UNIVERSITY AT ALBANY
State University of New York

VIRGINIA TECH

Deep Learning for Predicting Dynamic Uncertain Opinions in Network Data

Xujiang Zhao[1], Feng Chen[1], Jin-Hee Cho[2]

[1]University at Albany, [2]Virginia Tech

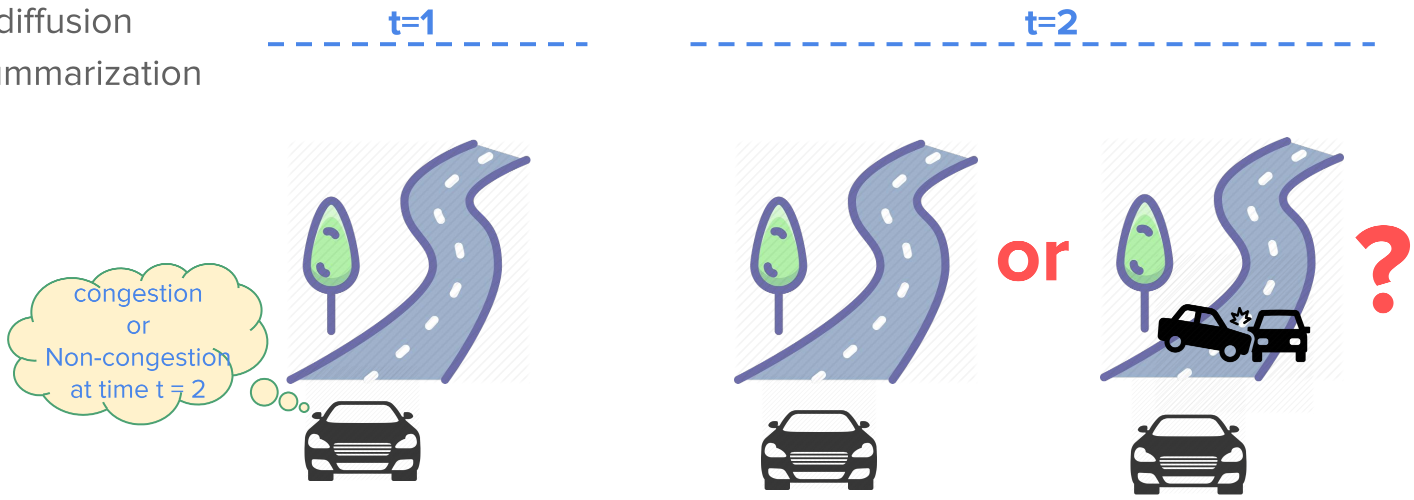
Outline

- Motivation
- Research Problem & Challenge
- Graph Convolutional Networks
- Proposed Approach
- Experimental Results
- Conclusion & Future Work

Motivation

- How do we make decisions with subjective, uncertain opinions?
- Applications
  - Trust in social networks
  - Opinion diffusion
  - Graph summarization

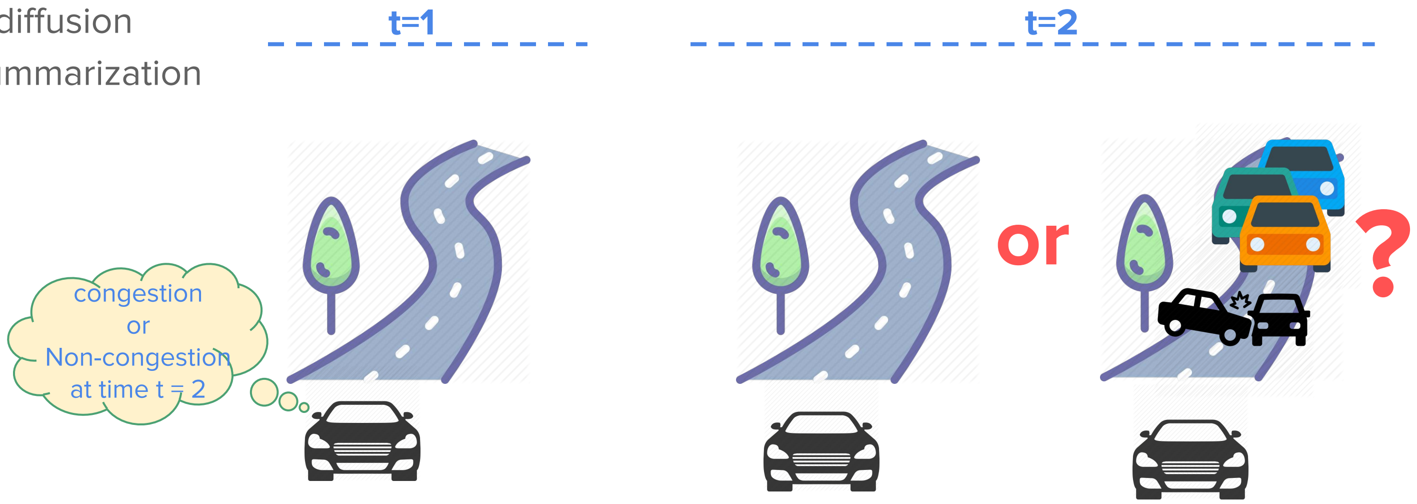
❖ In a traffic network, you want to know if a road will be congested or not in the near future.



Motivation

- How do we make decisions with subjective, uncertain opinions?
- Applications
  - Trust in social networks
  - Opinion diffusion
  - Graph summarization

❖ In a traffic network, you want to know if a road will be congested or not in the near future.



There is some **uncertainty component** that we need to consider while predicting the condition of a road link.

Binomial Opinion in Subjective Logic (SL)

How to model the uncertainty of a binary prediction?

Binomial Opinion

A binomial opinion ω is represented by

$$\omega = (b, d, u, a)$$

Where:

- b: belief (e.g., agree)
- d: disbelief (e.g., disagree)
- u: uncertainty (i.e., ignorance, vacuity, or lack of evidence)
- a: base rate (i.e., a prior knowledge)

$$b + d + u = 1$$

For example: an opinion of one road congestion could be $\omega = (0.4, 0.1, 0.5, 0.5)$, but if *we don't consider uncertainty*, the probability of congestion and non-congestion is: (0.8, 0.2)

Binomial Opinion in Subjective Logic (SL)

How to predict an binomial opinion?

Case 1: we have some recent observations of a road link.

Binomial Opinion

A binomial opinion ω is represented by

$$\omega = (b, d, u, a)$$

Where:

- b: belief (e.g., agree)
- d: disbelief (e.g., disagree)
- u: uncertainty (i.e., ignorance, vacuity, or lack of evidence)
- a: base rate (i.e., a prior knowledge, set a = 0.5)
- W: represents the amount of uncertain evidence, set W=2

$$b + d + u = 1$$

For example: Suppose we have several recent observations, 4 congestion, 2 non-congestion:

Evidence number: 2+4+W=8

$b = 4/8 = 0.5$

$d = 2/8 = 0.25$

$u = W/8=0.25$

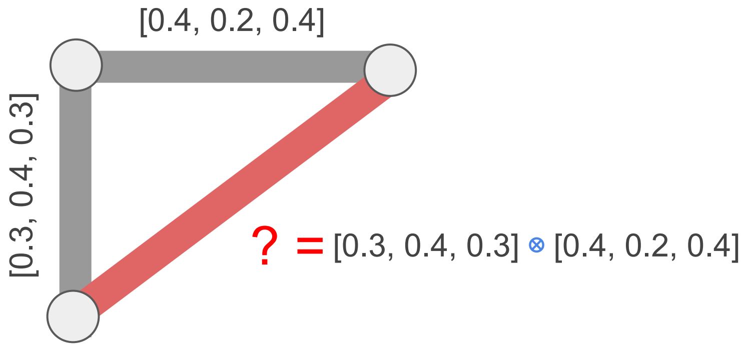
$\omega = (0.5, 0.25, 0.25, 0.5)$

Fusion Operators with Uncertain Opinions in SL

Case 2: we have some road links that do not have any observations.

We need to predict their uncertain opinions based on the fusion of the uncertain opinions of their nearby road links in the network that have observations.

- Discount operator ⊗: Discount trust of an entity one wants to interact when it does not have any direct interaction with the entity e.g.

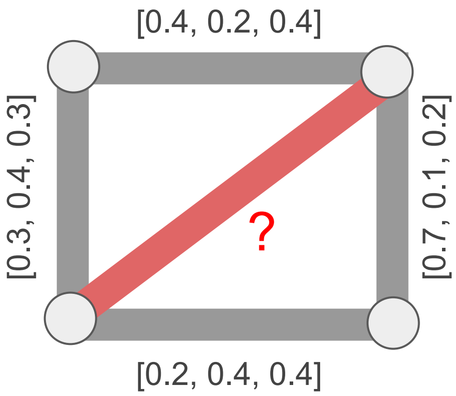


Fusion Operators with Uncertain Opinions in SL

Case 2: we have some road links that do not have any observations.

We need to predict their uncertain opinions based on the fusion of the uncertain opinions of their nearby road links in the network that have observations.

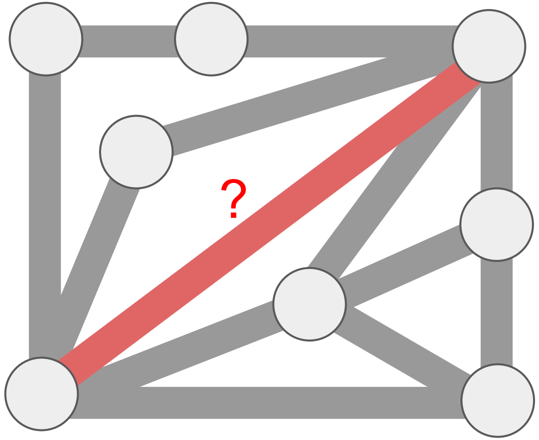
❖ Consensus operator ⊕: Find a consensus between two opinions where two entities observe a same entity, e.g.



$? = [0.3, 0.4, 0.3] \otimes [0.4, 0.2, 0.4] \oplus [0.2, 0.4, 0.4] \otimes [0.7, 0.1, 0.2]$

Scalability Issue in Subjective Logic

When a network is large, there are too many paths to consider for fusing them



Limitation:

SL's operators are good for fusing two opinions in dyadic relationships;
However, they are not scalable for fusing multiple opinions as large network data.

Why Deep Learning?

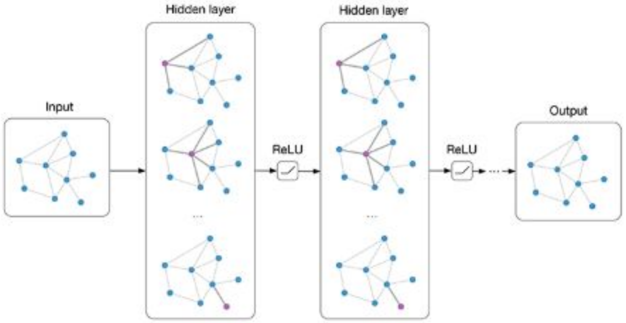
Both SL and CSL are:

- not scalable.
- not effectively dealing with heterogeneous data.

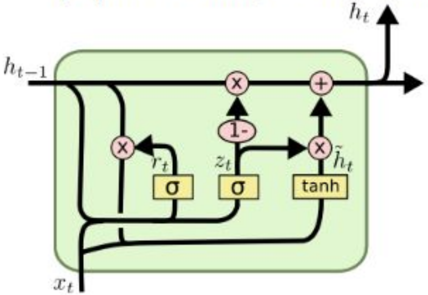
How to Solve These Challenge?

Graph Convolutional Network and Gated Recurrent Units can provide solutions for

- ❖ dealing with graph network data
- ❖ modeling topological and temporal heterogeneous dependency
- ❖ processing large-scale data (i.e., scalability)



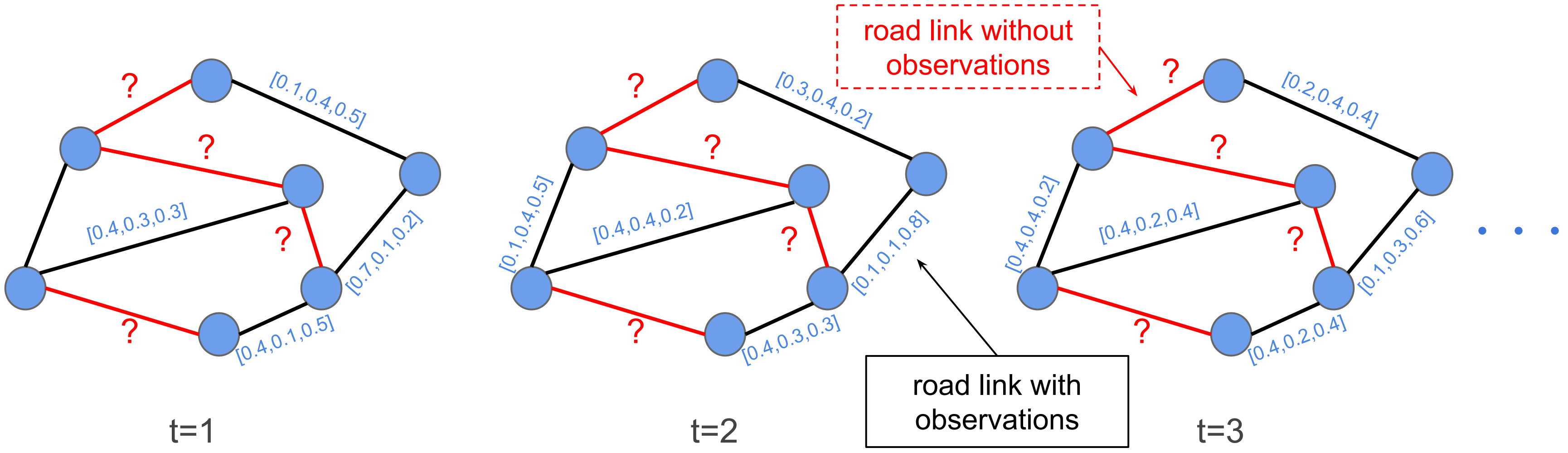
[Kipf & Welling, ICLR 2017]



[Cho, etc., arXiv 2014]

Research Problem & Challenges

Given an network G = (V, E), some of the edges in E have known uncertain opinions, the goal is to predict the uncertain opinions of the other edges.



How to accurately and efficiently predict unknown **dynamic** opinions with a large, heterogeneous, **uncertain** network data?

Research Goal & Contributions

Research goal: Develop a scalable, effective Deep Learning (DL)-based **dynamic** opinion inference algorithm for a large, heterogeneous, uncertain network data.

Key Contributions:

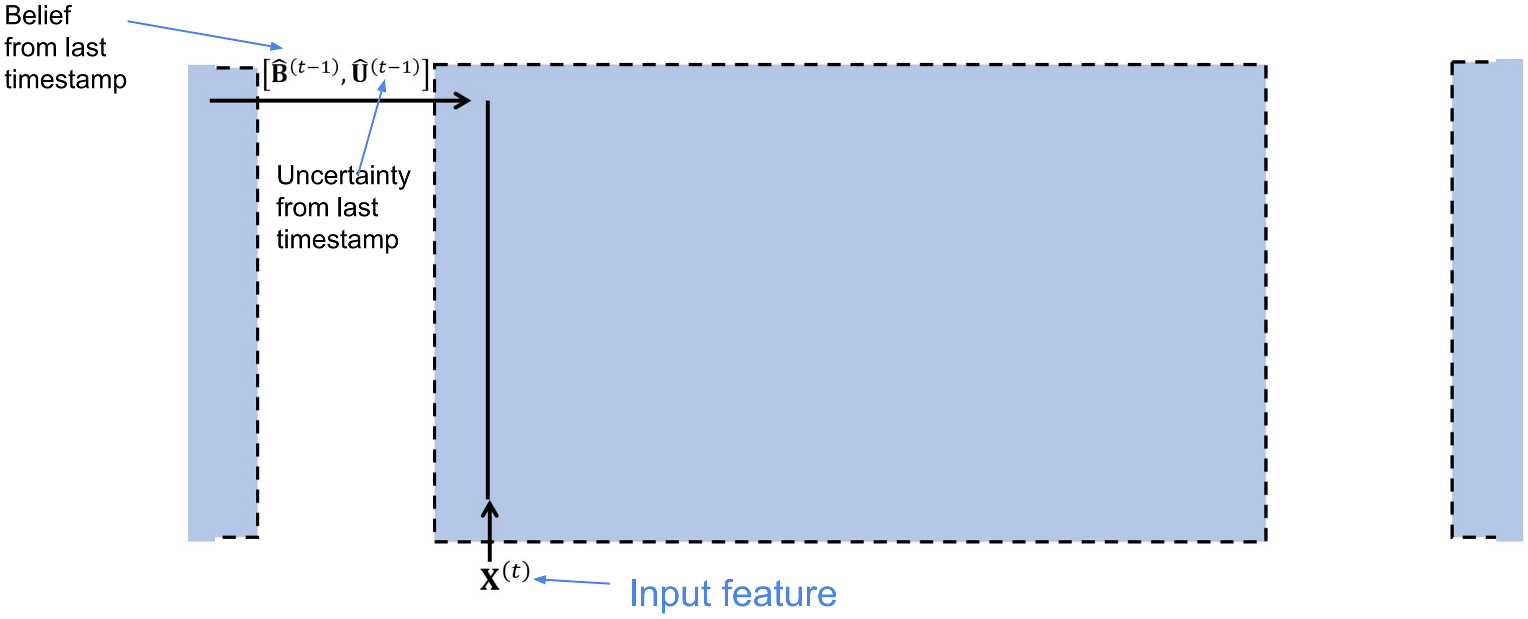
- Considered GCN and GRU for modeling the **topological** and **temporal** heterogeneous dependency information.
- Modeled **conflicting** opinions based on robust statistics.
- Developed a highly scalable inference algorithm to predict dynamic, uncertain opinions in a **linear** computation time.

Proposed Approach: GCN-GRU-based Opinion

How to model *topological* and *temporal* heterogeneous dependency?

- ❖ GCN can model *topological* heterogeneous dependency.
- ❖ GRU can model *temporal* heterogeneous dependency.

How to design a GRU cell based on a Graph convolution process?

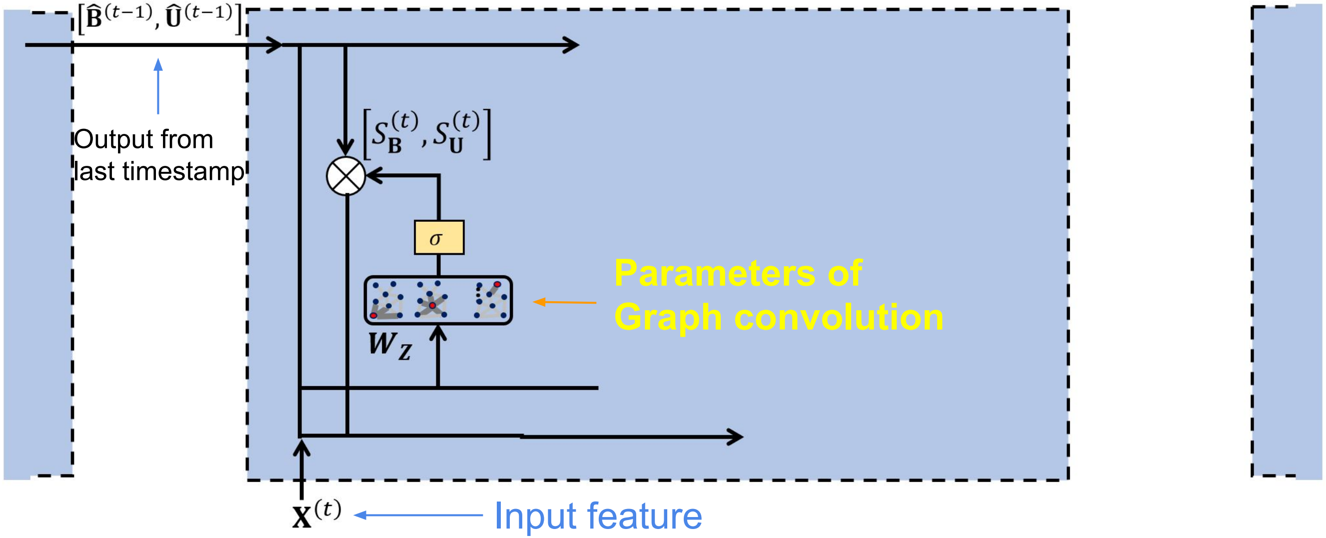


Proposed Approach: GCN-GRU-based Opinion

- The *reset gate* is obtained by

$$\left[\mathbf{S}_{\mathbf{B}}^{(t)}, \mathbf{S}_{\mathbf{U}}^{(t)}\right] = \sigma\left(\left[g_{\mathbf{W_S}} \star [\hat{\mathbf{B}}^{(t-1)}, \mathbf{X}^{(t)}], g_{\mathbf{W_S}} \star [\hat{\mathbf{U}}^{(t-1)}, \mathbf{X}^{(t)}]\right]\right)$$

Graph convolution operator

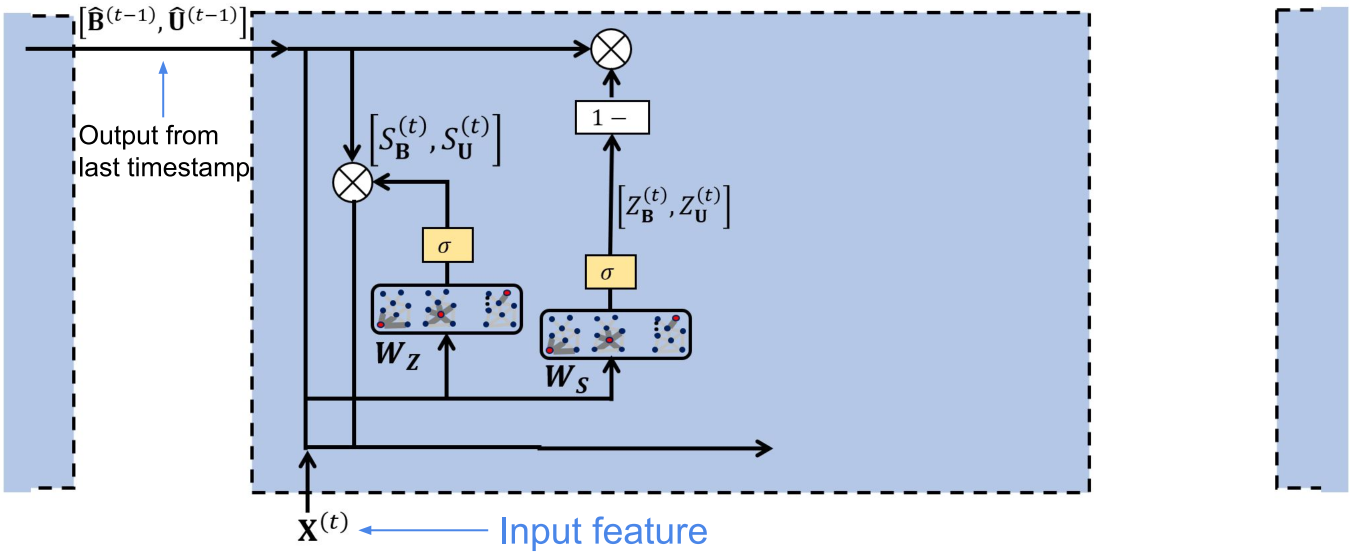


Proposed Approach: GCN-GRU-based Opinion

- The *update gates* $\mathbf{Z}_{\mathbf{B}}^{(t)}$ and $\mathbf{Z}_{\mathbf{U}}^{(t)}$ are computed by

$$\left[\mathbf{Z}_{\mathbf{B}}^{(t)}, \mathbf{Z}_{\mathbf{U}}^{(t)}\right] = \sigma\left(\left[g_{\mathbf{W}_{\mathbf{Z}}} \star [\hat{\mathbf{B}}^{(t-1)}, \mathbf{X}^{(t)}], g_{\mathbf{W}_{\mathbf{Z}}} \star [\hat{\mathbf{U}}^{(t-1)}, \mathbf{X}^{(t)}]\right]\right),$$

Graph convolution operator



Proposed Approach: GCN-GRU-based Opinion

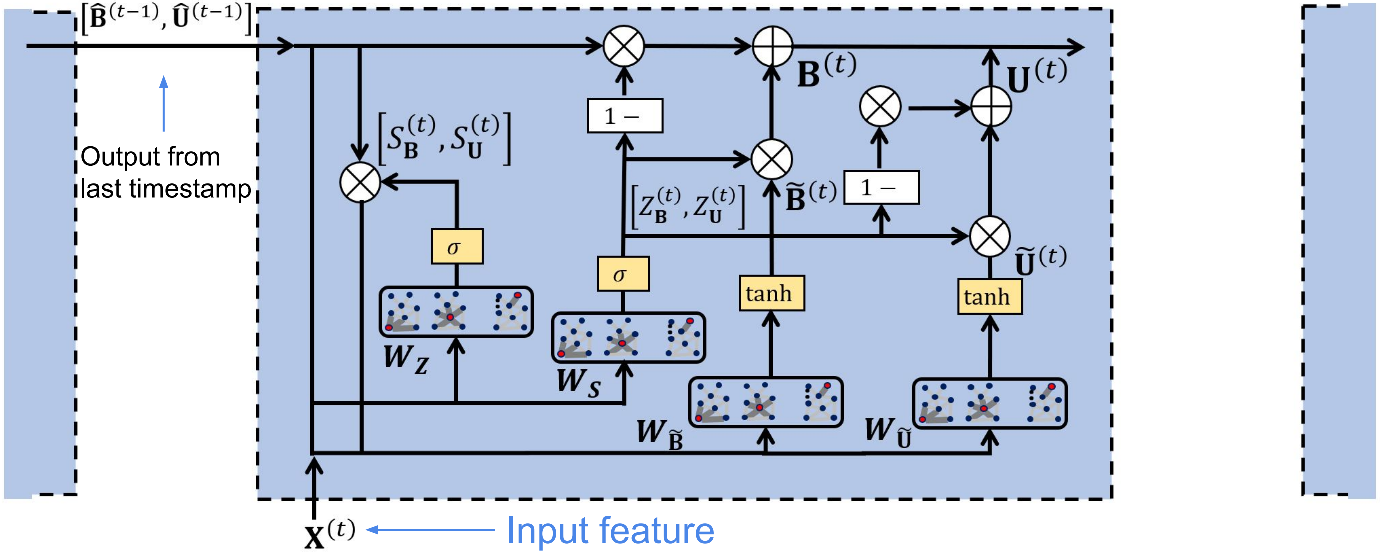
- *Current information* is given by

Graph convolution operator

$$\tilde{\mathbf{B}}^{(t)} = \tanh\left(g_{\mathbf{W}_{\tilde{\mathbf{B}}}} \star \left[\mathbf{S}_{\mathbf{B}}^{(t)} \odot \hat{\mathbf{B}}^{(t-1)}, \mathbf{X}^{(t)}\right]\right),$$

$$\tilde{\mathbf{U}}^{(t)} = \tanh\left(g_{\mathbf{W}_{\tilde{\mathbf{U}}}} \star \left[\mathbf{S}_{\mathbf{U}}^{(t)} \odot \hat{\mathbf{U}}^{(t-1)}, \mathbf{X}^{(t)}\right]\right)$$

Element-wise product



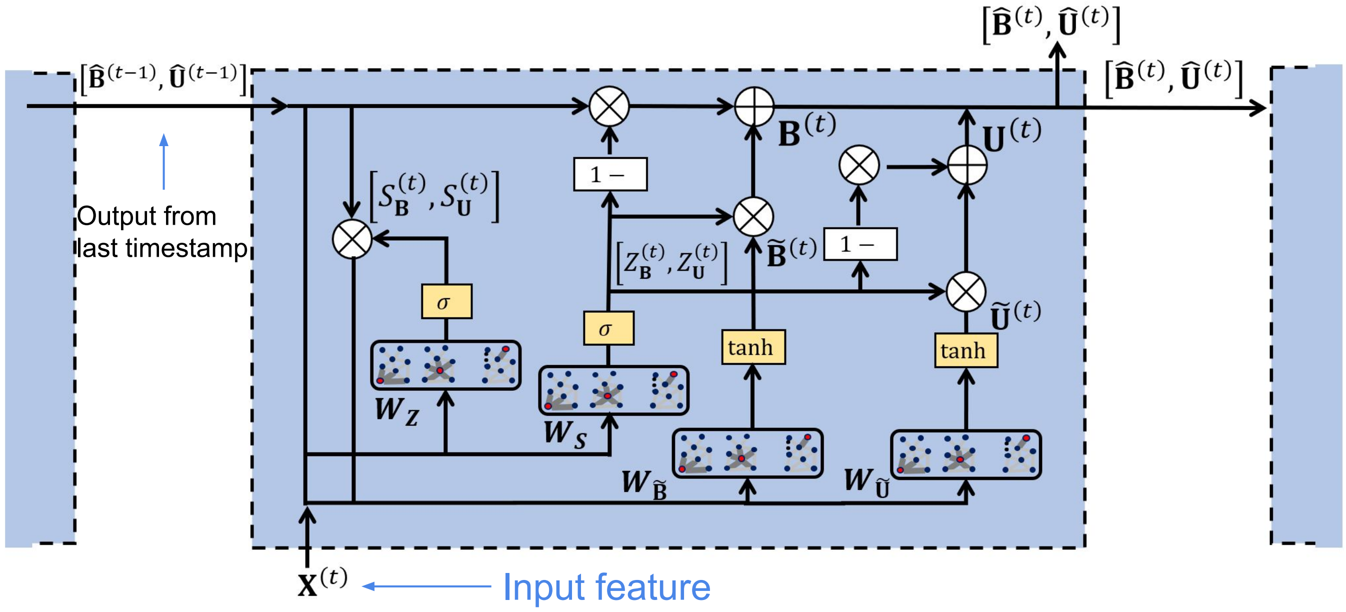
Proposed Approach: GCN-GRU-based Opinion

Output of GRU cell:

Element-wise product

$$\hat{\mathbf{B}}^{(t)} = \mathbf{Z}_{\mathbf{B}}^{(t)} \odot \hat{\mathbf{B}}^{(t-1)} + (1 - \mathbf{Z}_{\mathbf{B}}^{(t)}) \odot \tilde{\mathbf{B}}^{(t)},$$

$$\hat{\mathbf{U}}^{(t)} = \mathbf{Z}_{\mathbf{U}}^{(t)} \odot \hat{\mathbf{U}}^{(t-1)} + (1 - \mathbf{Z}_{\mathbf{U}}^{(t)}) \odot \tilde{\mathbf{U}}^{(t)},$$



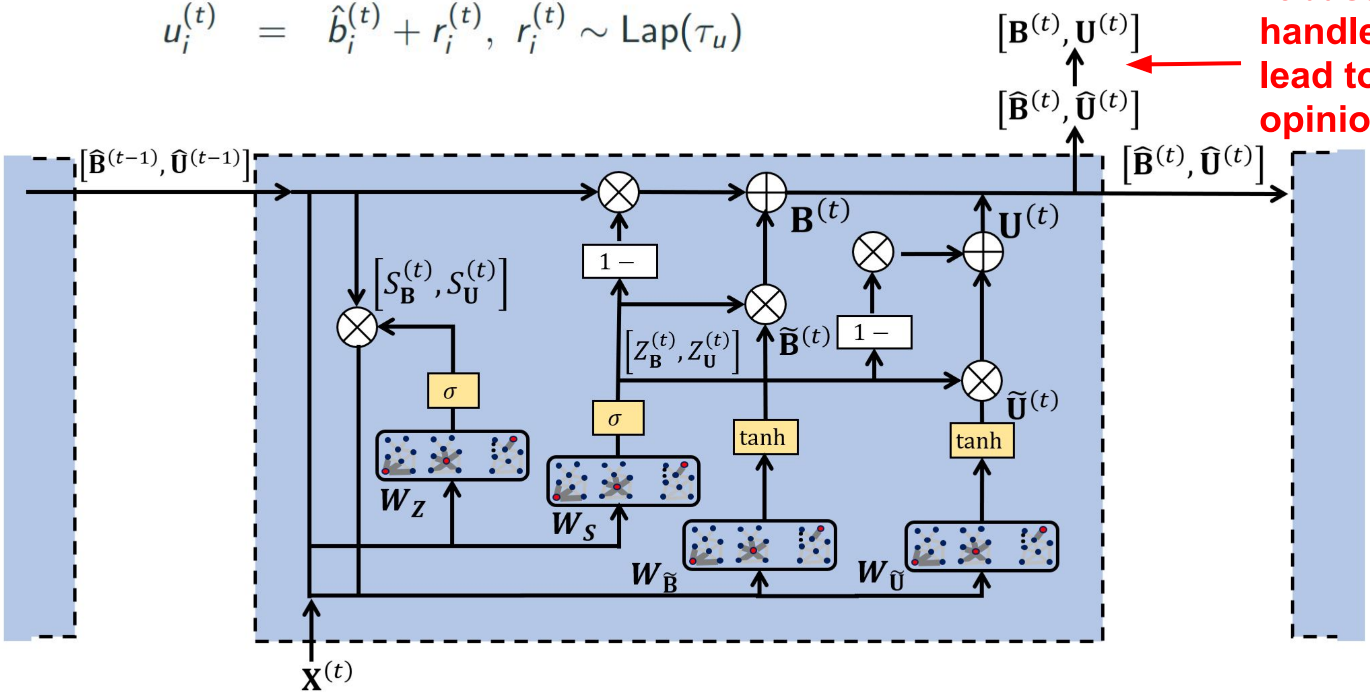
Proposed Approach: GCN-GRU-based Opinion

- The observed beliefs and uncertainties ($\mathbf{B}_{\mathbb{L}_t}^{(t)}$ and $\mathbf{U}_{\mathbb{L}_t}^{(t)}$) are considered as a noise version of their true values ($\hat{\mathbf{B}}_{\mathbb{L}_t}^{(t)}$ and $\hat{\mathbf{U}}_{\mathbb{L}_t}^{(t)}$):

$$
\begin{aligned}
b_i^{(t)} &= \hat{b}_i^{(t)} + e_i^{(t)}, \;\; e_i^{(t)} \sim \mathrm{Lap}(\tau_b) \\
u_i^{(t)} &= \hat{b}_i^{(t)} + r_i^{(t)}, \;\; r_i^{(t)} \sim \mathrm{Lap}(\tau_u)
\end{aligned}
$$

Laplacian distribution

robust statistics to handle anomalies that lead to conflicting opinions

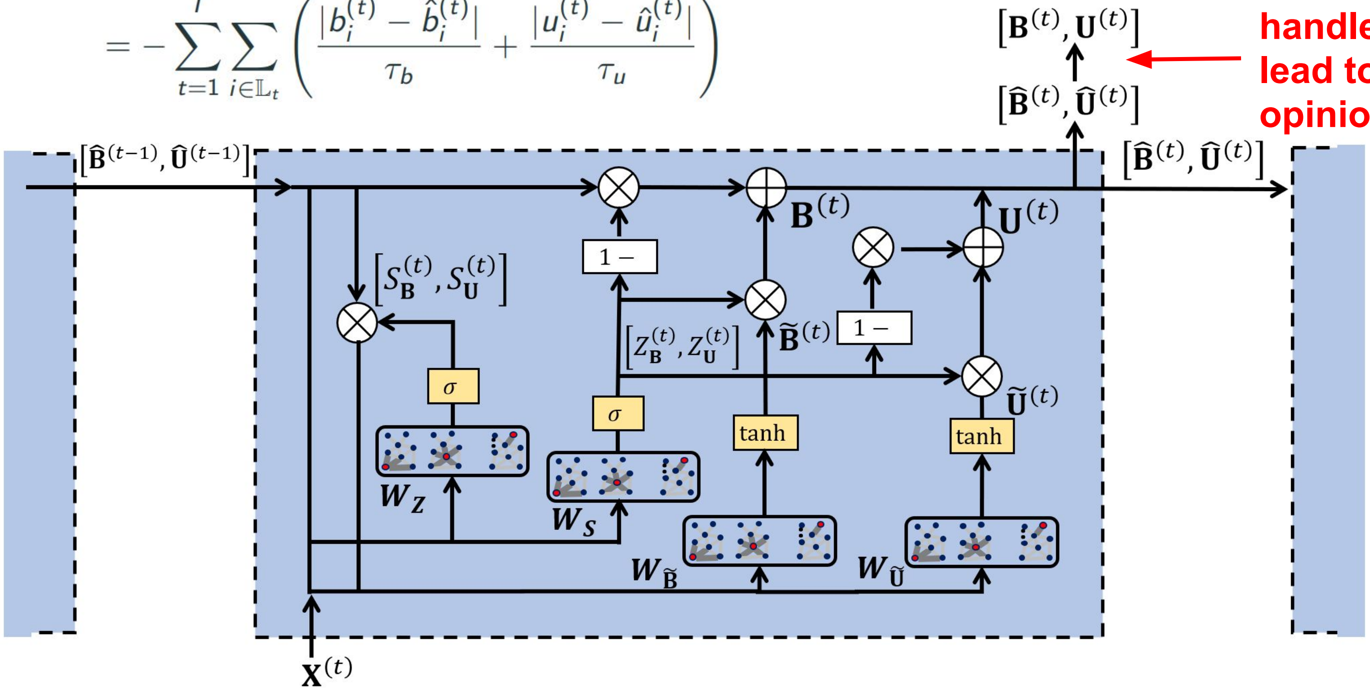


Proposed Approach: GCN-GRU-based Opinion

- Maximize the log probability function

$$\mathcal{L}(\Theta, \{\hat{\mathbf{B}}_{\mathcal{L}_t}, \hat{\mathbf{U}}_{\mathcal{L}_t}\}_{t=1}^{T}) = \log \prod_{t=1}^{T} \prod_{i \in \mathbb{L}_t} \left(\text{Prob}(b_i^{(t)} | \hat{b}_i^{(t)}; \tau_b) \text{Prob}(u_i^{(t)} | \hat{u}_i^{(t)}; \tau_u) \right)$$

$$= -\sum_{t=1}^{T} \sum_{i \in \mathbb{L}_t} \left(\frac{|b_i^{(t)} - \hat{b}_i^{(t)}|}{\tau_b} + \frac{|u_i^{(t)} - \hat{u}_i^{(t)}|}{\tau_u} \right)$$



robust statistics to handle anomalies that lead to conflicting opinions

Datasets & Experimental Setting

Trustiness prediction

Congestion prediction

Spammer prediction

- Road traffic datasets:

Dataset	# nodes	# edges	# weeks	# snapshots in total
Epinions	477,468	8,477,468	-	-
D.C.	1,522	5,028	43	3440
Philadelphia	607	1,772	43	3440
Spammer	165,410	2,441,388	-	-

- Parameter settings:
  - Time sequence: 10
  - Uncertainty ranges: [5%, 100%]
  - Test Ratio: $TR \in \{10\%, 20\%, 30\%, 40\%, 50\%, 60\%\}$
  - Conflict Ratio: $CR \in \{0\%, 5\%, 10\%, 15\%, 20\%\}$
- Performance metrics:

$$\text{B-MSE}(\omega_{\mathbb{V}\setminus\mathbb{L}}) = \frac{1}{N}\sum\nolimits_{i\in\mathbb{V}\setminus\mathbb{L}} |b_i - b_i^\star|$$

$$\text{U-MSE}(\omega_{\mathbb{V}\setminus\mathbb{L}}) = \frac{1}{N}\sum\nolimits_{i\in\mathbb{V}\setminus\mathbb{L}} |u_i - u_i^\star|$$

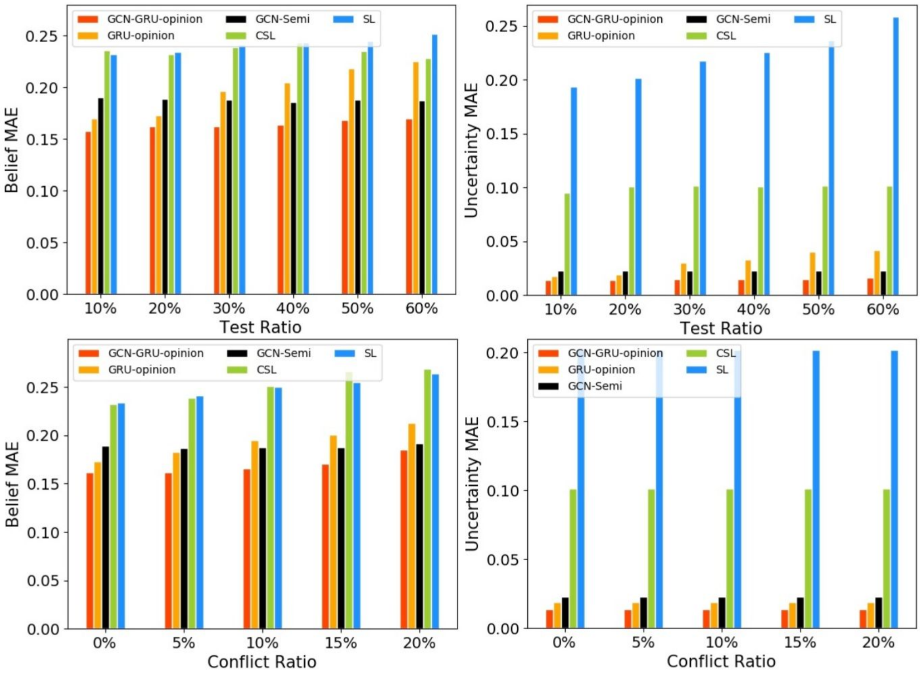
- Computation time metric: seconds

Comparing Schemes

Comparison Methods:

- The proposed: **GCN-GRU-opinion**
- GRU-opinion: RNN model
- GCN-Semi: Semi-supervised node classification
- CSL: combining Probabilistic Soft Logic (PSL) and Markov Random Fields (MRFs) with SL
- SL: Subjective Logic inference based on discount and consensus operators

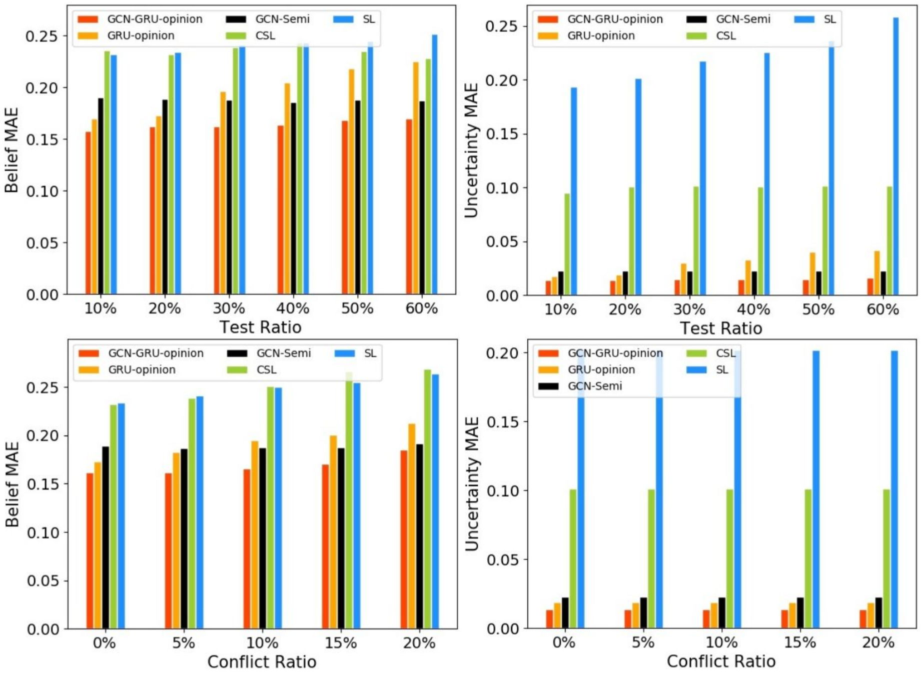
Results with Epinions Dataset



Effect of Test Ratio:

- ❖ Belief-MSE: **GCN-GRU** > GRU > GCN-Semi > SL > CSL
- ❖ Uncertainty-MSE: **GCN-GRU** > GRU > GCN-Semi > CSL > SL
- ❖ **Less sensitivity** under different test ratios

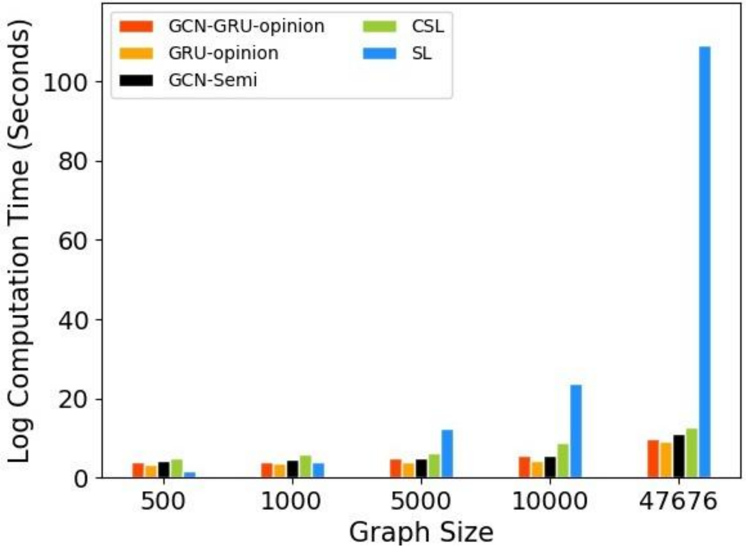
Results with Epinions Dataset



Effect of Conflict Ratio:

- ❖ Belief-MSE: **GCN-GRU** > GRU > GCN-Semi > CSL > SL
- ❖ Uncertainty-MSE: **GCN-GRU** > GRU > GCN-Semi > CSL > SL
- ❖ **Less sensitivity** under different conflict ratios

Results with Epinions Dataset



- Computation order: **GRU** > GCN-GRU > GCN-Semi > CSL > SL
- Complexity of SL increases **exponentially** while that of others (GCN-GRU, GRU, CSL) increases **linearly**

Conclusion

- GCN-GRU method **outperforms** among all in both B-MSE and U-MSE.
- GCN-GRU method shows **less sensitivity** over a wide range of test ratios and conflict ratios.
- GCN-GRU outperforms others because it integrates both **topological** and **temporal** heterogeneous dependency information.
- GCN-GRU scales almost **linearly** in proportion to the network size and is scalable for large-scale network

Thank You!

Question?

Reach Xujiang Zhao at
xzhao8@albany.edu

UAB 401, 1215 Western Ave, Albany, NY, USA
University at Albany - SUNY