	3.47	1.75	5.12
CS-LSTM	1.19	3.57	2.09	5.54
GNN-RNN	1.13	3.11	—	—
mmTransformer	1.29	3.50	—	—
Trajectron++	1.09	3.53	—	—
MTP-GO	**0.92**	**2.97**	**−0.22**	**3.85**

Summary and Future Research

LINKÖPING
UNIVERSITY

Summary

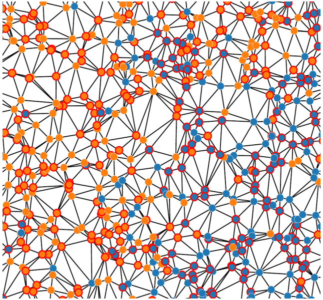
[Paper 1] DGMRF model for graph-structured data
- Probabilistic model for node-level prediction

[Paper 2] TGNNs for irregular data
- Time-continuous latent states

[Paper 3] TGNNs for multi-agent trajectory prediction
- Traffic agent interactions as graphs

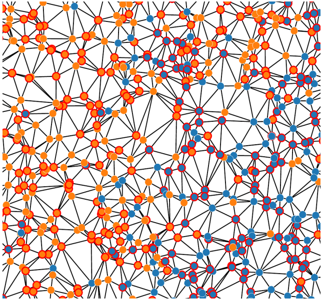
Future Research

Future Research

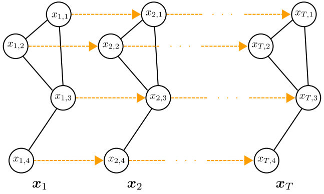


[Paper 1] extension:
Node classification

Future Research

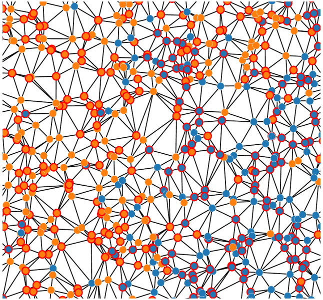


[Paper 1] extension:
Node classification

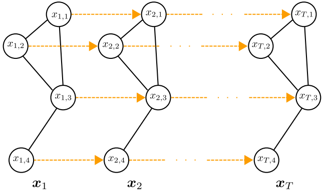


[Paper 1] extension:
Spatio-temporal data

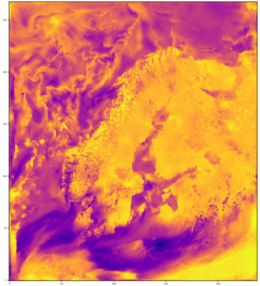
Future Research



[Paper 1] extension:
Node classification



[Paper 1] extension:
Spatio-temporal data



Application:
Climate and weather modeling

Joel Oskarsson, `joel.oskarsson@liu.se`

www.liu.se

LINKÖPING UNIVERSITY