End-to-End Risk-aware Reinforcement Learning to Detect Asymptomatic Cases in Healthcare Facilities

Yongjian Zhong, Weiyu Huang, Bijaya Adhikari

University of Iowa

ICHI 2024

Jun 3, 2024

University of Iowa
comp|epi
computational epidemiology research

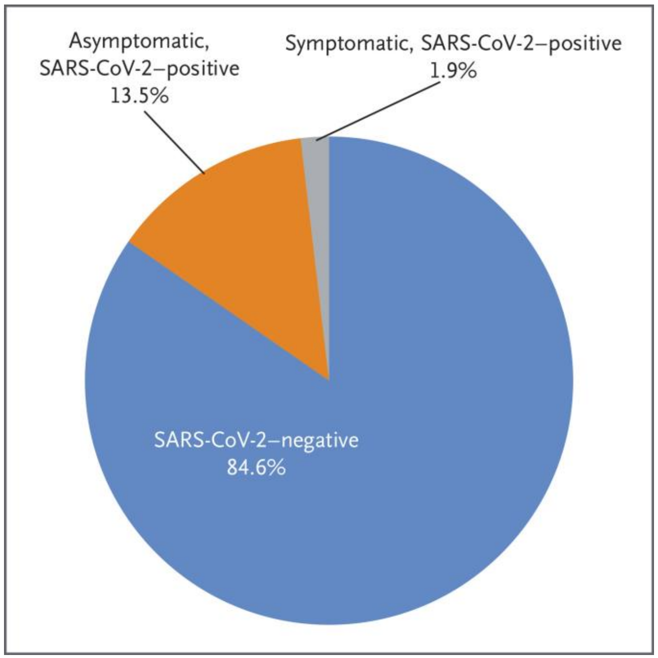
The University of Iowa

Outline

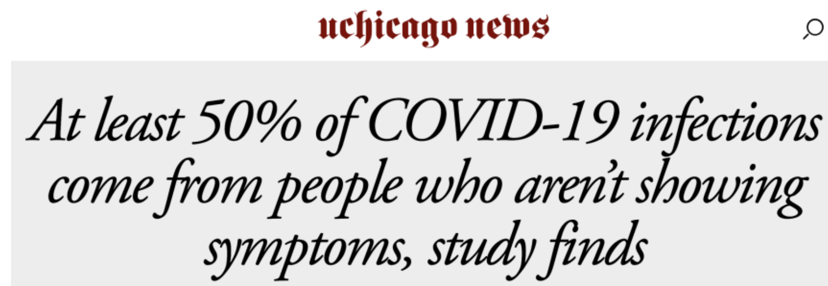
- Background & Challenges
- Problem Formulation
- Our Method
- Experiment
- Conclusion & Future Work

THE UNIVERSITY OF IOWA

Asymptomatic cases drives outbreak

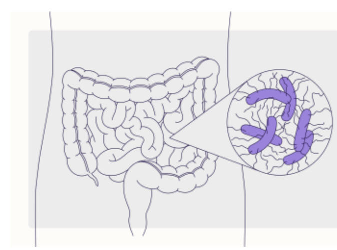


[1] **Asymptomatic cases are common**



[2]

They contribute to the outbreak



[3]

C. Difficile is another example

Detecting asymptomatic cases is crucial in combating pandemic outbreak!

The University of Iowa

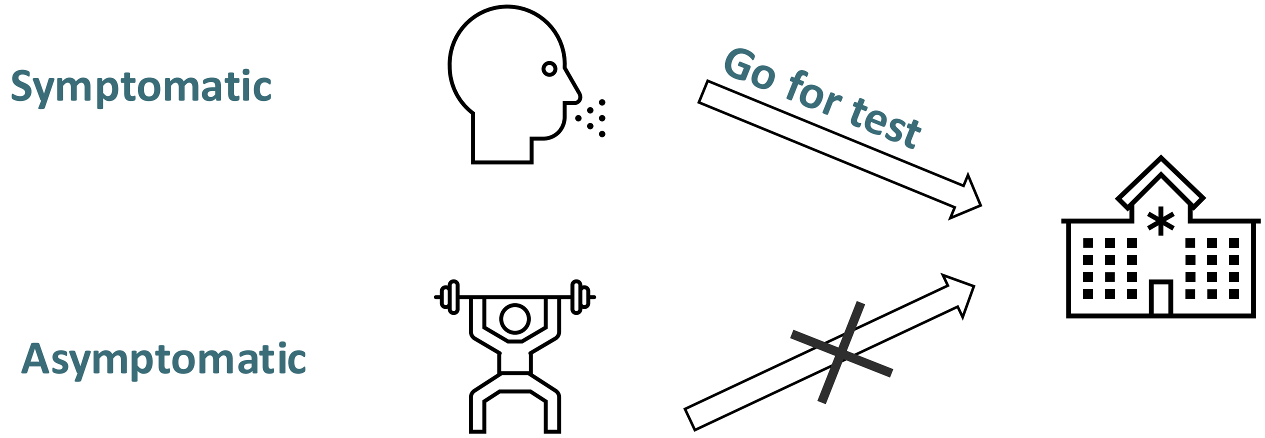
[2] Subramanian, Rahul, Qixin He, and Mercedes Pascual. "Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity." *Proceedings of the National Academy of Sciences* 118.9 (2021): e2019716118.

[3] Ziakas, Panayiotis D., et al. "Asymptomatic carriers of toxigenic C. difficile in long-term care facilities: a meta-analysis of prevalence and risk factors." PloS one 10.2 (2015): e0117195.

Challenges in Detecting Asymptomatic Cases

- Challenge 1: Data Scarcity

Most data don't include symptomatic information

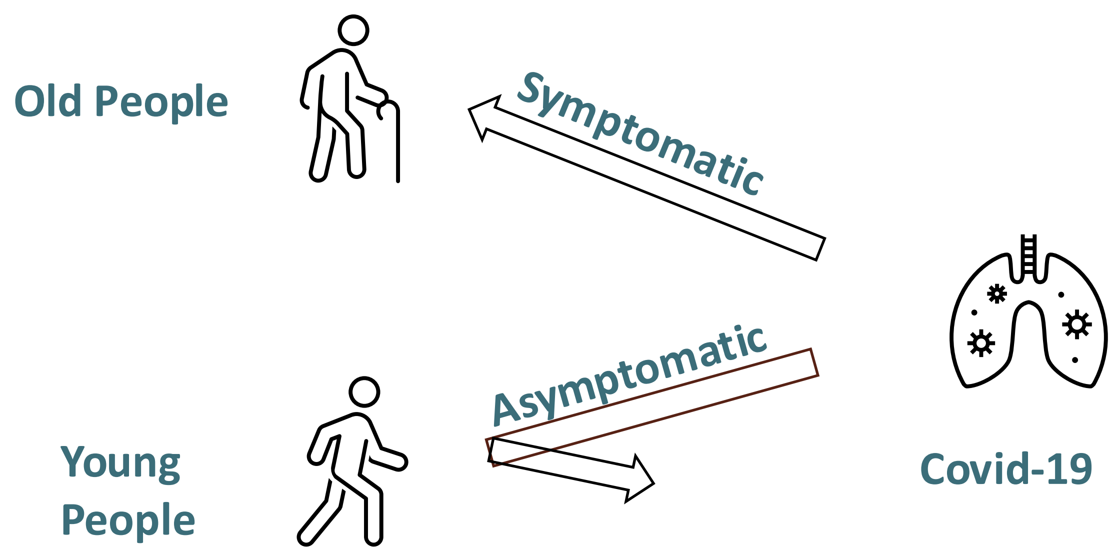


The University of Iowa

Challenges in Detecting Asymptomatic Cases

- ## Challenge 2: Bias on Risk Factors

The *risk factors* for symptomatic infections *differ* from those of asymptomatic infections.



The University of Iowa

Challenges in Detecting Asymptomatic Cases

- **Challenge 3: Systematic Bias**

  Severe cases get more attention **when it comes to testing when** ***capacity is limited*****.**

Scope of This Paper

- **Goal**

  Given *interactions* between people and some *positive* cases, *infer* the *asymptomatic* cases.

- **But …**

  Such *interactions do not exist* for most scenarios.

- **Health Care Facilities**

  Well-documented

THE UNIVERSITY OF IOWA

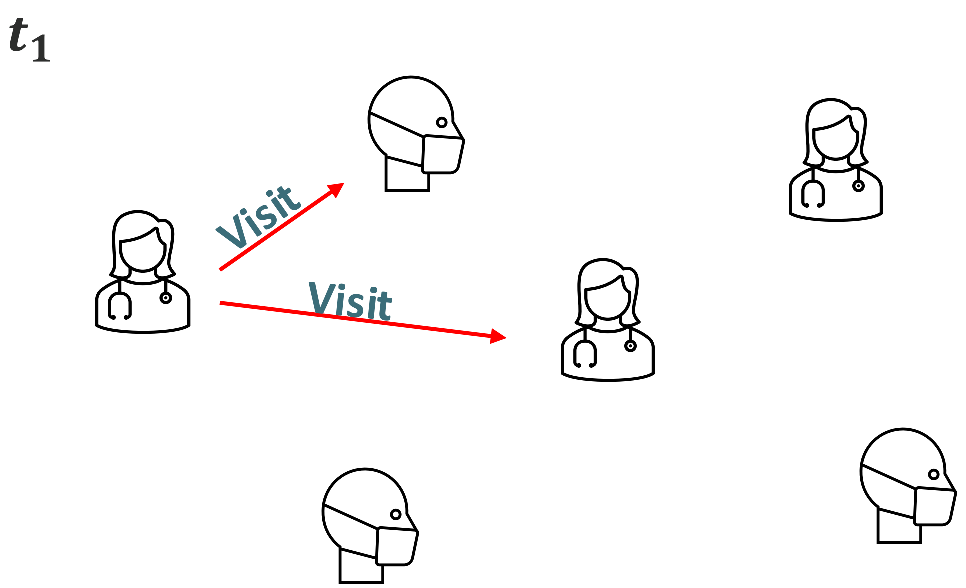
Outline

- Background & Challenges
- **Problem Formulation**
- Our Method
- Experiment
- Conclusion & Future Work

Construct Graph from Interactions

- Health care facilities

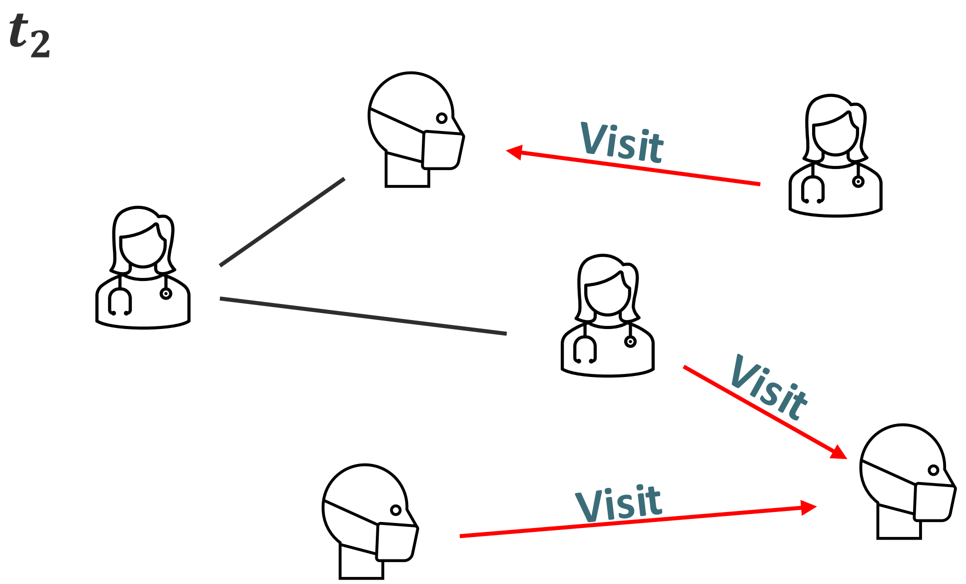
Interactions between *patients* and *HCPs*.



Construct Graph from Interactions

- Health care facilities

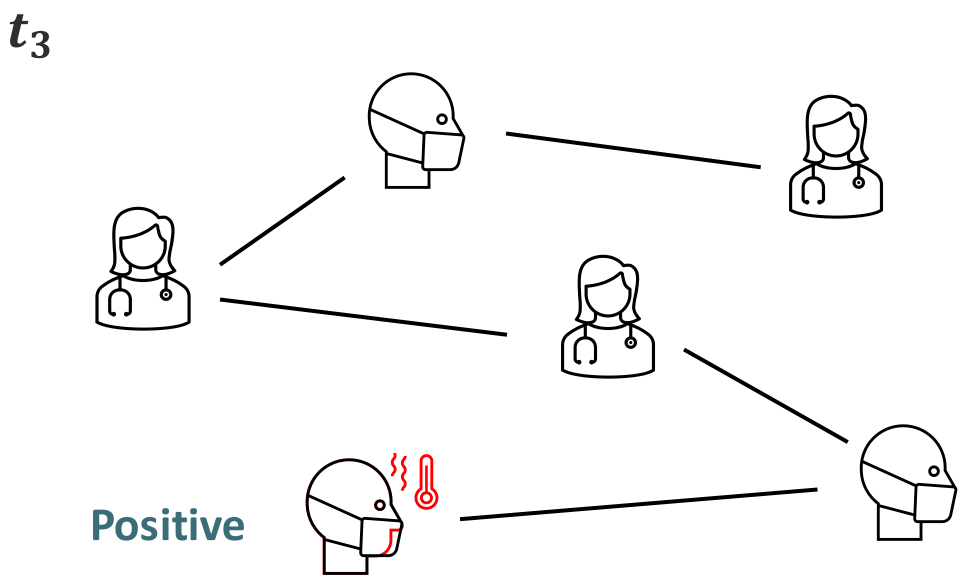
Interactions between *patients* and *HCPs*.



Construct Graph from Interactions

- Health care facilities

Interactions between *patients* and *HCPs*.



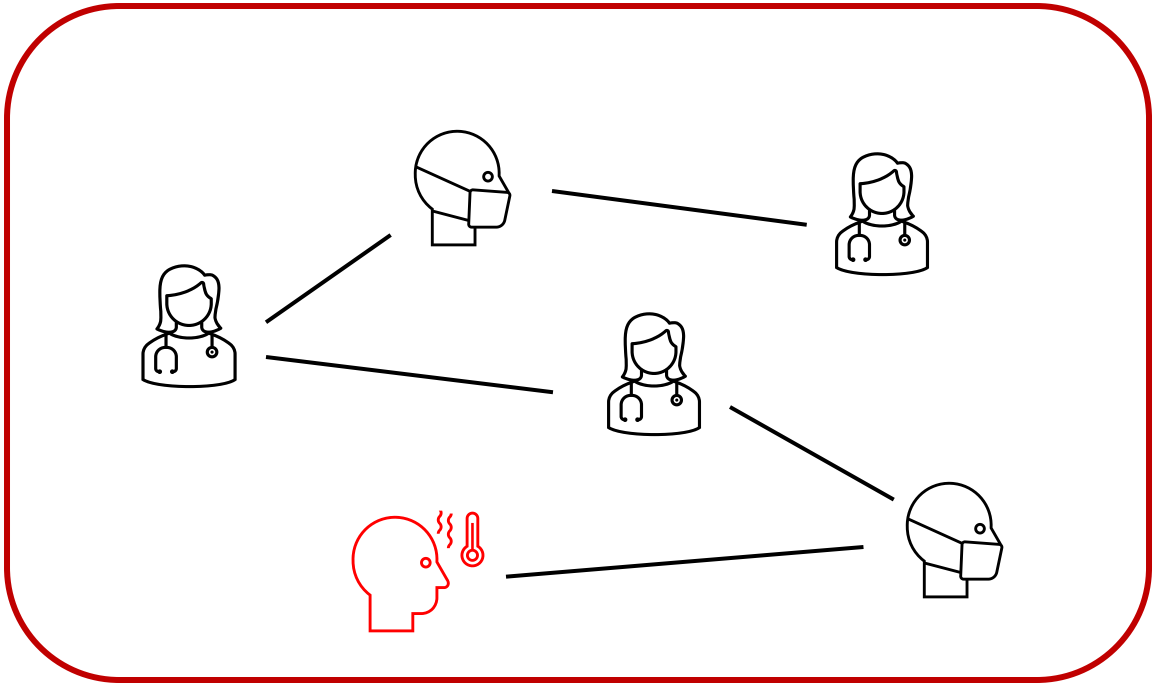
THE UNIVERSITY OF IOWA

Construct Graph from Interactions

- **Health care facilities**

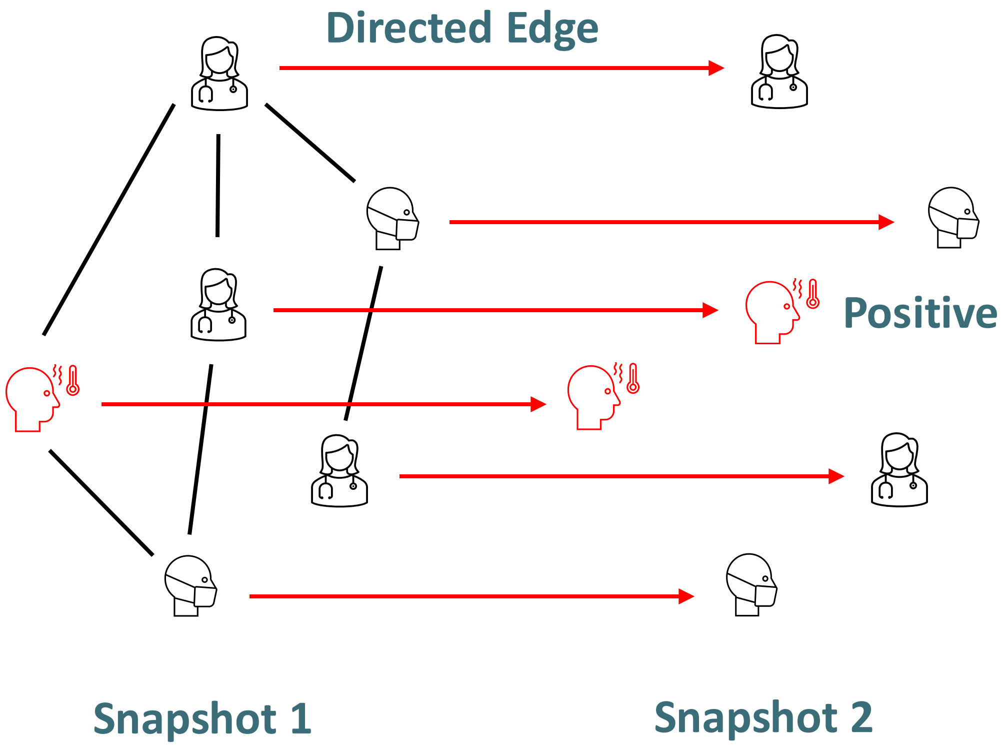
Interactions between *patients* and *HCPs*.

A Snapshot



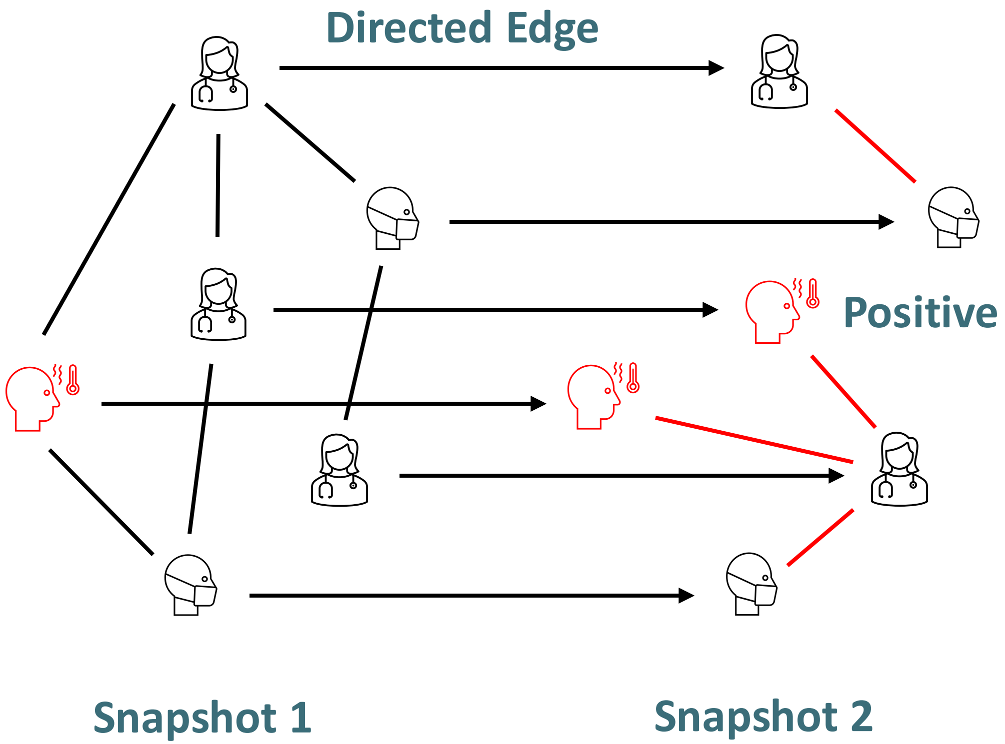
If we have more interactions …

- Build snapshot upon previous ones



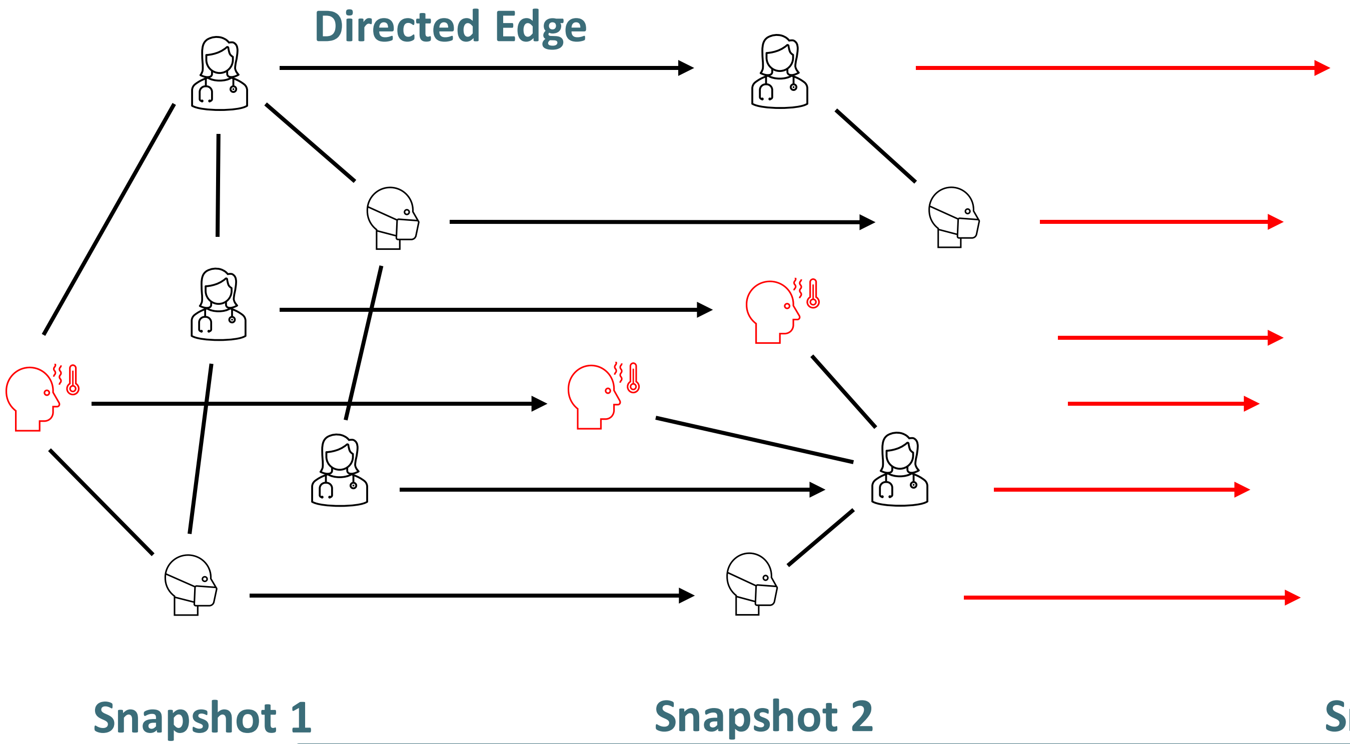
If we have more interactions ...

- Build snapshot upon previous ones



Still have more interactions ...

- Build snapshot upon previous ones

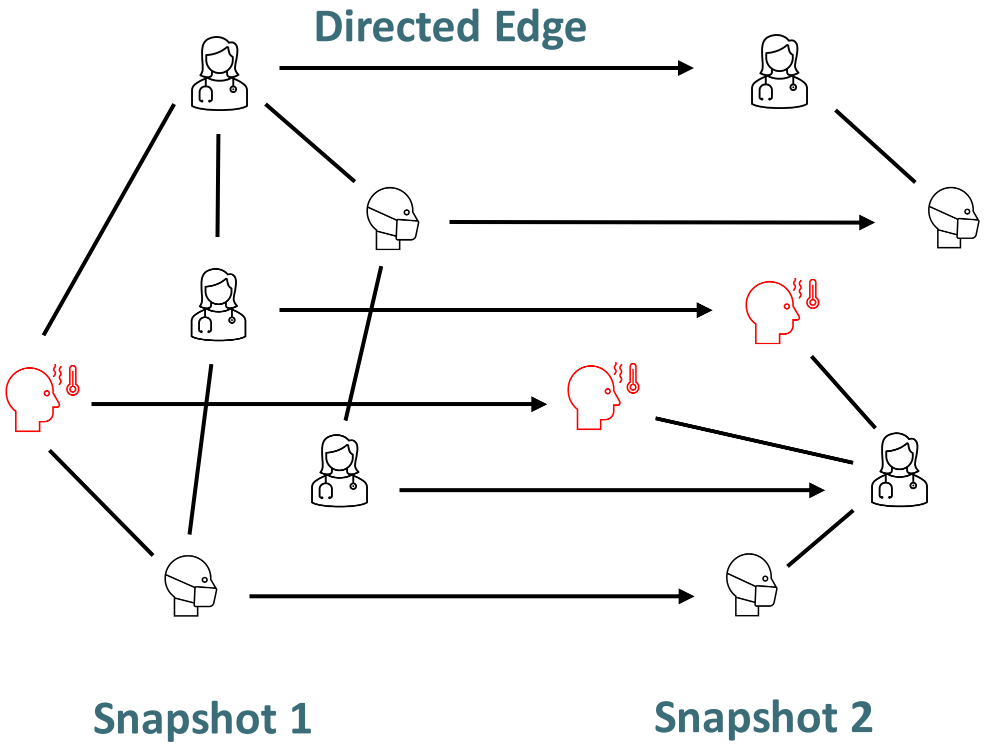


Such network also called *time-expanded network*

The University of Iowa

Problem Formulation

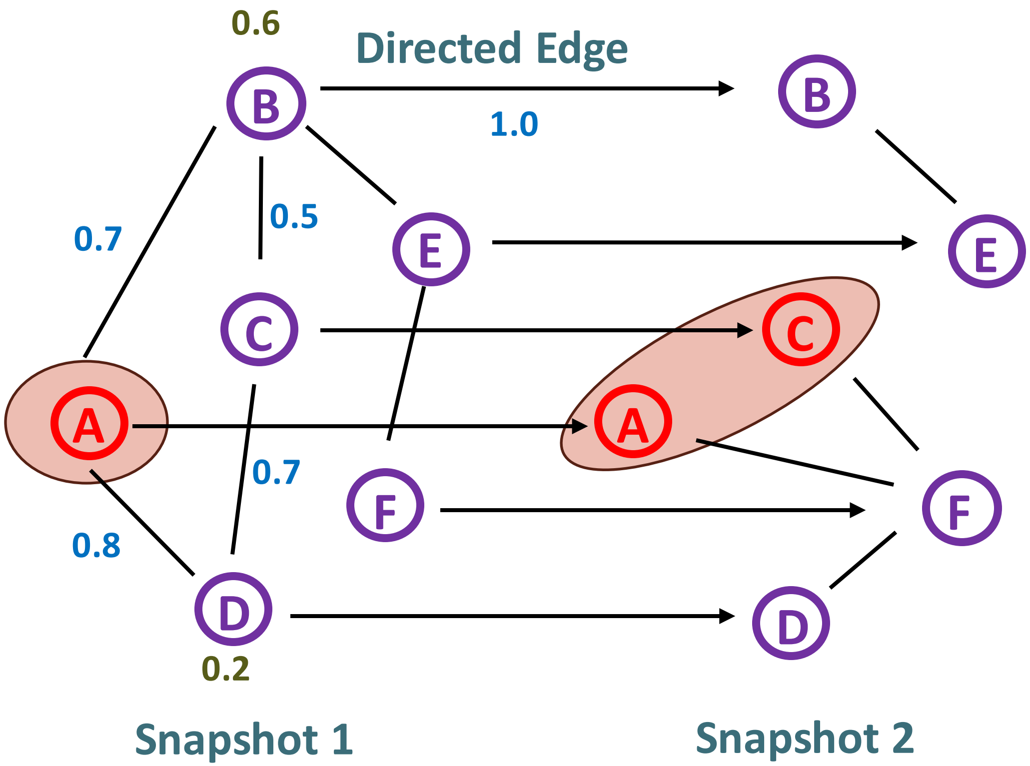
- Take *2 snapshots* as an example. We are focusing the following problem.



Simplify

Problem Formulation

- Take *2 snapshots* as an example. We are focusing the following problem.



Edge Weight: Probability of transmission

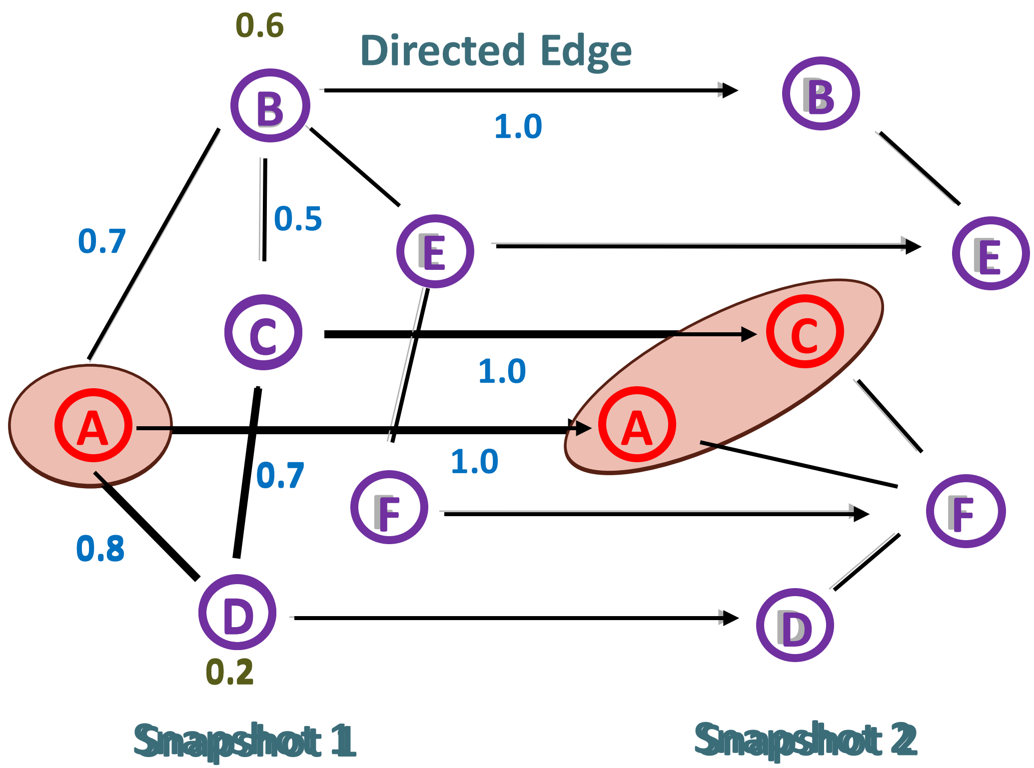
Node Prize: Probability of being asymptomatic

Goal: Cover positive nodes by a *tree*, and *maximize* the weight and prize

The University of Iowa

One Solution

Obj = 0.7 + 0.5 + 1.0 + 1.0 + 0.2 = 3.4



Edge Weight: Probability of transmission

Node Prize: Probability of being asymptomatic

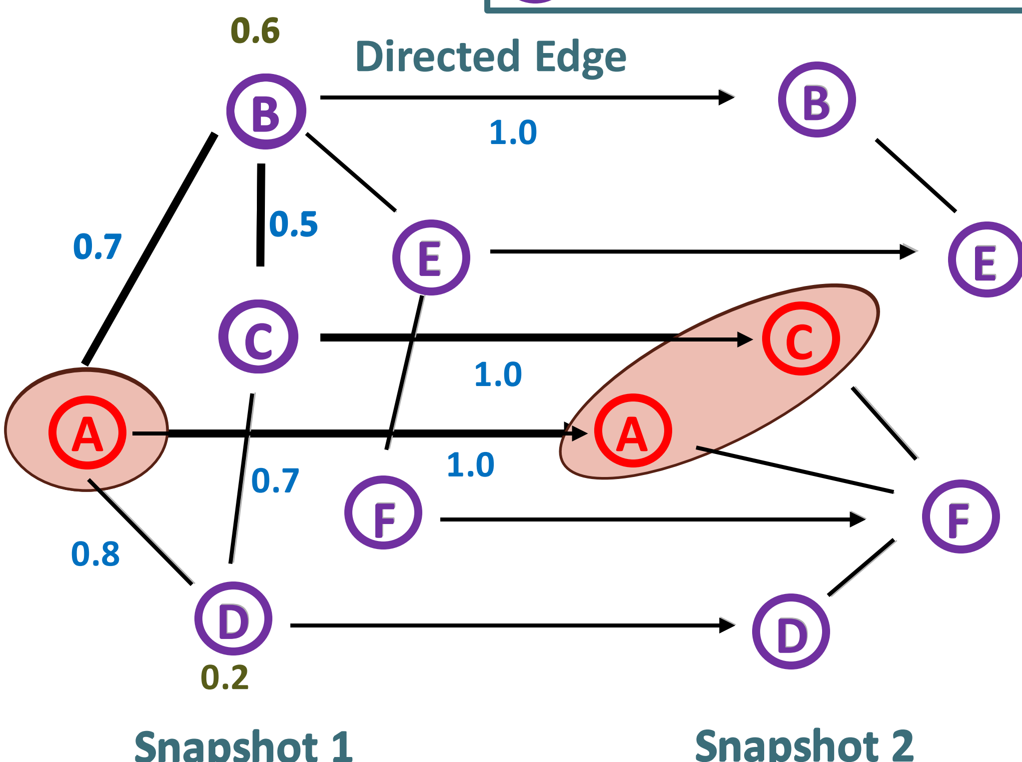
Goal: Cover positive nodes by a *tree*, and *maximize* the weight and prize

The University of Iowa

Another Solution

Obj = 0.7 + 0.5 + 1.0 + 1.0 + 0.6 = 3.8 > 3.4

B should be asymptomatic!



Edge Weight: Probability of transmission

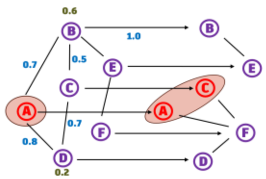
Node Prize: Probability of being asymptomatic

Goal: Cover positive nodes by a *tree*, and *maximize* the weight and prize

This problem is Prize-collective Steiner Tree.

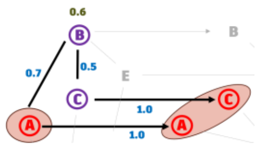
The University of Iowa

Directed Prize-collecting Steiner Tree



☐ Given

- Time-expanded Network
- Positive Cases



☐ Output

- Steiner Tree, and satisfies

$$T^* = \arg\min_{T} \sum_{(v_a,v_b)\in T} W_s(v_a, v_b) + \alpha \sum_{v_c \in V_s \setminus T} f(F_{v_c})$$

1 - Edge Weight

Node Prize of excluded nodes

Challenge 1: Estimate the Probability

Challenge 2: Construct the Tree

The University of Iowa

Outline

- Background & Challenges
- Problem Formulation
- **Our Method**
- Experiment
- Conclusion & Future Work

The University of Iowa

Related Works

- MCA [Jang et al.]
  - Prize-collecting Steiner Tree
  - Fixed prize, sub-optimal
- CuLT [Rozenshtein et al.]
  - Steiner Tree
  - Assume SI model, ignore risk-factors
- TopoLSTM [Wang et al.]
  - Cascade
  - Ignore risk-factors

Methods	Cascade	Steiner Tree	Risk Factors	End-to-end
TopoLSTM	✓	✗	✗	✗
CuLT	✓	✓	✗	✗
MCA	✓	✓	✓	✗
Ours	✓	✓	✓	✓

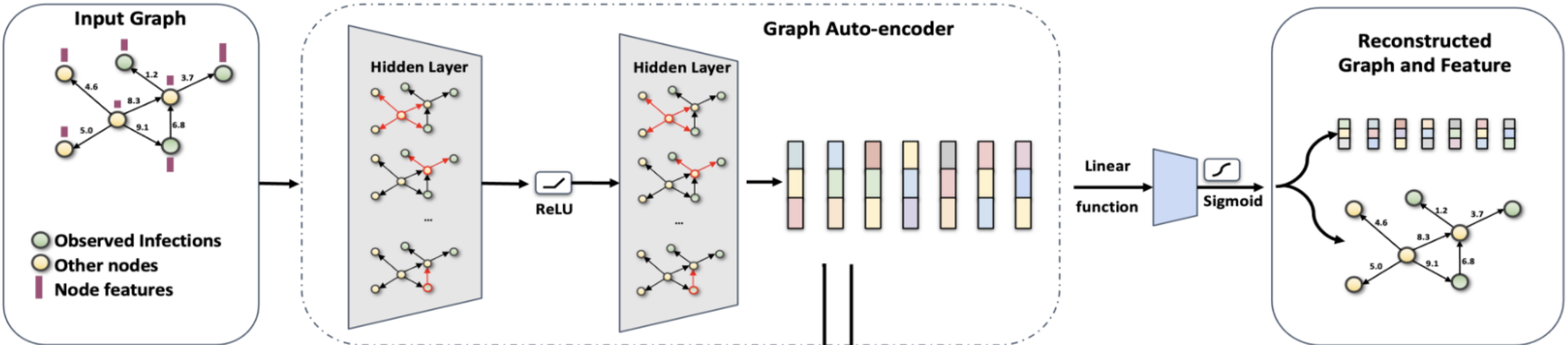
Main contribution: jointly optimize for estimating and constructing.

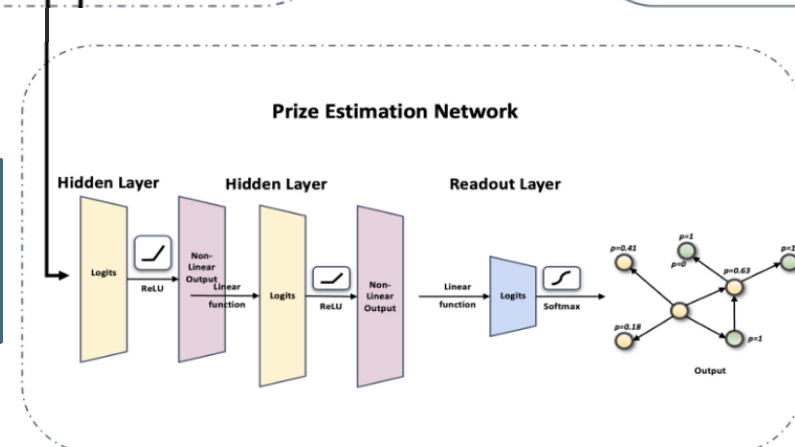
THE UNIVERSITY OF IOWA

[1] Jang, Hankyu, et al. "Risk-aware temporal cascade reconstruction to detect asymptomatic cases: For the cdc mind healthcare network." 2021 IEEE International Conference on Data Mining (ICDM). IEEE, 2021.
[2] Rozenshtein, Polina, et al. "Reconstructing an epidemic over time." Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016.
[3] Wang, Jia, et al. "Topological recurrent neural network for diffusion prediction." 2017 IEEE international conference ondata mining (ICDM). IEEE, 2017..

Estimate Probability

☐ Graph Autoencoder





- The *prize* is used in constructing tree
- Latent embeddings receive *gradient*

Construct the Tree – Reinforcement Learning

State:

- The state space is *all the possible Trees*. The starting state is $\{r\}$ for some random node.

Action:

- The action is to *select edge* (u, v) and $u \in S_t, v \notin S_t$. Then, the state will transit to $S_t \cup \{v\}$ with probability 1. We include one more node for each step.

Reward:

$$\alpha f(F_v) - \sum_{u \in S_t} W_s(u, v)$$

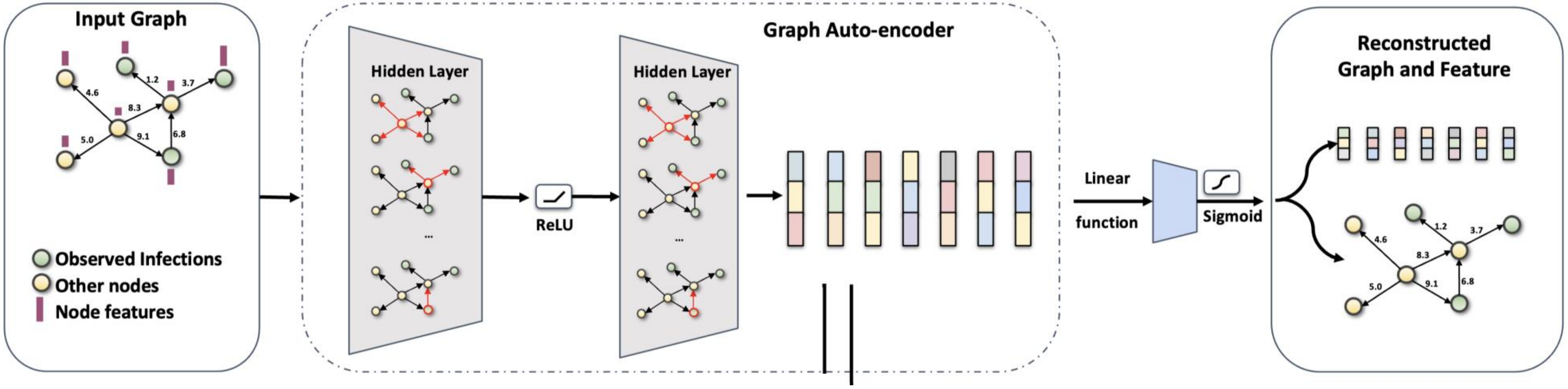
Sum up all steps

$$\alpha \sum_{v \in \mathcal{S}_T} f(v) - \sum_{(u,v) \in \mathcal{S}_T} W(u, v)$$

Exactly the objective!

The University of Iowa

Jointly Optimization



Content

- Background & Challenges
- Prize-collecting Steiner Tree
- Our Method
- **Experiment**
- Conclusion & Future Work

Experiment

Data

- University of Iowa Hospitals and Clinics **(UIHC)**
- Interactions between **patients** and **healthcare workers**
- 500 **(UIHC1**), 2000 **(UIHC2),** and 5000**(UIHC3)**

Tasks

- Simulated CDI Outbreak
- Real CDI Outbreak
- Simulated Covid-19 Outbreak

No asymptomatic ground truth

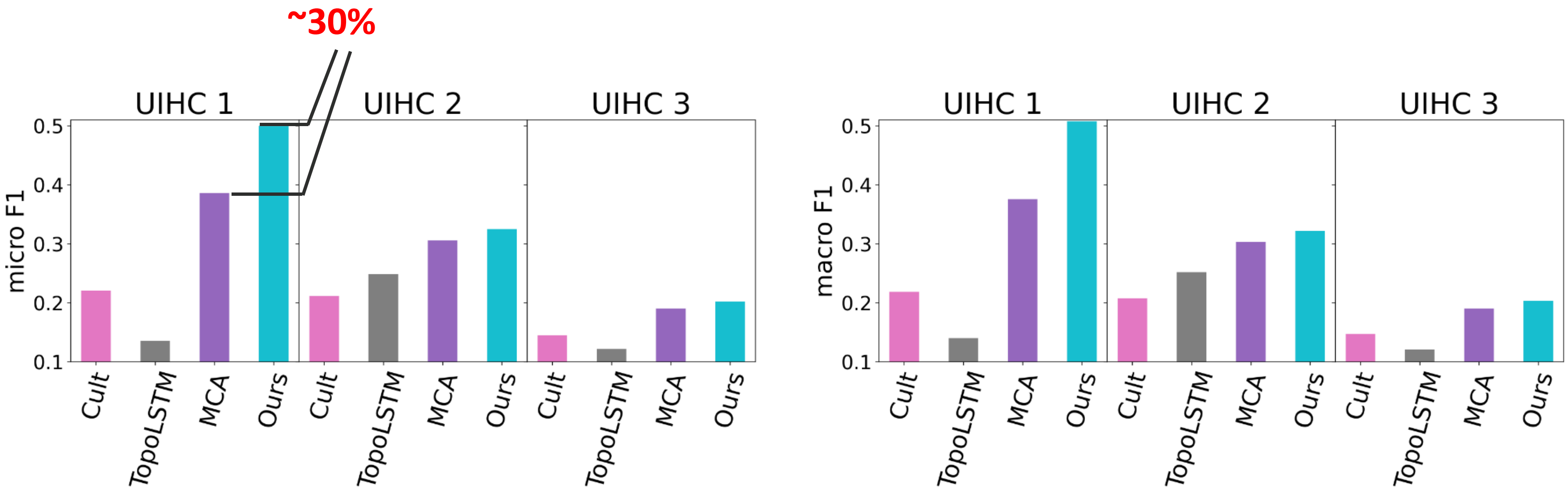
The University of Iowa

Simulated CDI Outbreak - Setup

- **Use** ***Biased-SIS*** model to generate ***symptomatic*** and ***asymptomatic*** infections.
- The model used known risk factors for C. Difficle
- **On 500** ***(UIHC1)*****, 2000** ***(UIHC2)*****,** and **5000** ***(UIHC3)***
- **Based on** ***symptomatic cases, infer the asymptomatic***
- **Use micro-F1 and macro-F1 to evaluate**
- **Run** ***5 time*** **and report the mean.**

The University of Iowa

Simulated CDI Outbreak - Result



Joint optimization enables us to avoid the limitations of MCA and achieve greater accuracy.

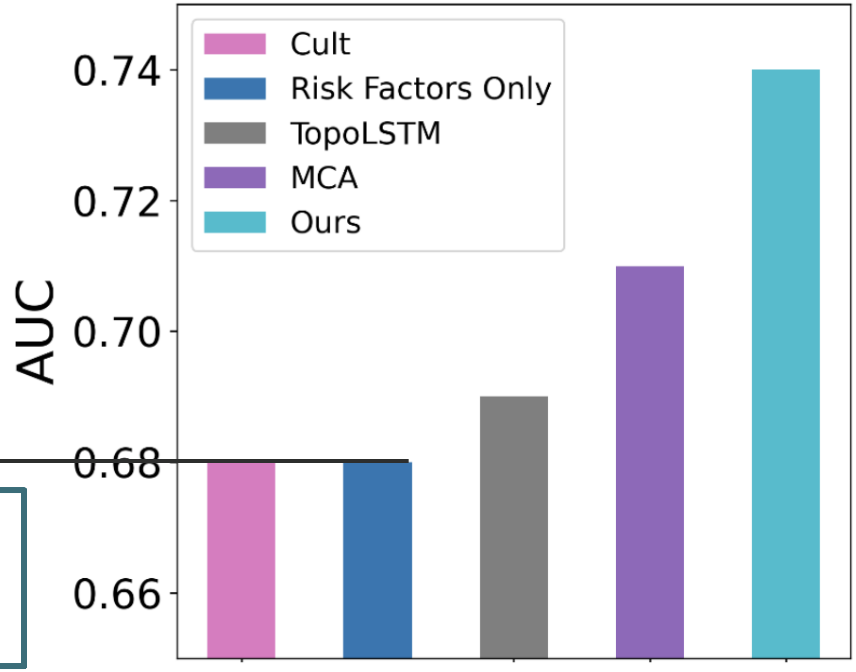
Real CDI Outbreak - Setup

- Sample ***one month of interactions*** from UIHC to construct graph.
- **68** positive cases.
- **80%** training
- Each method infers the asymptomatic cases based on training set
- Based on the prediction, we compute the *asymptomatic pressure.* (i.e., a normalized metric computing interaction frequency with asymptomatic cases)
- Use the *asymptomatic pressure* as extra features, we train a MLP to predict the rest positive cases.

The University of Iowa

Real CDI Outbreak

Results



CuLT is unable to improve upon the Risk-factors-only baseline

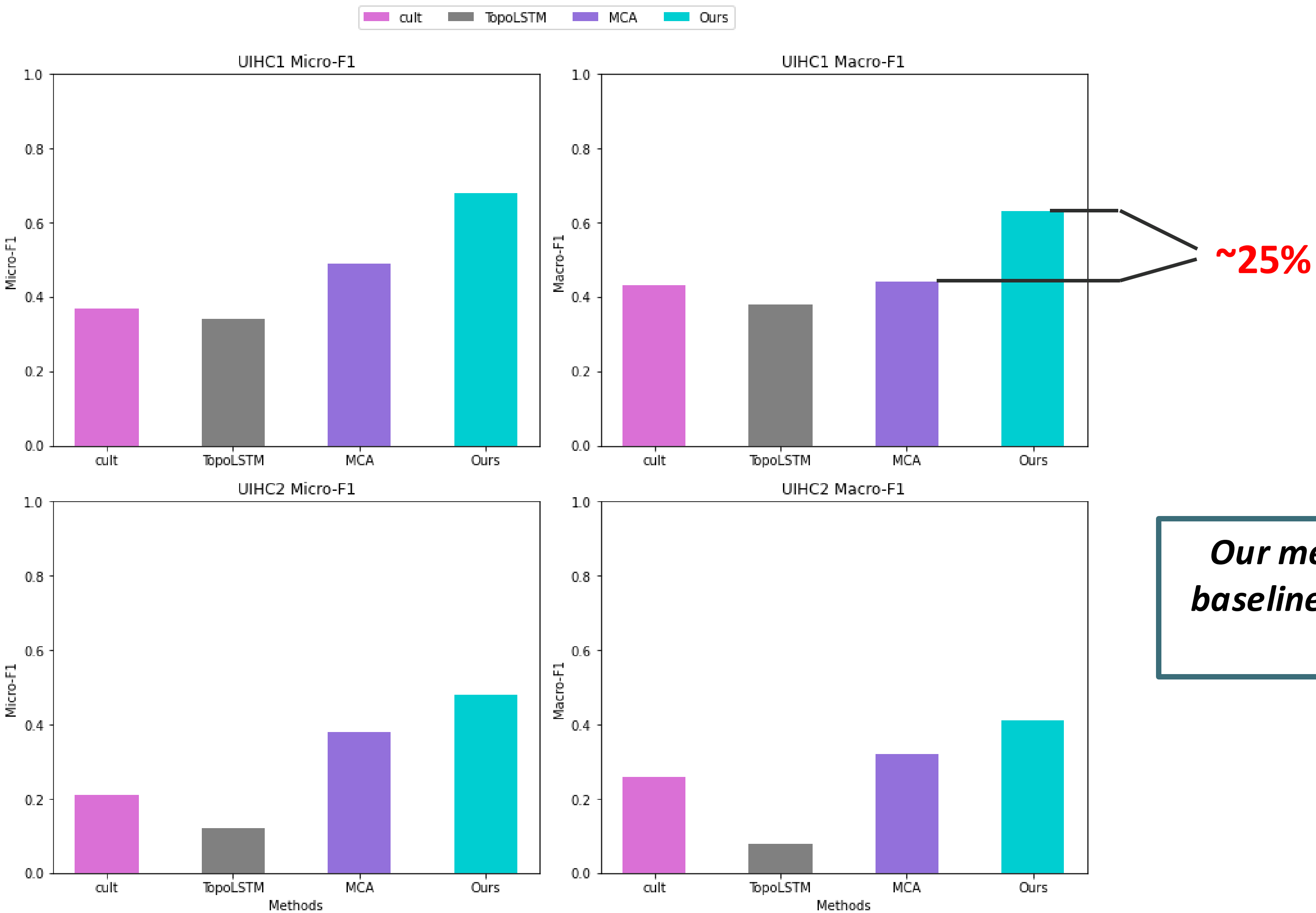
Inferred asymptomatic cases by our method are more accurate

Simulated Covid-19 Outbreak - Setup

- ***CovaSim*** model
- Generate **symptomatic** and **asymptomatic** infections.
- **On 500 *(UIHC1)*, 2000 *(UIHC2)***
- Based on **symptomatic, models infer the asymptomatic cases.**
- **Use micro-F1 and macro-F1 to evaluate**

THE UNIVERSITY OF IOWA

Simulated Covid-19 Outbreak - Result



Our method outperform baselines even for another disease

THE UNIVERSITY OF IOWA

Outline

- Background & Challenges
- Problem Formulation
- Our Method
- Experiment
- **Conclusion & Future Work**

THE UNIVERSITY OF IOWA

Conclusion & Future Work

- Our results on synthetic outbreaks show that the proposed approach is able to identify asymptomatic infection with high accuracy while the baseline approaches are less accurate.

- The tree affect the probability.

$$T^* = \arg\min_T \sum_{(v_a, v_b) \in T} W_s(v_a, v_b) + \alpha \sum_{v_c \in V_s \setminus T} f(F_{v_c})$$

$$f(F_{v_c}; T)$$

THE UNIVERSITY OF IOWA

THANKS