

Conformal Prediction for Multi-Source Detection on a Network

Xingchao Jian^{1,2}, Purui Zhang³, Lan Tian³, Feng Ji³, Wenfei Liang³, Wee Peng Tay³, Bihan Wen³, Felix Krahrmer^{1,2}

¹School of Computation, Information and Technology, Technical University of Munich

²Munich Center for Machine Learning (MCML)

³School of Electrical and Electronic Engineering, Nanyang Technological University

xingchao.jian@tum.de, purui001@e.ntu.edu.sg, lant0003@e.ntu.edu.sg, jifeng@ntu.edu.sg, wenfei001@e.ntu.edu.sg, wptay@ntu.edu.sg, bihan.wen@ntu.edu.sg, felix.krahrmer@tum.de

Abstract

Detecting the origin of information or infection spread in networks is a fundamental challenge with applications in misinformation tracking, epidemiology, and beyond. We study the multi-source detection problem: given snapshot observations of node infection status on a graph, estimate the set of source nodes that initiated the propagation. Existing methods either lack statistical guarantees or are limited to specific diffusion models and assumptions. We propose a novel conformal prediction framework that provides statistically valid recall guarantees for source set detection, independent of the underlying diffusion process or data distribution. Our approach introduces principled score functions to quantify the alignment between predicted probabilities and true sources, and leverages a calibration set to construct prediction sets with user-specified recall and coverage levels. The method is applicable to both single- and multi-source scenarios, supports general network diffusion dynamics, and is computationally efficient for large graphs. Empirical results demonstrate that our method achieves rigorous coverage with competitive accuracy, outperforming existing baselines in both reliability and scalability.

Code — <https://github.com/xcjian/Conformalized-Network-Source-Detection>

Extended version — <https://arxiv.org/abs/2511.08867>

1 Introduction

Online social networks have experienced substantial expansion over the past few decades, becoming vital platforms for information dissemination (Bakshy et al. 2011; Mitchell et al. 2013). Their dense connectivity and increasing role in news distribution allow rumors, often started by multiple users, to spread quickly across large network segments. When misinformation causes harm, identifying its original sources is crucial for investigators, akin to tracing patient zero in epidemiological challenges like COVID-19 (Hu et al. 2020) to inform containment strategies. This source detection problem (Shah and Zaman 2011; Luo, Tay, and Leng 2013) is the focus of our study.

The propagation of rumors or diseases on a network can be modeled as the spread of *information*, where any node that

acquires the information is considered *infected*. The process typically begins with a finite set of source nodes, denoted by \mathcal{Y} . Each infected node may transmit the information to its neighbors and, depending on the model, may recover (e.g., in rumor spreading, a recovered node stops propagating the information) and become susceptible again. There are multiple models such as SIS, SIR, SIRS that describe different propagation patterns (Shelke and Attar 2019). Given one or more snapshots of node statuses after the outbreak, the *multi-source detection problem* aims to estimate the original set of source nodes \mathcal{Y} .

The multi-source detection problem is inherently challenging because different source sets can produce identical observed infection patterns, and exhaustively enumerating all possible source sets is computationally infeasible (Shah and Zaman 2011). Early research primarily explored centrality-based methods (Shah and Zaman 2011; Chen, Zhu, and Ying 2014; Zhu and Ying 2014; Wang et al. 2015; Luo, Tay, and Leng 2017; Tang, Ji, and Tay 2018), which estimate the sources by maximizing various graph centrality measures as proxies for the maximum likelihood estimator, relying on *single snapshots* and without assuming available simulated data. However, these approaches offer provable guarantees only for specific graph structures, such as trees (Shah and Zaman 2011; Luo, Tay, and Leng 2013), and require heuristic adaptations for general graphs (Shah and Zaman 2011; Luo, Tay, and Leng 2017; Ji, Tang, and Tay 2019).

More recently, data-driven machine learning techniques, particularly those based on Graph Neural Networks (GNNs), have been applied to multi-source detection by framing it as a node status pattern recognition problem (Shah et al. 2020; Yan, Fang, and He 2023), with some leveraging *multiple snapshots* (Sha, Al Hasan, and Mohler 2021) and requiring knowledge about propagation model and parameter ranges for simulating training data. Despite their empirical success, these methods also lack statistical performance guarantees.

In this work, we address this gap by developing a conformal prediction (CP) framework (Angelopoulos and Bates 2023; Angelopoulos, Barber, and Bates 2025) combined with a backbone GNN for multi-source detection based on multiple snapshots of node status. Unlike previous similar work using the data-driven approach, ours provides rigorous performance guarantees. Specifically, our goal is to construct a prediction set \hat{C} of nodes such that its recall rate for identi-

fyng the true source nodes exceeds $1 - \beta$ with probability at least $1 - \alpha$, where $\alpha, \beta \in [0, 1]$ are user-specified nominal levels. The adjustable β allows users to balance coverage and prediction set size. A smaller β improves coverage, but also increases the size of detected node set, raising the subsequent investigation costs.

To obtain a pre-trained GNN and a calibration dataset for CP, we assume access to a diffusion path dataset. This dataset can be derived from real-world diffusion examples, such as intranet network propagation, or generated through simulations when the propagation models and parameter ranges are known (Sha, Al Hasan, and Mohler 2021). While this additional information is considered expert knowledge and may require external input, it enables more accurate modeling compared to centrality-based methods. For compatibility with the SD-STGCN model (Sha, Al Hasan, and Mohler 2021), we assume multiple snapshots are available as inputs. However, our CP approach is flexible and can accommodate various input formats.

CP is a statistical approach that provides confident prediction sets with guaranteed coverage, using any predictor as backbone (Vovk, Gammerman, and Shafer 2005). It is model-free and distribution-free, ensuring the statistical coverage guarantee regardless of the predictor or data distribution. CP is widely applied in image classification (Bhatnagar et al. 2023; Angelopoulos, Barber, and Bates 2024; Romano, Sesia, and Candès 2020), text classification (Tyagi and Guo 2023), time series prediction (Bhatnagar et al. 2023; Gibbs and Candès 2024; Angelopoulos, Barber, and Bates 2024), and other inference tasks.

In *single-source detection*, traditional CP yields a node set that contains the true source node with statistical guarantee. However, in *multi-source detection* where there is a *set* of source nodes, the traditional CP methods produces a collection of node sets that contains the target node set with statistical guarantees (Maltoudoglou et al. 2022; Tyagi and Guo 2023; Katsios and Papadopoulos 2024). This deviates from the target of finding a node set to *include* a proportion $(1 - \beta)$ of the true source set with statistical guarantee. Existing methods like ArbiTree and PGM (Cauchois, Gupta, and Duchi 2021) addressed the cases when $\beta = 0$, however, they are restricted to the special case where the graph is a tree. ADiT (Dawkins, Li, and Xu 2021) established coverage guarantee for network source detection using a hypothesis-testing approach, applicable only to the single-source case.

Another related line of research studies CP methods for semi-supervised node classification on graphs (Huang et al. 2023; H. Zargarbashi, Antonelli, and Bojchevski 2023; Zargarbashi and Bojchevski 2024). However, these approaches do not address the set inclusion objective required for source detection. First, while multi-source detection can be cast as a binary node classification problem, it is not a semi-supervised setting. Second, these methods provide only marginal (node-wise) coverage guarantees, rather than guarantees on the inclusion of the source set.

To the best of our knowledge, this is the first work to apply CP to the network source detection problem using multiple snapshots and with the goal of predicting a multi-source set while providing *rigorous performance guarantees*. We estab-

lish sufficient conditions for the design of non-conformity scores and the construction of prediction sets to achieve a user-specified nominal recall rate. Our approach is computationally efficient, scaling well with the graph size, and does not rely on restrictive assumptions about the underlying network structure. Importantly, the statistical inclusion guarantee of our method holds under the minimal assumption of exchangeability between the calibration and test data, in line with standard CP methodology.

In summary, the main contributions of this work are as follows:

1. We propose non-conformity score designs for set estimation problems, targeting a nominal recall rate with statistical guarantee. These scores yield efficient (i.e., small) prediction sets.
2. We apply the proposed score to the multi-source detection problem. Compared to existing methods that only deal with single-source detection, this approach works under arbitrary number of sources and propagation models.
3. We numerically validate the advantage of our proposed method on the multi-source detection problem under real-world network propagation.

Notation: We use $Q(\mathcal{U}, a)$ to denote the sample lower a -quantile of a set \mathcal{U} . The power set of a set \mathcal{U} is written as $2^{\mathcal{U}}$. For a positive integer k , we write $[k]$ to denote the set $\{1, 2, \dots, k\}$.

2 Related Work

2.1 CP for Set Estimation

In this subsection, we present a brief review of the literature on CP for set estimation problems. These works address the multi-label classification problem, where the target output is a set of labels, similar to the multi-source detection problem.

In multi-label classification, where the label set contains N elements, any label subset can be represented as an N -dimensional binary vector. This allows the set estimation problem to be reformulated as a point estimation problem. However, since there are 2^N possible binary vectors, directly applying traditional CP becomes computationally infeasible (Maltoudoglou et al. 2022; Tyagi and Guo 2023). To address this, some works organize the labels into a hierarchical tree structure and apply multiple hypothesis testing to control the family-wise error rate (Tyagi and Guo 2023). Additionally, incorporating the covariance structure between labels using the Mahalanobis distance, rather than the standard Euclidean norm, has been shown to improve prediction efficiency (Katsios and Papadopoulos 2024).

In all the aforementioned works that address CP for label set estimation, the focus is on finding a collection of subsets that *contains* the ground truth label subset with probability guarantee. This approach may not always be appropriate. In particular, for set estimation tasks such as network multi-source detection, returning a collection of possible node subsets is often impractical or uninformative. Instead, it is more desirable to provide a single node set that *includes* the true source set \mathcal{Y} , or achieves a specified recall rate with high probability.

Existing methods such as PGM and ArbiTree (Cauchois, Gupta, and Duchi 2021) address set inclusion problems by constructing inner and outer sets, \widehat{C}_{in} and \widehat{C}_{out} , that satisfy $\widehat{C}_{\text{in}} \subset \mathcal{Y} \subset \widehat{C}_{\text{out}}$ with statistical guarantees. These approaches rely on learning a hierarchical tree structure over the label set and designing non-conformity scores tailored to this structure. The prediction sets are then efficiently computed via message passing on the label tree. However, these methods have notable limitations: (i) the tree structure learning in PGM does not scale to large label sets, as it requires solving $O(N^2)$ convex optimization problems without closed-form solutions; and (ii) both the non-conformity score design and prediction set construction are restricted to tree-structured label relationships, making them unsuitable for more general or arbitrary network structures. Conformal risk control (CRC) (Angelopoulos et al. 2024) is a CP approach for set estimation. Our approach can be seen as its score-based version targeting at controlling the recall rate but with different score designs. We present detailed explanation and comparison in the extended version of this work (Jian et al. 2025).

2.2 Hypothesis Testing-based Confident Source Detection

We next briefly review the confident network source detection method ADiT (Dawkins, Li, and Xu 2021). This method addresses the single-source detection problem over general networks under the SI model, providing a prediction set that contains the true source with statistical guarantees. For each node, a hypothesis test is formulated: the null hypothesis is that the node is the source, and the alternative is that it is not. To construct the test statistic, multiple infection paths are simulated from each node, and the average fitness of these paths with respect to the observed infection status is computed. A higher fitness indicates that the node is more likely to be the true source. Based on this fitness measure, hypothesis tests are performed for all nodes, and the prediction set is formed by including those nodes for which the null hypothesis is not rejected.

In summary, ADiT provides assumption-free prediction sets for the source detection problem with statistical guarantees. Its fitness measure, which incorporates infection order, often leads to reasonably small prediction sets. However, the method has notable limitations: (i) it cannot be extended to multi-source scenarios, as hypothesis testing for all possible node subsets is computationally infeasible; and (ii) it cannot be directly applied to more complex models such as SIR, since the fitness function does not account for recovered nodes.

3 Preliminaries

3.1 The Source Detection Problem

In this subsection, we introduce the multi-source detection problem over networks, using a compartmental epidemic model to describe information spreading. The network is represented by a graph $G = (\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the set of nodes and \mathcal{E} is the set of edges, with $|\mathcal{V}| = N$. While our approach is not restricted to a particular diffusion model, *for illustration purposes*, we focus on the SIR model, which is

widely used in both epidemiology and information diffusion studies (Sha, Al Hasan, and Mohler 2021; Shah et al. 2020). In this propagation model, each node belongs to one of three compartments: susceptible (S), infected (I), or recovered (R). A susceptible node can become infected through contact (i.e., an edge) with an infected neighbor, with a certain probability. Infected nodes may recover and transition to the recovered state.

Let $\theta_v(t) \in \{S, I, R\}$ denote the status of node v at time t . The discrete-time dynamics are given by (Keeling and Rohani 2011; Shah et al. 2020; Yan, Fang, and He 2023):

$$\mathbb{P}(\theta_v(t) = I \mid \theta_v(t-1) = S) = 1 - \prod_{\substack{u \in \mathcal{N}(v), \\ \theta_u(t-1) = I}} (1 - \sigma_{\text{inf}}),$$

$$\mathbb{P}(\theta_v(t) = R \mid \theta_v(t-1) = I) = \sigma_{\text{rec}},$$

where $\mathcal{N}(v)$ denotes the neighbors of v , and σ_{inf} and σ_{rec} are the infection and recovery rates, respectively. The *basic reproduction number* is defined as $R_0 := \sigma_{\text{inf}} \lambda_1 / \sigma_{\text{rec}}$, where λ_1 is the largest eigenvalue of the adjacency matrix of G .

The SIR model can be implemented as an independent cascade process: starting from a set of source (infected) nodes $\mathcal{Y} \subset \mathcal{V}$, each infected node attempts to infect its neighbors independently at each time step. The SIR model generalizes the widely used SI model, which is recovered by setting $\sigma_{\text{rec}} = 0$ (i.e., no recovery).

The *multi-source detection problem* is the inverse problem that estimates the *set of infection sources*, assuming snapshots of node status are observed at time points $t_1 < t_2 < \dots < t_M$. For any prediction set \widehat{C} , its performance can be measured by precision $:= \frac{|\widehat{C} \cap \mathcal{Y}|}{|\widehat{C}|}$ and recall $:= \frac{|\widehat{C} \cap \mathcal{Y}|}{|\mathcal{Y}|}$.

3.2 Conformal Prediction

We briefly review the basic pipeline and theoretical results of conformal prediction (CP). There are two main variants: full CP and split CP (Angelopoulos, Barber, and Bates 2025). Split CP is a computationally efficient special case of full CP and is widely used in practice. In this work, we focus on split CP.

Split CP consists of the following components: (i) A *point estimator*, such as a pre-trained neural network, which maps inputs to outputs. Let \mathcal{X} and \mathcal{W} denote the input and output spaces, respectively. In classification tasks, \mathcal{W} is the set of possible labels, and the neural network typically outputs estimated probabilities for each label. (ii) A *non-conformity score* $s : \mathcal{X} \times \mathcal{W} \rightarrow \mathbb{R}$, which quantifies how atypical a given input-output pair is with respect to the model; higher values indicate less conformity. (iii) A *calibration dataset* $\{(\mathbf{X}_i, w_i) : i \in [n]\} \subset \mathcal{X} \times \mathcal{W}$. By evaluating the non-conformity scores on this set, we can empirically estimate the distribution of prediction errors and thus construct a prediction set. Given a new test input \mathbf{X}_{n+1} and a user-specified nominal level α , split CP defines the prediction set as $\widehat{C}(\mathbf{X}_{n+1}) := \{w' \in \mathcal{W} : s(\mathbf{X}_{n+1}, w') \leq \widehat{q}_\alpha\}$, where $\widehat{q}_\alpha := Q(\{s(\mathbf{X}_i, w_i)\}_{i=1}^n, (1 - \alpha)(1 + \frac{1}{n}))$ is the empirical $(1 - \alpha)$ quantile of the calibration scores. This construction yields the following statistical guarantee.

Theorem 1 ((Vovk, Gammernan, and Shafer 2005)). *If the calibration set and the test data are exchangeable, then*

$$\mathbb{P}(w_{n+1} \in \widehat{C}(\mathbf{X}_{n+1})) \geq 1 - \alpha. \quad (1)$$

Among all methods satisfying the *coverage* guarantee (1), a smaller prediction set $\widehat{C}(\mathbf{X}_{n+1})$ is preferred, as its size reflects the *efficiency* of the CP procedure.

3.3 Conformalized Multi-source Detection

We now formulate the *conformalized network multi-source detection* problem. Recall that we have N nodes in a graph G and we observe M snapshots of the node status at M time instances. The input space is $\mathcal{X} = \{S, I, R\}^{N \times M}$, representing the status of each node at each observed time instance. The output space is $2^{\mathcal{V}}$, i.e., the set of all possible subsets of nodes. Suppose we have a pre-defined function $f(\mathbf{X}) = (\widehat{\pi}(v))_{v \in \mathcal{V}} \in \mathbb{R}^{\mathcal{V}}$, where a larger $\widehat{\pi}(v)$ indicates a higher likelihood of $v \in \mathcal{Y}$ given $\mathbf{X} \in \mathcal{X}$. For instance, f can be a pre-trained neural network, and $\widehat{\pi}(v)$ can be an estimate of $\mathbb{P}(v \in \mathcal{Y} | \mathbf{X})$. As in standard CP, we assume access to a calibration set $\{(\mathbf{X}_i, \mathcal{Y}_i) : i \in [n]\} \subset \mathcal{X} \times 2^{\mathcal{V}}$, and denote the test data point as $(\mathbf{X}_{n+1}, \mathcal{Y}_{n+1})$. Given the observed snapshots $\mathbf{X}_{n+1} \in \mathcal{X}$, our goal is to construct a prediction set $\widehat{C}(\mathbf{X}_{n+1}; \alpha, \beta)$ such that

$$\mathbb{P}\left(\frac{|\mathcal{Y}_{n+1} \cap \widehat{C}(\mathbf{X}_{n+1}; \alpha, \beta)|}{|\mathcal{Y}_{n+1}|} \geq 1 - \beta\right) \geq 1 - \alpha, \quad (2)$$

where α and β are user-specified nominal levels. In other words, the prediction set should achieve a recall rate of at least $1 - \beta$ with probability at least $1 - \alpha$.

4 Conformal Prediction for Set Estimation

In this section, we introduce a principled framework for constructing prediction sets that satisfy the recall guarantee in (2), while also achieving high efficiency in both prediction accuracy and computational cost.

4.1 Conformal Prediction for Set Inclusion

We first develop CP approaches for the special case where $\beta = 0$, and then extend this to the general case where $\beta \in [0, 1]$. Note that when $\beta = 0$, (2) becomes

$$\mathbb{P}(\mathcal{Y}_{n+1} \subset \widehat{C}(\mathbf{X}_{n+1}; \alpha, 0)) \geq 1 - \alpha. \quad (3)$$

In what follows, we write $\widehat{C}(\mathbf{X}_{n+1}; \alpha, 0)$ as $\widehat{C}(\mathbf{X}_{n+1}; \alpha)$ for simplicity.

We define the following map (see Fig. 1 for illustration):

$$\begin{aligned} \gamma : \mathcal{X} \times 2^{\mathcal{V}} &\rightarrow 2^{\mathcal{V}} \\ (\mathbf{X}, \mathcal{U}) &\mapsto \{v \in \mathcal{V} : \widehat{\pi}(v) \geq \min_{z \in \mathcal{U}} \widehat{\pi}(z)\}. \end{aligned} \quad (4)$$

Recall that $\widehat{\pi}(v)$ is the output of $f(\mathbf{X})$ on the vertex v for any $v \in \mathcal{V}$. The intuition behind γ is that, adding those nodes from \mathcal{V} with larger predicted probabilities than $\min_{z \in \mathcal{U}} \widehat{\pi}(z)$ to \mathcal{U} should yield better conformity and thus worth considering.

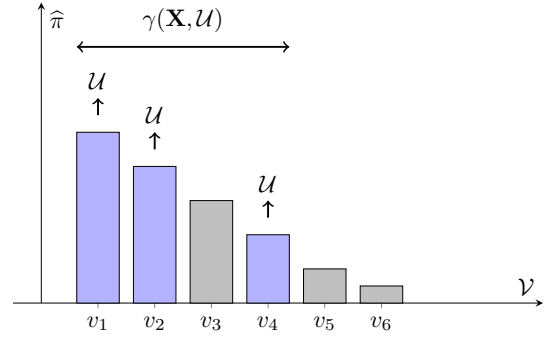


Figure 1: Illustration of $\gamma(\mathbf{X}, \mathcal{U})$. Suppose the nodes are indexed such that $\widehat{\pi}(v_1) \geq \dots \geq \widehat{\pi}(v_N)$. When $\mathcal{U} = \{v_1, v_2, v_4\}$, $\gamma(\mathbf{X}, \mathcal{U}) := \{v_1, v_2, v_3, v_4\}$.

Let the non-conformity score be a function

$$\begin{aligned} s : \mathcal{X} \times 2^{\mathcal{V}} &\rightarrow \mathbb{R} \\ (\mathbf{X}, \mathcal{U}) &\mapsto s(\mathbf{X}, \gamma(\mathbf{X}, \mathcal{U})), \end{aligned} \quad (5)$$

i.e., it operates on $2^{\mathcal{V}}$ through the function γ . We suppose that it is monotone: if $\gamma(\mathbf{X}, \mathcal{U}_1) \subset \gamma(\mathbf{X}, \mathcal{U}_2)$,

$$s(\mathbf{X}, \gamma(\mathbf{X}, \mathcal{U}_1)) \leq s(\mathbf{X}, \gamma(\mathbf{X}, \mathcal{U}_2)). \quad (6)$$

Below we provide two examples:

$$s_{\text{pre}} : (\mathbf{X}, \mathcal{U}) \mapsto -\frac{1}{|\gamma(\mathbf{X}, \mathcal{U})|} \sum_{z \in \gamma(\mathbf{X}, \mathcal{U})} \widehat{\pi}(z), \quad (7)$$

$$s_{\text{rec}} : (\mathbf{X}, \mathcal{U}) \mapsto \frac{\sum_{z \in \gamma(\mathbf{X}, \mathcal{U})} \widehat{\pi}(z)}{\sum_{z \in \mathcal{V}} \widehat{\pi}(z)}. \quad (8)$$

It can be shown that s_{pre} and s_{rec} satisfy (5) and (6).

Remark 1. *The score designs (7) and (8) are meaningful since they represent the non-conformity in terms of precision and recall rate respectively, under the oracle case where $\widehat{\pi}(v) = \mathbb{P}(v \in \mathcal{Y}_{n+1} | \mathbf{X}_{n+1})$ for all $v \in \mathcal{V}$. For any fixed $\mathcal{U} \subset \mathcal{V}$, the expected precision of the set $\gamma(\mathbf{X}, \mathcal{U})$ given \mathbf{X}_{n+1} can be calculated as*

$$\begin{aligned} &\mathbb{E}\left[\frac{|\gamma(\mathbf{X}, \mathcal{U}) \cap \mathcal{Y}_{n+1}|}{|\gamma(\mathbf{X}, \mathcal{U})|} \mid \mathbf{X}_{n+1}\right] \\ &= \frac{1}{|\gamma(\mathbf{X}, \mathcal{U})|} \sum_{z \in \gamma(\mathbf{X}, \mathcal{U})} \mathbb{E}[\mathbf{1}\{z \in \mathcal{Y}_{n+1}\} \mid \mathbf{X}_{n+1}], \end{aligned}$$

which coincides with (7) under the oracle case up to a minus sign. Using similar arguments, the recall rate of the set $\gamma(\mathbf{X}, \mathcal{U})$ given \mathbf{X}_{n+1} can be calculated as

$$\frac{\mathbb{E}[|\gamma(\mathbf{X}, \mathcal{U}) \cap \mathcal{Y}_{n+1}| \mid \mathbf{X}_{n+1}]}{\mathbb{E}[|\mathcal{Y}_{n+1}| \mid \mathbf{X}_{n+1}]} = \frac{\sum_{z \in \gamma(\mathbf{X}, \mathcal{U})} \widehat{\pi}(z)}{\sum_{z \in \mathcal{V}} \widehat{\pi}(z)},$$

which coincides with (8).

Let $\widehat{q}_\alpha := Q(\{s(\mathbf{X}_i, \mathcal{Y}_i)\}_{i=1}^n, (1 - \alpha)(1 + \frac{1}{n}))$. The prediction set is defined as

$$\widehat{C}(\mathbf{X}_{n+1}; \alpha) := \{v \in \mathcal{V} : s(\mathbf{X}_{n+1}, \{v\}) \leq \widehat{q}_\alpha\}. \quad (9)$$

Next, we prove that $\widehat{C}(\mathbf{X}_{n+1}; \alpha)$ satisfies (3). This is the result of the following key observation.

Proposition 1. Under the conditions (5) and (6), the following two events are equivalent:

$$\mathcal{Y}_{n+1} \subset \widehat{C}(\mathbf{X}_{n+1}; \alpha); \quad (10)$$

$$s(\mathbf{X}_{n+1}, \mathcal{Y}_{n+1}) \leq \widehat{q}_\alpha. \quad (11)$$

Proposition 1 yields the inclusion guarantee (3) as follows.

Theorem 2. Assume that $\{(\mathbf{X}_i, \mathcal{Y}_i) : i \in [n+1]\}$ are exchangeable. Then $\widehat{C}(\mathbf{X}_{n+1}; \alpha)$ (cf. (9)) satisfies the inclusion guarantee (3).

Furthermore, we show that (9) can also be obtained by applying our proposed scores within existing CP frameworks (Jian et al. 2025, Section VII).

4.2 Conformal Prediction with Recall Rate Guarantee

Theorem 2 establishes the validity of the prediction set construction in (9) for the special case where the nominal recall rate is 1 (i.e., $\beta = 0$ in (2)). We now generalize this approach to handle arbitrary recall levels $\beta \in [0, 1]$. Our strategy consists of two steps: first, we define a map that shrinks the ground truth set \mathcal{Y} by retaining only a $(1 - \beta)$ fraction of its elements; second, we apply the prediction set construction from (9) to ensure inclusion of this shrunken set.

To shrink sets, we define the map $\nu : \mathcal{X} \times 2^{\mathcal{Y}} \rightarrow 2^{\mathcal{Y}}$ such that

$$\nu(\mathbf{X}, \mathcal{U}) \subset \mathcal{U} \quad (12)$$

$$|\nu(\mathbf{X}, \mathcal{U})| \geq (1 - \beta)|\mathcal{U}|. \quad (13)$$

for any $\mathcal{U} \subset \mathcal{Y}$. Then replacing \mathcal{Y}_i by the shrunken sets $\nu(\mathbf{X}_i, \mathcal{Y}_i)$ for all $i \in [n+1]$ and formulating the prediction set via (9), we obtain a set $\widehat{C}(\mathbf{X}_{n+1}; \alpha, \beta)$ such that

$$\mathbb{P}(\nu(\mathbf{X}_{n+1}, \mathcal{Y}_{n+1}) \subset \widehat{C}(\mathbf{X}_{n+1}; \alpha, \beta)) \geq 1 - \alpha. \quad (14)$$

Note that $\nu(\mathbf{X}_{n+1}, \mathcal{Y}_{n+1}) \subset \widehat{C}(\mathbf{X}_{n+1}; \alpha, \beta)$ implies

$$\frac{|\mathcal{Y}_{n+1} \cap \widehat{C}(\mathbf{X}_{n+1}; \alpha, \beta)|}{|\mathcal{Y}_{n+1}|} \geq 1 - \beta.$$

Combining this with (14) we obtain the following result.

Theorem 3. Suppose the map ν satisfies (12) and (13). Let

$$\widehat{q}_\alpha := \mathbb{Q}\left(\{s(\mathbf{X}_i, \nu(\mathbf{X}_i, \mathcal{Y}_i))\}_{i=1}^n, (1 - \alpha)\left(1 + \frac{1}{n}\right)\right),$$

and define the prediction set as

$$\widehat{C}(\mathbf{X}_{n+1}; \alpha, \beta) := \{v \in \mathcal{Y} : s(\mathbf{X}_{n+1}, \{v\}) \leq \widehat{q}_\alpha\}. \quad (15)$$

Then, this set satisfies the recall rate guarantee (2).

In the above argument, the only requirements on the shrinking map ν are (12) and (13), i.e., $\nu(\mathbf{X}, \mathcal{U})$ should be a subset of \mathcal{U} while maintaining at least $(1 - \beta)$ proportion of elements. One natural construction of such a map is to retain the $(1 - \beta)$ proportion of elements with the largest predicted probability (cf. Fig. 2), i.e.,

$$\begin{aligned} & \nu(\mathbf{X}_{n+1}, \mathcal{Y}_{n+1}) \\ & := \{v \in \mathcal{Y}_{n+1} : \widehat{\pi}(v) \geq -\mathbb{Q}(\{-\widehat{\pi}(v)\}_{v \in \mathcal{Y}_{n+1}}, 1 - \beta)\}. \end{aligned} \quad (16)$$

Finally, we note that this shrinking strategy can be combined with any existing CP approach that achieves the inclusion guarantee in (3), such as ArbiTree (Cauchois, Gupta, and Duchi 2021). By applying the same reasoning, the resulting method also satisfies the recall rate guarantee in (2).

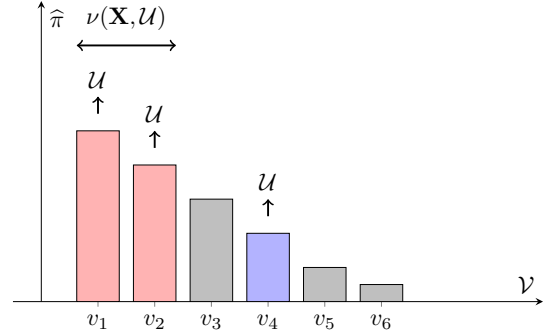


Figure 2: Illustration of $\nu(\mathbf{X}, \mathcal{U})$. Suppose the nodes are indexed such that $\widehat{\pi}(v_1) \geq \dots \geq \widehat{\pi}(v_N)$. When $\mathcal{U} = \{v_1, v_2, v_4\}$, $\beta = \frac{1}{3}$, $\nu(\mathbf{X}, \mathcal{U}) = \{v_1, v_2\}$.

5 Experiments

In this section, we present numerical results for conformalized multi-source detection under various propagation patterns on real-world networks. Our objectives are twofold: (i) to achieve the recall rate guarantee in (2); and (ii) to maximize efficiency by minimizing prediction set sizes.

We implement our method using two non-conformity score designs s_{pre} and s_{rec} (see (7) and (8)), referred to as setCP-pre and setCP-rec, respectively. For comparison, we include the following baselines: ADiT (Dawkins, Li, and Xu 2021) and ArbiTree (Cauchois, Gupta, and Duchi 2021). All methods except ADiT are based on the conformal prediction (CP) framework and thus require a pre-trained neural network. We use the SD-STGCN model (Sha, Al Hasan, and Mohler 2021) as the backbone, modifying its output to $N \times 2$ channels and adapting the loss function for binary node classification in the multi-source detection setting.

We evaluate the proposed and baseline methods across a range of nominal levels (α, β) to assess their ability to achieve the desired recall guarantees. In addition, we systematically investigate performance under various practical scenarios, including: (i) varying the number of sources; (ii) different diffusion models (e.g., SI and SIR); (iii) a range of propagation speeds; and (iv) multiple real-world network topologies.

Datasets. We simulate SIR and SI propagation processes over three social networks: highSchool, bkFratB, and sfhh, which are obtained by aggregating contact records (Sha, Al Hasan, and Mohler 2021) within groups of people.

Experimental setup. For all experiments, we use a calibration set of size $n = 7600$ and a test set of size 400. The pre-trained model is trained on a separate set of 20,000 samples. Each experiment is repeated over 50 random splits of the calibration and test sets to ensure robustness of the results.

Method	β	Inclusion Rates			Prediction Set Size		
		$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.15$	$\alpha = 0.05$	$\alpha = 0.10$	$\alpha = 0.15$
setCP-rec	0.1	1.000 \pm 0.000	1.000 \pm 0.000	1.000 \pm 0.000	774.000 \pm 0.000	774.000 \pm 0.000	774.000 \pm 0.000
	0.3	1.000 \pm 0.000	0.897 \pm 0.012	0.845 \pm 0.016	774.000 \pm 0.000	16.246 \pm 0.479	15.017 \pm 0.434
	0.5	0.950 \pm 0.010	0.899 \pm 0.013	0.847 \pm 0.015	11.856 \pm 0.338	9.874 \pm 0.279	8.912 \pm 0.248
	0.7	0.947 \pm 0.010	0.898 \pm 0.014	0.848 \pm 0.016	8.949 \pm 0.273	6.570 \pm 0.186	5.307 \pm 0.148
setCP-pre	0.1	0.951 \pm 0.011	0.899 \pm 0.016	0.852 \pm 0.016	25.069 \pm 0.566	22.576 \pm 0.511	20.640 \pm 0.463
	0.3	0.951 \pm 0.012	0.899 \pm 0.016	0.850 \pm 0.018	21.333 \pm 0.486	18.622 \pm 0.429	16.599 \pm 0.391
	0.5	0.949 \pm 0.009	0.898 \pm 0.015	0.847 \pm 0.017	16.493 \pm 0.384	<u>13.670 \pm 0.327</u>	<u>11.719 \pm 0.291</u>
	0.7	0.948 \pm 0.009	0.899 \pm 0.014	0.849 \pm 0.016	<u>13.056 \pm 0.307</u>	<u>9.669 \pm 0.283</u>	<u>7.860 \pm 0.273</u>
ArbiTree	0.1	0.993 \pm 0.005	0.980 \pm 0.008	0.961 \pm 0.010	771.022 \pm 1.054	764.048 \pm 4.216	750.241 \pm 10.979
	0.3	0.992 \pm 0.005	0.973 \pm 0.009	0.941 \pm 0.014	<u>763.480 \pm 3.888</u>	741.736 \pm 7.911	709.067 \pm 10.800
	0.5	0.983 \pm 0.007	0.948 \pm 0.013	0.910 \pm 0.014	748.169 \pm 6.251	710.206 \pm 10.059	669.028 \pm 11.733
	0.7	0.969 \pm 0.008	0.928 \pm 0.012	0.881 \pm 0.014	727.488 \pm 9.241	681.732 \pm 10.512	637.896 \pm 12.525

Table 1: Inclusion rates and prediction set sizes under the SIR model with random parameters ($|\mathcal{Y}| \in [15]$, $R_0 \in [1, 15]$, $\sigma_{\text{rec}} \in [0.1, 0.4]$) over the highSchool network ($N = 774$) for methods setCP-pre, setCP-rec, and ArbiTree. The best and the second-best results under each (α, β) configuration are **boldfaced** and underlined, respectively.

Model	Method	Inclusion Rates			Prediction Set Size		
		$ \mathcal{Y} = 1$	$ \mathcal{Y} = 7$	$ \mathcal{Y} = 10$	$ \mathcal{Y} = 1$	$ \mathcal{Y} = 7$	$ \mathcal{Y} = 10$
SI	setCP-rec	0.901 \pm 0.015	0.903 \pm 0.016	0.899 \pm 0.014	23.315 \pm 0.626	29.053 \pm 0.288	40.309 \pm 0.437
	setCP-pre	0.902 \pm 0.013	0.900 \pm 0.014	0.904 \pm 0.015	25.691 \pm 0.324	<u>37.438 \pm 0.372</u>	<u>52.139 \pm 0.513</u>
	ADiT	0.988 \pm 0.005	-	-	24.432 \pm 0.698	-	-
	ArbiTree	0.896 \pm 0.014	1.000 \pm 0.001	1.000 \pm 0.001	624.442 \pm 14.545	770.490 \pm 1.087	770.933 \pm 0.774
SIR	Method	$ \mathcal{Y} =1$	$ \mathcal{Y} =7$	$ \mathcal{Y} =10$	$ \mathcal{Y} =1$	$ \mathcal{Y} =7$	$ \mathcal{Y} =10$
	setCP-rec	0.901 \pm 0.015	0.903 \pm 0.017	0.903 \pm 0.015	23.045 \pm 0.607	27.996 \pm 0.308	37.732 \pm 0.311
	setCP-pre	0.900 \pm 0.015	0.902 \pm 0.017	0.901 \pm 0.015	23.551 \pm 0.366	34.451 \pm 0.287	49.255 \pm 0.502
	ArbiTree	0.902 \pm 0.016	1.000 \pm 0.000	1.000 \pm 0.000	607.128 \pm 13.636	768.980 \pm 0.863	769.771 \pm 1.178

Table 2: Inclusion rates and prediction set sizes under the SI model ($\sigma_{\text{inf}} = 0.25$) and SIR model ($\sigma_{\text{inf}} = 0.25$, $\sigma_{\text{rec}} = 0.15$) over the highSchool network ($N = 774$) with different number of sources (denoted by $|\mathcal{Y}|$). When $|\mathcal{Y}| > 1$, nominal levels are set as $\alpha = 0.1$, $\beta = 0.3$. When $|\mathcal{Y}| = 1$, $\beta = 0$.

Method	R_0	Inclusion Rates			Prediction Set Size		
		highSchool	bkFratB	sfhh	highSchool	bkFratB	sfhh
setCP-rec	[1, 15]	0.897 \pm 0.012	0.902 \pm 0.016	0.901 \pm 0.014	16.246 \pm 0.479	13.415 \pm 0.426	14.740 \pm 0.485
	[11, 25]	0.894 \pm 0.018	0.902 \pm 0.015	0.902 \pm 0.016	26.340 \pm 0.783	19.425 \pm 0.569	22.862 \pm 0.859
	[21, 35]	0.897 \pm 0.015	0.895 \pm 0.017	0.901 \pm 0.017	37.055 \pm 1.270	23.763 \pm 0.547	30.245 \pm 0.965
setCP-pre	[1, 15]	0.899 \pm 0.016	0.896 \pm 0.014	0.899 \pm 0.013	18.622 \pm 0.429	16.256 \pm 0.478	16.107 \pm 0.441
	[11, 25]	0.898 \pm 0.018	0.900 \pm 0.016	0.900 \pm 0.015	<u>31.705 \pm 0.902</u>	<u>24.867 \pm 0.778</u>	<u>25.854 \pm 0.651</u>
	[21, 35]	0.904 \pm 0.016	0.895 \pm 0.015	0.901 \pm 0.013	<u>47.044 \pm 1.553</u>	<u>31.433 \pm 0.842</u>	<u>36.330 \pm 0.941</u>
ArbiTree	[1, 15]	0.973 \pm 0.009	0.988 \pm 0.010	0.973 \pm 0.013	741.736 \pm 7.911	56.618 \pm 0.683	384.167 \pm 4.975
	[11, 25]	0.992 \pm 0.004	0.996 \pm 0.005	0.980 \pm 0.008	761.811 \pm 3.511	57.528 \pm 0.398	390.277 \pm 3.583
	[21, 35]	0.990 \pm 0.006	0.998 \pm 0.004	0.987 \pm 0.006	762.320 \pm 4.470	57.816 \pm 0.264	394.993 \pm 2.315

Table 3: Inclusion rates and prediction set sizes under the SIR model with random parameters ($|\mathcal{Y}| \in [15]$, $\sigma_{\text{rec}} \in [0.1, 0.4]$) over the highSchool, bkFratB, and sfhh networks ($N = 774, 58, 403$, respectively). R_0 are drawn from different ranges to represent different propagation speed. Nominal levels are set as $\alpha = 0.1$, $\beta = 0.3$.

All performance differences are significant in terms of a level 0.95 Wilcoxon signed-rank test.

Further details regarding the dataset and experimental settings are included in the extended version (Jian et al. 2025).

5.1 Conformalized Detection under Different Nominal Levels

We evaluate the quality of prediction sets derived by different methods under different combinations of nominal levels α and β . According to (2), it is expected that the prediction sets cover at least $(1 - \beta)$ proportion of the source set with probability at least $(1 - \alpha)$. Besides, the size of the prediction sets should be as small as possible. We implement the methods under SIR models with parameters σ_{inf} , σ_{rec} , R_0 randomly drawn from uniform distributions over certain ranges. Since ADiT cannot be applied under multi-source detection problem, it is not included in this comparison.

In Table 1, we report the empirical results for different combinations of (α, β) . Across all settings, each method achieves the desired recall rate $(1 - \beta)$ with probability at least $(1 - \alpha)$ (up to a standard error). This provides empirical support for Theorem 3, confirming the validity of the statistical guarantees for our proposed non-conformity scores and prediction set constructions. Furthermore, these results demonstrate that the shrinking strategy using the map ν (see (16)) effectively calibrates the recall rate to the user-specified nominal level, not only for our proposed set construction in (15), but also for alternative methods such as ArbiTree.

Across most nominal level settings, setCP-rec and setCP-pre consistently achieve the best or second-best efficiency in terms of predict set size, producing substantially smaller prediction sets than ArbiTree. In contrast, ArbiTree yields prediction sets that are nearly as large as the entire graph ($N = 774$), making them impractically large for informative inference. When the nominal recall rate is high (i.e., β is small), setCP-rec also produces large, uninformative prediction sets. This limitation arises from its score definition in (8), which does not penalize the size of the prediction set. In the extreme case where $\beta = 0$ (requiring all source nodes to be included), if the GNN assigns zero probability to any true source node, s_{rec} will result in the entire node set \mathcal{V} being selected. In contrast, s_{pre} , which incorporates set size into its formulation, yields more efficient prediction sets.

Finally, we remark that, since all methods fulfill the statistical guarantees, in practice, different CP approaches can be utilized, and the smallest prediction set can be chosen.

5.2 Conformalized Detection under Different Propagation Models

We next compare the prediction performance of different methods across various propagation models and network topologies. The observed inclusion rates further confirm the statistical guarantees established in Theorem 3, while the sizes of the prediction sets provide insight into the inherent difficulty of each scenario.

In Table 2, we present the numerical results for both SI and SIR models with varying numbers of sources. All methods achieve the desired recall rate guarantee, while setCP-rec

setCP-rec	setCP-pre	ADiT	ArbiTree
1.392 ± 0.008	1.559 ± 0.008	792.170 ± 23.936	79.637 ± 1.208

Table 4: Average time (in seconds) for computing on 400 test samples under the SI model ($\sigma_{\text{inf}} = 0.25$), $|\mathcal{Y}| = 1$, with $\alpha = 0.05, 0.07, 0.10, 0.15, 0.20$.

consistently yields the smallest prediction sets. Additionally, for a fixed number of sources and method, the prediction sets under the SIR model are always smaller than those under the SI model. This is because the SIR model provides additional information through the recovery status, making the inverse problem easier and resulting in more efficient (smaller) prediction sets.

Table 3 reports results across different social networks. As before, all methods satisfy the recall rate guarantee, and setCP-rec consistently achieves the smallest prediction sets. For a fixed network and method, the prediction set size increases with larger values of R_0 . This is expected, as a higher R_0 corresponds to a faster infection rate σ_{inf} , making the inverse problem more challenging.

The execution time for evaluating each method on the test data is shown in Table 4. ADiT and ArbiTree are implemented by parallel computing to manage their computational demands, as using a single thread would result in excessive time costs. In contrast, setCP-rec and setCP-pre operate on a single thread without parallel computing. The results demonstrate that setCP-rec and setCP-pre are substantially faster than both ADiT and ArbiTree. Although the computational complexity for evaluating the prediction set in (15) is $O(N \log N)$ —slightly larger than ArbiTree and ADiT with complexity $O(N)$ —the practical runtimes differ significantly. This is because ArbiTree requires two recursive message-passing steps over the label tree, whereas setCP-rec and setCP-pre only require a single pass to compute scores for all nodes followed by sorting. In contrast, ADiT performs a hypothesis test for each node, which involves estimating the distribution of the test statistic via multiple Monte Carlo simulations, further increasing its computational cost.

6 Conclusion and Limitations

In this paper, we introduced a systematic CP framework for network multi-source detection, capable of achieving user-specified recall rates with rigorous statistical guarantees. Theoretically, our approach relies only on the minimal assumption of data exchangeability, offering a general solution to conformalized set estimation problems. Empirically, the method demonstrates superior efficiency compared to existing approaches, both in terms of prediction accuracy and computational performance. One limitation of our approach is its dependence on having access to a calibration set of diffusion data. In some applications, this can be synthetically generated based on prior knowledge of the propagation model and parameter ranges, which may require external expert input. Future work involves studying how mismatch in this prior knowledge affects the performance of our method, and how to adaptively update our prediction using new data.

Acknowledgments

X.J. and F.K. acknowledge support by the German Ministry of Education and Research (BMBF) in the context of the Munich Center for Machine Learning (MCML) and by the German Science Foundation (DFG) in the context of the project Solving linear inverse problems with end-to-end neural networks: expressivity, generalization, and robustness project number 464123524 as part of the Priority Program 2298.

References

- Angelopoulos, A. N.; Barber, R.; and Bates, S. 2024. Online conformal prediction with decaying step sizes. In *ICML*, volume 235, 1616–1630.
- Angelopoulos, A. N.; Barber, R. F.; and Bates, S. 2025. Theoretical Foundations of Conformal Prediction. *arXiv preprint arXiv:2411.11824*.
- Angelopoulos, A. N.; and Bates, S. 2023. *Conformal Prediction: A Gentle Introduction*. Now Foundations and Trends.
- Angelopoulos, A. N.; Bates, S.; Fisch, A.; Lei, L.; and Schuster, T. 2024. Conformal risk control. In *ICLR*.
- Bakshy, E.; Hofman, J. M.; Mason, W. A.; and Watts, D. J. 2011. Everyone’s an influencer: quantifying influence on Twitter. In *Proc. ACM Int. Conf. Web Search Data Min.*, 65–74.
- Bhatnagar, A.; Wang, H.; Xiong, C.; and Bai, Y. 2023. Improved online conformal prediction via strongly adaptive online learning. In *ICML*. Honolulu, Hawaii, USA.
- Cauchois, M.; Gupta, S.; and Duchi, J. C. 2021. Knowing what You Know: valid and validated confidence sets in multi-class and multilabel prediction. *J. Mach. Learn. Res.*, 22(81): 1–42.
- Chen, Z.; Zhu, K.; and Ying, L. 2014. Detecting multiple information sources in networks under the SIR model. In *Proc. Annu. Conf. Inf. Sci. Syst.*
- Dawkins, Q. E.; Li, T.; and Xu, H. 2021. Diffusion source identification on networks with statistical confidence. In *ICML*.
- Gibbs, I.; and Candès, E. 2024. Conformal inference for online prediction with arbitrary distribution shifts. *J. Mach. Learn. Res.*, 25(1).
- H. Zargarbashi, S.; Antonelli, S.; and Bojchevski, A. 2023. Conformal Prediction Sets for Graph Neural Networks. In *ICML*.
- Hu, G., T.; et al. 2020. Building an open resources repository for COVID-19 research. *Data and Information Management*, 4(3): 130–147.
- Huang, K.; Jin, Y.; Candès, E.; and Leskovec, J. 2023. Uncertainty quantification over graph with conformalized graph neural networks. In *NeurIPS*. New Orleans, LA, USA.
- Ji, F.; Tang, W.; and Tay, W. P. 2019. On the properties of Gromov matrices and their applications in network inference. *IEEE Trans. Signal Process.*, 67(10): 2624–2638.
- Jian, X.; Zhang, P.; Tian, L.; Ji, F.; Liang, W.; Tay, W. P.; Wen, B.; and Kraemer, F. 2025. Conformal Prediction for Multi-Source Detection on a Network. *arXiv preprint: arXiv:2511.08867*.
- Katsios, K.; and Papadopoulos, H. 2024. Multi-label conformal prediction with a mahalanobis distance nonconformity measure. *Proceedings of Machine Learning Research*, 230: 1–14.
- Keeling, M. J.; and Rohani, P. 2011. *Modeling infectious diseases in humans and animals*. Princeton University Press.
- Luo, W.; Tay, W. P.; and Leng, M. 2013. Identifying Infection Sources and Regions in Large Networks. *IEEE Trans. Signal Process.*, 61(11): 2850–2865.
- Luo, W.; Tay, W. P.; and Leng, M. 2017. On the universality of Jordan centers for estimating infection sources in tree network. *IEEE Trans. Inf. Theory*, 63(7): 4634–4657.
- Maltoudoglou, L.; Paisios, A.; Lenc, L.; Martínek, J.; Král, P.; and Papadopoulos, H. 2022. Well-calibrated confidence measures for multi-label text classification with a large number of labels. *Pattern Recognition*, 122: 108271.
- Mitchell, A.; Kiley, J.; Gottfried, J.; and Guskin, E. 2013. The role of news on Facebook: Common yet incidental. *Pew Research Center, Tech. Rep.*
- Romano, Y.; Sesia, M.; and Candès, E. J. 2020. Classification with valid and adaptive coverage. In *NeurIPS*. Vancouver, BC, Canada.
- Sha, H.; Al Hasan, M.; and Mohler, G. 2021. Source detection on networks using spatial temporal graph convolutional networks. In *IEEE DSAA*.
- Shah, C.; Dehmamy, N.; Perra, N.; Chinazzi, M.; Barabasi, A.-L.; Vespignani, A.; and Yu, R. 2020. Finding patient zero: Learning contagion source with graph neural networks. *arXiv preprint: arXiv:2006.11913v2*.
- Shah, D.; and Zaman, T. 2011. Rumors in a Network: Who’s the Culprit? *IEEE Trans. Inf. Theory*, 57(8): 5163–5181.
- Shelke, S.; and Attar, V. 2019. Source detection of rumor in social network – A review. *Online Social Networks and Media*, 9: 30–42.
- Tang, W.; Ji, F.; and Tay, W. P. 2018. Estimating Infection Sources in Networks Using Partial Timestamps. *IEEE Trans. Inf. Foren. Secur.*, 13(12): 3035–3049.
- Tyagi, C.; and Guo, W. 2023. Multi-label Classification under Uncertainty: A Tree-based Conformal Prediction Approach. In *Proceedings of the Twelfth Symposium on Conformal and Probabilistic Prediction with Applications*, 488–512.
- Vovk, V.; Gammerman, A.; and Shafer, G. 2005. *Algorithmic learning in a random world*. Springer.
- Wang, Z.; Dong, W.; Zhang, W.; and Tan, C. W. 2015. Rooting our rumor sources in online social networks: The value of diversity from multiple observations. *IEEE J. Sel. Top. Signal Process.*, 9(4): 663–677.
- Yan, X.; Fang, H.; and He, Q. 2023. Diffusion Model for Graph Inverse Problems: Towards Effective Source Localization on Complex Networks. In *NeurIPS*, 22326–22350.
- Zargarbashi, S. H.; and Bojchevski, A. 2024. Conformal inductive graph neural networks. In *ICML*.
- Zhu, K.; and Ying, L. 2014. A robust information source estimator with sparse observations. In *IEEE INFOCOM*, 2211–2219.