TCCD: Tree-guided Continuous Causal Discovery via Collaborative MCTS-Parameter Optimization

Jingjin Liu[1] , **Yingkai Xiao**[1] , **Hankz Hankui Zhuo**[2,3*] **and Wushao Wen**[1]
[1]School of Computer Science and Engineering, Sun Yat-sen University
[2]National Key Laboratory for Novel Software Technology, Nanjing University
[3]School of Artificial Intelligence, Nanjing University
{liujj263@mail2, xiaoyk8@mail2, wenwsh@mail}.sysu.edu.cn, hankz@nju.edu.cn

Abstract

Learning causal relationships in directed acyclic graphs (DAGs) from multi-type event sequences is a challenging task, especially in large-scale telecommunication networks. Existing methods struggle with the exponentially growing search space and lack global exploration. Gradient-based approaches are limited by their reliance on local information and often fail to generalize. To address these issues, we propose TCCD, a framework that combines Monte Carlo Tree Search (MCTS) with continuous gradient optimization. TCCD balances global exploration and local optimization, overcoming the shortcomings of purely gradient-based methods and enhancing generalization. By unifying various causal structure learning approaches, TCCD offers a scalable and efficient solution for causal inference in complex networks. Extensive experiments validate its superior performance on both synthetic and real-world datasets. Code and Appendix are available at https://github.com/jzephyrl/TCCD.

1 Introduction

In various real-world applications, uncovering causal relationships among event types within multi-type event sequences, which are often represented as Directed Acyclic Graphs (DAGs), is a critical and inherently challenging task. In modern telecommunication networks, a single fault can trigger a cascade of alarm events across interconnected devices, driven by complex topological relationships between them. As the scale of networks expands and the complexity of inter-device topologies increases, accurately discerning the causal relationships among these alarm events becomes increasingly challenging [Chickering, 1996; Li *et al.*, 2022].

Recently, Granger causality analysis [Shojaie and Fox, 2022] has become central to causal structure learning. Approaches such as multivariate Hawkes processes [Liniger, 2009] and Poisson point processes [Itô, 1971] have been demonstrated to effectively capture event dynamics and inter-type interactions. However, traditional causal learning methods [Spirtes and Glymour, 1991; Zhu *et al.*, 2019; Idé *et al.*, 2021; Bhattacharjya *et al.*, 2022; Qiao *et al.*, 2023] typically assume i.i.d. event sequences, which contradicts the topological dependencies in telecommunication networks. Consequently, the task of effectively learning causal relationships within such complex network topologies has become one of the main challenges in causal learning for alarm-type events.

To overcome the problem of learning Granger causal structures from event sequences under the non-i.i.d. assumption, [Cai *et al.*, 2022] proposed the Topological Hawkes Process (THP), which extends the traditional Hawkes process to time-graph domains for modeling topological event sequences. However, THP struggles with capturing complex generation mechanisms and nonlinear dependencies. Additionally, its reliance on gradient-free optimization and decoupled parameter learning results in inefficient training. To improve modeling capabilities, [Liu *et al.*, 2024] introduced the Topological Neural Poisson Autoregressive model (TNPAR), which utilizes neural point processes to better model event sequences. Despite its enhanced expressiveness, TNPAR suffers from limited training efficiency and scalability due to the lack of analytical likelihood functions. In contrast, S^2GCSL [Li *et al.*, 2024] employs a linear kernel to model event activation effects and adopts gradient descent optimization to improve the learning efficiency while maintaining model expressiveness.

Despite advances in causal structure learning, existing methods face critical challenges in large-scale telecommunication networks: (1) The search space for causal structures grows exponentially with network scale and event types, hindering efficient explorations; (2) Dynamic topologies and increasing event types limit the generalization ability of current approaches, making it difficult to accurately estimate spatio-temporal continuous parameters; (3) The lack of effective integration between structure search and parameter optimization makes gradient-based methods fall into local optima, constraining overall model performance.

In response to the above limitations, we present a novel framework, namely TCCD, which integrates Monte Carlo Tree Search (MCTS) with continuous optimization. This combination fully utilizes the exploration-exploitation trade-off of MCTS to adaptively balance global exploration and local optimization, enabling effective joint optimization of structure search and parameter estimation. Inspired by

*Corresponding author

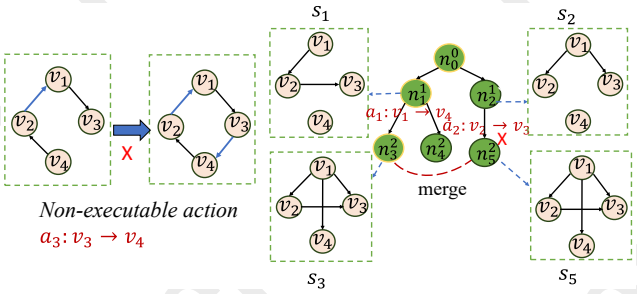


Figure 1: The example of Non-executable action and merged states.

S^{2}GCSL [Li *et al.*, 2024], we refine the likelihood function design to transform the discrete DAG search problem into a continuous optimization task, thereby enhancing the generalizability of the model. Specifically, this hybrid approach ensures effective exploration of large-scale search spaces while simultaneously improving the accuracy of causal structure inference. Ultimately, our TCCD framework offers a novel solution for large-scale causal structure learning in telecommunications networks, enabling more precise and efficient handling of complex causal relationships.

In summary, we provide the following key contributions:

- We design a tree-guided dynamic parameter learning strategy to achieve joint optimization of structure search and parameter estimation, providing richer search directions for gradient optimization and forming a hierarchical framework of global exploration and local optimization;
- We propose a local optimization mechanism based on influence propagation that enables efficient parameter updates by identifying the impact scope of causal structure changes;
- We develop a hierarchical evaluation approach that unifies different scoring methods, demonstrating strong scalability for large-scale causal inference across multiple event types, complex topologies, and dynamic characteristics.

In light of this, we aim to effectively learn complex causal structures when handling large-scale multi-event data, providing accurate fault prediction and diagnosis support for network operators. Experiments on synthetic and real-world telecommunication fault diagnosis datasets validate the accuracy of the inference results of the TCCD framework in large-scale scenarios.

2 Related Work

Causal structure learning [Zheng *et al.*, 2018; Tsamardinos *et al.*, 2006; Lachapelle *et al.*, 2019; Wang *et al.*, 2021; Cheng *et al.*, 2023] has emerged as a vital area of research due to its wide-ranging applications. Broadly, existing methods are mainly categorized into constraint-based and score-based approaches. The PC algorithm [Spirtes *et al.*, 2000] constructs causal graphs through conditional independence tests, while score-based methods search for optimal DAGs using criteria like Bayesian Information Criterion (BIC) or Minimum Description Length (MDL). However, these conventional approaches often struggle with large-scale data. Recent work has focused on multi-type event sequences, particularly through multivariate Hawkes processes based on Granger causality [Shojaie and Fox, 2022; Zhang *et al.*, 2020; Qiao *et al.*, 2023; Xu *et al.*, 2016; Achab *et al.*, 2018]. ADM4 [Zhou *et al.*, 2013] uses kernel methods and L_1 regularization for sparse learning, while deep point process methods like RPPN [Zhang *et al.*, 2020] and CAUSE incorporate attention mechanisms and neural networks. To handle non-i.i.d. data, THP [Cai *et al.*, 2022] accounts for topological influences, and TNPAR [Liu *et al.*, 2024] combines topological networks with Granger causality. Recently, S^{2}GCSL [Li *et al.*, 2024] enhanced learning efficiency through linear kernel modeling and gradient descent optimization. Building on these advances, we propose a novel framework TCCD that utilizes tree-guided continuous causal discovery through collaborative optimization of MCTS parameters, with the objective of addressing the challenges of learning causal structures.

3 Preliminaries

In this work, we focus on revealing the causal structure among various event types within sequences of events. Concretely, a multi-type **event dataset** (or **event stream**) $\mathcal{X} = \{\mathcal{X}(t_j)\}_{j=1}^{M}$ is a sequence of events, where each event $\mathcal{X}(t_j) = (d_j, v_j, t_j)$ indicates the occurrence of an event label v_j (or synonymously type or variable) at a spatial location d_j at a timestamp t_j. Here, $v_j \in \mathcal{V}$ denotes the event type, $d_j \in \mathcal{D}$ corresponds to the nodes in the topological graph $\mathcal{G}_\mathcal{D} = (\mathcal{D}, \mathcal{E}_\mathcal{D})$, where $\mathcal{E}_\mathcal{D}$ represents the edges in the graph, and $t_j \in \mathbb{R}^+$ is the event timestamp. The events occur sequentially in the continuous time interval $[t_1, t_T]$, with $t_1 < t_2 < \cdots < t_T$. The objective is to explore the causal structure graph (or a directed acyclic graph) $\mathcal{G}_\mathcal{V} = (\mathcal{V}, \mathcal{E}_\mathcal{V})$ among event types, grounded in event sequences $\mathcal{X}$ and the topological graph $\mathcal{G}_\mathcal{D}$. $\mathcal{V} = \{v_1, v_2, \dots\}$ indicates the set of event types and $\mathcal{E}_\mathcal{V} = \{e_1, e_2, \dots\}$ indicates the set of directed edges in the causal graph, where each edge $e_1 : v_i \to v_j$ denotes v_i causally influences v_j.

Definition 1 (State space) $\mathcal{S}$. The state space $\mathcal{S}$ represents the set of all possible Directed Acyclic Graphs (DAGs) among event types. When prior causal knowledge A_{prior} is available, state space $\mathcal{S}$ is constructed based on initial state s^0 and A_{prior}. Specifically, causal relationships with $v_{i,j} = 1$ in A_{prior} are integrated to generate different DAGs, forming $\mathcal{S}$.

Definition 2 (Action) $\mathcal{A}$. The action set $\mathcal{A}$ consists of all executable edges that avoid forming cycles. When prior causal knowledge A_{prior} exists, $\mathcal{A}$ is generated based on pairs where $v_{i,j} = 0$ in A_{prior}, allowing different edge connections. The action space is defined as $\mathcal{A} = \{a_{v_1,v_2} : v_1 \to v_2, ..., a_{v_4,v_3} : v_4 \to v_3\}$. Each state s has a corresponding set of executable actions $\mathbb{A}(s) \subseteq \mathcal{A}$ ensuring acyclicity. As shown in Fig. 1, which depicts the causal structures in different states and their corresponding executable actions, blue arrows indicate executable actions, while the red 'X' denotes non-executable actions, such as a_3 in state s^1. It is important to note that, since the optimal causal structure cannot be predefined, there

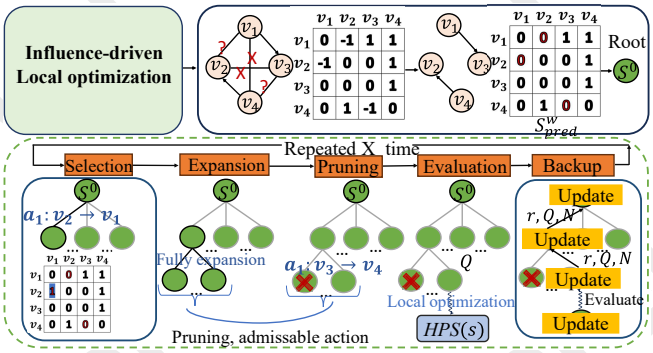


Figure 2: The overview of framework.

is no explicit termination state. Therefore, a termination action $\texttt{Termination}(\hat{a})$ is defined, indicating that the termination state has been reached when executed.

Definition 3 (MCTS Tree). The Monte Carlo Tree Search (MCTS) tree consists of nodes and edges that store different types of information. *Node* n^{l_k}: The k-th node at level l in the search tree is defined as a tuple:

$$n^{l_k} = \{s^{l_k},\ \mathbb{A}(s^{l_k}),\ N(s^{l_k}),\ \text{HPS}(s^{l_k})\} \tag{1}$$

where s^{l_k} represents the current state, $\mathbb{A}(s^{l_k})$ is the set of executable actions including $\texttt{Termination}(\hat{a})$, $N(s^{l_k})$ tracks visit frequency, and the Hierarchical Potential Score (HPS), denoted as $\text{HPS}(s^{l_k})$, quantifies the overall quality of the state. *Edge* ε^{l_k}: The edge information for transitions from node n^{l_k} is defined as:

$$\varepsilon^{l_k} = \{Q(s^{l_k}, a^{l_k}_{v_i,v_j}),\ \Theta^{l_k} = (S^{w,l_k}, \mu^{l_k})\} \tag{2}$$

where $Q(s^{l_k}, a^{l_k}_{v_i,v_j})$ represents the action value for $a^{l_k}_{v_i,v_j} \in \mathbb{A}(s^{l_k})$, and Θ^{l_k} contains the dynamically optimized causal strength matrix $S^{w,l_k} \in \mathbb{R}^{n\times n}$ and baseline intensity vector $\mu^{l_k} \in \mathbb{R}^n$.

Definition 4 (Intensity function) λ. The event occurrence $O^t_{v,d}$ depends on historical event sequences $\mathcal{X}_h = \{(d_i, v_i, t_i)|t_i < t\}$, the state $s \in \mathcal{S}$, and $B_{0:k}$. B_k is a set of binary $|D| \times |D|$ matrices representing physical connections between devices at geodesic distance k, with $B^k_{d_i,d_j} = 1$ indicating a connection. The event intensity $\lambda^{v,d}_t$ for event type v at device d is given by,

$$\lambda^{v,d}_t = \mu^{v,d} + \sum_{j\in Nei(d)\cap Act(t)} \alpha_{v_j,v} s B^k_{d_j,d} \phi(t-t_j)_{v_j,v} \tag{3}$$

where $\mu^{v,d}$ is the baseline intensity, $Nei(d)$ represents the set of neighbors for device d, $Act(t)$ is the set of active event types at time t, and $\alpha_{v_j,v}$ is the impact of event type v_j on event v. $\phi(t-t_j)_{v_j,v} = e^{-\beta(t-t_j)}$ is a time-decaying function that models the influence of event v_j on v at time t. For simplicity, $Nei(d)$ is abbreviated as $N(d)$, and $Act(t)$ will be denoted as $A(t)$, and we define $AN(d,t) = Nei(d) \cap Act(t)$ to represent the active neighbors of d at time t.

4 Our TCCD Approach

In this section, we describe how the TCCD framework integrates MCTS with continuous parameter optimization to jointly optimize the causal structure and parameters. In addition, a detailed algorithmic procedure is provided in Appendix A.

4.1 Overview of Our TCCD Approach

Traditional causal structure learning methods typically treat structure search and parameter optimization as separate tasks. In contrast, we propose the TCCD framework, which innovatively combines Monte Carlo Tree Search (MCTS) with continuous parameter optimization for collaborative optimization. As shown in Fig. 2, TCCD initializes the root node with influence-driven local optimization and employs a tree-based dynamic parameter learning strategy, where parameter estimation guides structure exploration, and structural updates trigger efficient local parameter adjustments. Additionally, TCCD introduces a Hierarchical Potential Score function that combines BIC scores, likelihood values, and global consistency to effectively guide the MCTS search and iterative causal structure optimization.

4.2 Influence-Driven Local Optimization

To efficiently process large-scale telecommunication data, we propose a localized optimization mechanism based on influence propagation. The optimization process consists of two main components: log-likelihood function and local parameter updates.

Log-Likelihood Function. Inspired by S^{2}GCSL [Li *et al.*, 2024], we further design a log-likelihood function that integrates current event occurrences $\mathcal{X}(t_0)$, historical sequences $\mathcal{X}_h = \{(d_i, v_i, t_i)|t_i < t\}$, and topological structure $B_{0:k}$ to derive the weighted Granger causal matrix $S^w_{v_i,v_j} = \alpha_{v_i,v_j} \times s_{v_i,v_j}$ captures the causal influence of event type v_i on v_j through the Granger causality statistic α_{v_i,v_j}. Combined with the baseline intensity matrix μ, these parameters form the set $\Theta = \{S^w, \mu\}$, which we optimize using event sequences $\mathcal{X} = \{(d_i, v_i, t_i)|i \in 1, \ldots, m\}$ through the following likelihood function,

$$\begin{aligned} L(S^w, \mu) &= \sum_d \left(\sum_{i:d_i=d} \log \lambda^{v_i,d}_{t_i} - \sum_v \int_0^T \lambda^{v,d}_t \, dt \right) \\ &= \sum_d \Bigg(\underbrace{\sum_{i:d_i=d} \log\Big(\mu^{v_i,d} + \sum_{j\in AN(d,t)} S^w_{v_j,v_i} B^k_{d_j,d} \phi(t_i - t_j)\Big)}_{\text{Log-Likelihood}} \\ &\quad \underbrace{- \sum_v \sum_{i:t_i=d} \int_{t_i}^T \sum_{j\in N(d)} S^w_{v_j,v_i} B^k_{d_j,d} \phi(s-t_j)\, ds}_{\text{compensation integral term}} \Bigg) \\ &\quad \underbrace{- T \sum_v \mu^{v,d}}_{\text{baseline intensity}} \end{aligned} \tag{4}$$

Here, the log-likelihood represents the occurrence probability of the current event at time t_i, considering both the baseline intensity $\mu_{v_i,d}$ and the weighted influence from neighboring events ($t_j < t_i$ and still active) through the exponential decay function $\phi(t) = e^{-\beta t}$. The baseline intensity term reflects the intrinsic occurrence rate of event type v over $[0, T]$, independent of external influences. The compensation integral term corrects for potential future event impacts by integrating the decay effects and causal dynamics of neighbors $N(n)$ and active events $A(t_i)$. In addition, the steps to deduce the integral term in the log-likelihood function are as follows,

$$\int_{t_i}^{T} \phi(s - t_j)\, ds = \int_{t_i}^{T} e^{-\beta\cdot(s-t_j)}\, ds = e^{\beta t_j} \cdot \left[\frac{e^{-\beta\cdot s}}{-\beta}\right]_{t_i}^{T}$$
$$= e^{\beta t_j} \cdot \left[\frac{e^{-\beta t_i} - e^{-\beta T}}{\beta}\right] = \frac{e^{-\beta\cdot(t_i - t_j)} - e^{-\beta\cdot(T-t_j)}}{\beta} \tag{5}$$

The final optimization problem follows the S^{2}GCSL[Li *et al.*, 2024] and is defined as,

$$S^\star, \mu^\star = \arg\min_{S^w,\mu} -L(S^w, \mu) + \lambda_1 ||S^w||_1 + \lambda_2 h(S^w) \tag{6}$$

where λ_1 and λ_2 are regularization hyperparameters, $||A'||_1$ is the l_1-norm of A', promoting sparsity in the causal matrix to reduce unnecessary connections and improve interpretability, and $h(A')$ ensures that the deduced causal graph is a directed acyclic graph (DAG) [Zheng *et al.*, 2018].

Local Parameter Updates. To effectively combine continuous optimization with search and improve search efficiency, we propose a novel localized optimization mechanism that dynamically updates parameters only for regions affected by each action. For each action $a_{(v_i,v_j)}$, we identify the set of affected nodes $\mathcal{N}_{\text{affected}} = \{v_i, v_j\} \cup \{v_k \mid (v_i, v_k) \in \mathcal{E}_V \text{ or } (v_k, v_j) \in \mathcal{E}_V\}$ that includes both event-type nodes v_i, v_j and their immediate neighbors.

Based on this, we then construct adaptive parameter masks $M_S \in \{0,1\}^{|\mathcal{V}|\times|\mathcal{V}|}$ and $M\mu \in \{0,1\}^{|\mathcal{V}|\times|\mathcal{D}|}$, setting elements to 1 for affected nodes (rows and columns in M_S and corresponding elements in M_μ) and 0 otherwise. The optimization process employs masked gradient updates with local parameters by $S_{\text{local}} = S_{\text{opt}} \odot M_S$ and $\mu_{\text{local}} = \mu_{\text{opt}} \odot M_\mu$, where $\odot$ denotes the Hadamard product. During backpropagation, gradients are constrained to affected nodes through $\nabla S_{\text{opt}} = \nabla S_{\text{opt}} \odot M_S$ and $\nabla \mu_{\text{opt}} = \nabla \mu_{\text{opt}} \odot M_\mu$. Local optimization is performed using Eq.4, tracking the parameter configuration that results in minimal loss. Finally, after the optimization has been completed, only the optimal parameters for the affected nodes are updated to the MCTS nodes. This localized approach significantly reduces computational complexity while maintaining optimization quality by focusing only on relevant parameter subspaces.

4.3 Tree-Guided Continuous-Optimization Casual Discovery

The Monte Carlo Tree Search (MCTS) aims to uncover causal graph structures from observational data by introducing a tree-guided dynamic parameter learning strategy that integrates search and optimization. In the search tree, each node n^{l_k} maintains the visit count $N(s^{l_k})$, tracking the frequency of visits to state s^{l_k}, and the hierarchical potential score HPS(s^{l_k}), which evaluates the overall score of s^{l_k}. For each edge (s^{l_k}, a^{l_k}), assuming a transition from s^{l_k} to $s^{(l+1)_k}$ along (s^{l_k}, a^{l_k}), the action value is defined as $Q(s^{l_k}, a^{l_k}) =$ HPS$(s^{(l+1)_{k'}})$ − HPS(s^{l_k}). The Q-function $Q(s^{l_k}, a^{l_k})$ and parameters $\Theta^{l_k} = (S^{w,l_k}, \mu^{l_k})$ are recorded, where the causal strength matrix $S^{w,l_k} \in \mathbb{R}^{n\times n}$ and baseline intensity vector $\mu^{l_k} \in \mathbb{R}^n$ are dynamically optimized as the tree expands. Furthermore, in each layer of child nodes, we incorporate the action termination identifier $Termination(\hat{a})$ to ensure the optimal node is found.

The search process begins at the root state s^0, progressively selecting and expanding nodes by optimizing the combination of the value function and the exploration reward $u(s)$. Upon reaching a leaf node n', all possible child nodes are expanded based on the executable action set $\mathcal{A}(s')$, where each child node is evaluated by its hierarchical potential score HPS(s_{child}). These child nodes are then ranked and pruned using a pruning mechanism, after which the updated Q-values and node visit counts $N(s)$ are propagated upwards. Rather than starting with an empty root node, the search space is optimized using an influence-driven local optimization method. Specifically, by utilizing the initialized root node parameters $\Theta_{\text{init}} = (S^w_{\text{init}}, \mu_{\text{init}})$, the weight matrix $S^w_{\text{pred}} = \text{softplus}(S^w)$ is derived through global optimization, as prescribed by Eq.4, since the root node state is an empty graph with no masks or influence nodes. The action space is constrained by iteratively removing edges with the smallest weights, where the edges to remove are determined by,

$$e_{\text{remove}(v_i,v_j)} = \arg\min_{(v_i,v_j)\in\mathcal{A}} S^w_{\text{pred}}[v_i, v_j] \tag{7}$$

After this transformation, the resulting binarized matrix is updated to the root node n^0 as the initial state s^0 for the subsequent tree search. This transformation reduces the search space by constraining the action set $\mathcal{A}(s^0)$ starting from s_0. Subsequently, tree-guided continuous optimization is employed to further guide causal discovery, instead of relying solely on the results of continuous optimization. The entire search process consists of five steps, as illustrated in Fig.2.

Selection. Starting from the root node, we employ the Upper Confidence thresholds (UCB) algorithm [Kocsis and Szepesvári, 2006] to select an unexplored node, striking a balance between exploitation and exploration. At each step, an action a is selected to move to a child node, continuing until an unexplored node is found. The selected action a^* is determined by:

$$a^* = \arg\max_{a\in\mathbb{A}(s)} \left[Q(s,a) + C \cdot \frac{\sqrt{N(s)}}{1 + N(s')}\right] \tag{8}$$

where s' is the state after action a^* from state s, $Q(s,a)$ denotes the value of action a, and C controls the exploration-exploitation trade-off.

Expansion. For non-terminal leaf nodes, valid actions are generated as:

$$\mathcal{A}(s) = \{a_{(v_i,v_j)} \mid s[v_i, v_j] = 0 \text{ and } \textit{IsValid}(s \cup \{a_{(v_i,v_j)}\})\} \tag{9}$$

	Wireless-18				Micro-24				Micro-25			
Methods	**Precision**	**Recall**	**F1**	**SHD**	**Precision**	**Recall**	**F1**	**SHD**	**Precision**	**Recall**	**F1**	**SHD**
PC	0.5128	0.2899	0.3704	-	0.2553	0.1752	0.2078	155	0.2921	0.1757	0.2194	150
GES	0.5000	0.2464	0.3301	62	0.2605	0.2263	0.2422	160	0.2078	0.2162	0.2119	179
ADM4	0.3727	0.5942	0.4581	49	0.2671	0.8832	0.4102	135	0.2833	0.9189	0.4331	134
PCMCI	0.3402	0.4783	0.3976	60	0.3333	0.4307	0.3758	126	0.2477	0.3581	0.2928	168
Corl	0.5000	0.5362	0.6167	69	0.2778	0.0365	0.0645	135	0.2632	0.0338	0.0599	147
THP	0.7255	0.5362	0.6167	37	0.4632	0.3212	0.3793	128	0.4118	0.2365	0.3004	140
S^{2}GCSL	0.6364	0.5072	0.5645	46	0.2540	**0.6934**	0.3718	155	0.4786	0.3784	0.4226	**124**
TCCD	**0.7288**	**0.6232**	**0.6719**	**34**	**0.5470**	0.4672	**0.5039**	**115**	**0.4846**	**0.4257**	**0.4532**	132

Table 1: Performance Results on Real-World datasets.

where *IsValid* checks action feasibility. Non-root nodes consider a subset of initial state actions to reduce search space. Each node generates a termination-action, Termination($\hat{a}$), inheriting parent parameters and marked terminal. The selected node s^* is expanded with child nodes n initialized with $Q(s, a) = 0$.

Evaluation. Each new node state $s' = s \cup a_{(v_i,v_j)}$ inherits and local optimizes its parameters Θ of parent to obtain Θ'. The potential score of node is measured by the **H**ierarchical **P**otential **S**core (HPS) function:

$$\text{HPS}(s') = \alpha_1 \cdot \text{BIC}(s') + \alpha_2 \cdot L_{local}(s') + \alpha_3 \cdot L_{global}(s') \tag{10}$$

where α_1, α_2 and α_3 are weights that balance the contributions of BIC, local likelihood and global likelihood in the score function. $L_{local}(s')$ and $L_{global}(s')$ are derived from Eq.4, substituting S^w with $S^w \cdot s'$, and BIC(s') as in THP [Cai *et al.*, 2022]. Following this, the immediate reward is $r(s') = \text{HPS}(s') - \text{HPS}(s)$. Hierarchical normalization standardizes child node rewards for comparability, setting identical rewards to zero to indicate no advantage. For terminal states, the potential score of the parent state is used, and visit statistics are initialized for future decisions.

Pruning. After evaluation, a pruning step is conducted to discard child nodes with lower potential scores. Specifically, a threshold θ is determined as $\theta = \text{sorted}(\text{HPS}(s), \text{reverse=True})[\lfloor k_{\text{threshold}} \times |\text{HPS}(s)| \rfloor]$, where $k_{\text{threshold}}$ represents the proportion of child nodes to retain. This process ensures that only the most promising child nodes, likely to yield high returns, are preserved, thereby further reducing the search space.

Backup. During backpropagation, the visit counts $N(\cdot)$ and Q-values for all nodes along the path from the leaf to the root are updated iteratively. The updates propagate bottom-up, ensuring that the parent nodes integrate the contributions from their children. The immediate reward of child node $r(s')$ is calculated as $\text{HPS}(s') - \text{HPS}(s)$, where it denotes the difference between the HPS of a child node and its parent node. The update rules are as follows: For terminal nodes, the Q-value is set directly to the current reward:

$$Q(s, \hat{a}) \leftarrow r(s) \tag{11}$$

For non-terminal nodes, the Q-value is updated iteratively based on a weighted incremental formula:

$$Q(s, a) \leftarrow Q(s, a) + \frac{\left(\gamma^d Q(s', a) - Q(s, a)\right)}{N(s)} \tag{12}$$

where s represents a node on the path, and after performing action $\hat{a}$, its successor s' (also on the path) becomes the child node; d represents the depth of the node in the tree, γ is the discount factor controlling the influence of rewards from deeper nodes, and $N(s) \leftarrow N(s) + 1$ denotes the visit count of node s. The update process begins at the leaf node and proceeds sequentially along the path up to the root using a breadth-first traversal method. At each step, the cumulative reward is adjusted using the discount factor γ, ensuring that rewards closer to the leaf node contribute more significantly. Furthermore, node expansion is prohibited after the execution of $Termination(\hat{a})$.

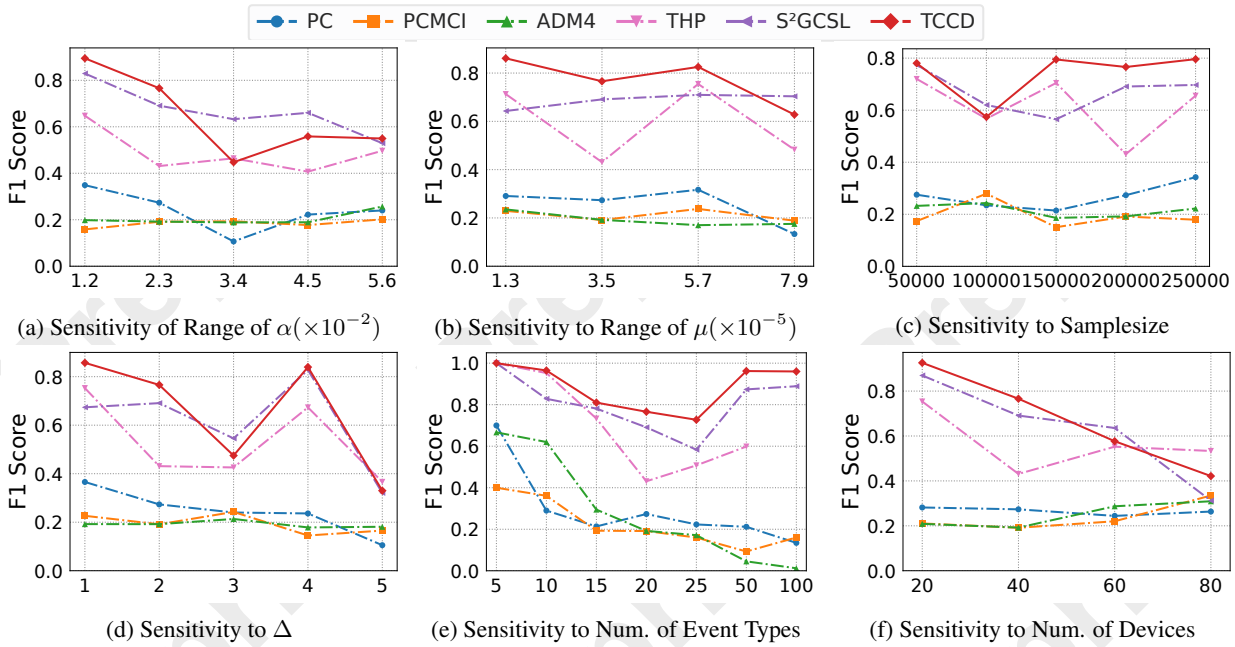
5 Experiments

5.1 Experimental Setup

Evaluation Metrics. We evaluated the performance of our TCCD approach with comparison to baselines using real-world metropolitan cellular network alarm data and synthetic datasets. Similar to [Cai *et al.*, 2022; Li *et al.*, 2024], we test the following metrics: *Precision, Recall, F1 Score, and Structural Hamming Distance (SHD).* Precision measures the proportion of predicted edges that are true among the total predicted edges. Recall quantifies the fraction of actual edges that have been correctly predicted. F1 Score represents the harmonic mean of Precision and Recall, calculated as $F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$. SHD denotes the number of edge insertions, deletions, or flips needed to transform a graph into another graph. Also, the detailed implementation and hyperparameters of our experiments are provided in Appendix B.1.

Real-World Metropolitan Cellular Network Alarm Data. We evaluate TCCD on three challenging datasets derived from real-world metropolitan cellular networks, showcased in the PCIC 2021 competition[1]. The first dataset, 18V_55N_Wireless (Wireless-18), consists of 18 alarm types, 55 topological network elements, and a total of 34,839

[1]https://competition.huaweicloud.com/information/1000041487/dataset



(a) Sensitivity of Range of $\alpha(\times 10^{-2})$ (b) Sensitivity to Range of $\mu(\times 10^{-5})$ (c) Sensitivity to Samplesize

(d) Sensitivity to Δ (e) Sensitivity to Num. of Event Types (f) Sensitivity to Num. of Devices

Figure 3: Sensitivity analysis of different methods with respect to various parameters.

alarm events. The other two datasets, 24V_439N_Microwave (Micro-24) and 25V_474N_Microwave (Micro-25), contain 24 and 25 alarm types, 439 and 474 topological network elements, and 64,599 and 48,573 alarm events, respectively. Remarkably, the Wireless-18 lacks an underlying topological network structure, which demonstrates the applicability of our model. Causal discovery on such datasets aims to uncover causal relationships among alarm types, further validating the effectiveness and generalization capability of our method. Detailed statistics of the datasets are shown in Appendix B.2.

Synthetic Datasets. In addtion, we generate a range of synthetic datasets via the gcatle [Zhang *et al.*, 2021] API, which simulates event sequences using topological Hawkes process [Cai *et al.*, 2022]. The details of the synthetic data generation process and the parameter configurations are provided in Appendix B.3.

Baselines. In experiments, we selected the following causal discovery methods as baselines for comparison: GES [Chickering, 2002], ADM4 [Zhou *et al.*, 2013], PC [Spirtes *et al.*, 2001], PCMCI [Runge *et al.*, 2019], Corl [Wang *et al.*, 2021], THP [Cai *et al.*, 2022] and S^{2}GCSL [Li *et al.*, 2024]. For fairness, we used the optimal parameters as specified in their respective papers (Appendix C). We ran all experiments on a platform with Ubuntu 16.04 on 256GB of memory and an NVIDIA GeForce GTX 1080 Ti GPU.

5.2 Experimental Results

Results on Real-World datasets. Based on the experimental results in Table 1, TCCD consistently outperforms baselines across three real-world datasets. Methods utilizing topological network structures (THP, S^{2}GCSL, and TCCD) achieve higher F1 scores than traditional point process-based approaches (PC, GES, PCMCI), highlighting the importance of spatiotemporal modeling. TCCD outperforms all baselines with the highest F1 score, demonstrating a strong balance between Precision and Recall. On the Wireless-18 dataset, it achieves an F1 of 0.6719, with high Precision 0.7288, Recall 0.6232, and the lowest SHD 34, ensuring accurate causal graph reconstruction. For Micro-24, TCCD maintains optimal performance with the highest F1 0.5039 and Precision 0.5470, alongside the lowest SHD 115, despite higher Recall 0.6934 of S^{2}GCSL. In networks with 25 event types, TCCD

Datasets	Methods	Pre	Recall	F1	SHD
Wireless-18	TCCD$^{-\text{MCTS}}$	0.7273	0.5797	0.6452	37
	TCCD	**0.7288**	**0.6232**	**0.6719**	**34**
Micro-24	TCCD$^{-\text{MCTS}}$	0.5556	0.4380	0.4898	115
	TCCD	0.5470	**0.4672**	**0.5039**	115
Micro-25	TCCD$^{-\text{MCTS}}$	0.4841	0.4122	0.4453	132
	TCCD	**0.4846**	**0.4257**	**0.4532**	132

Table 2: Ablation study on three real-world datasets.

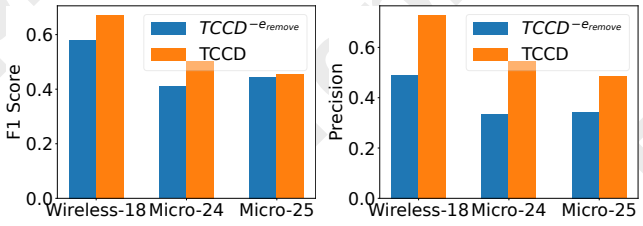
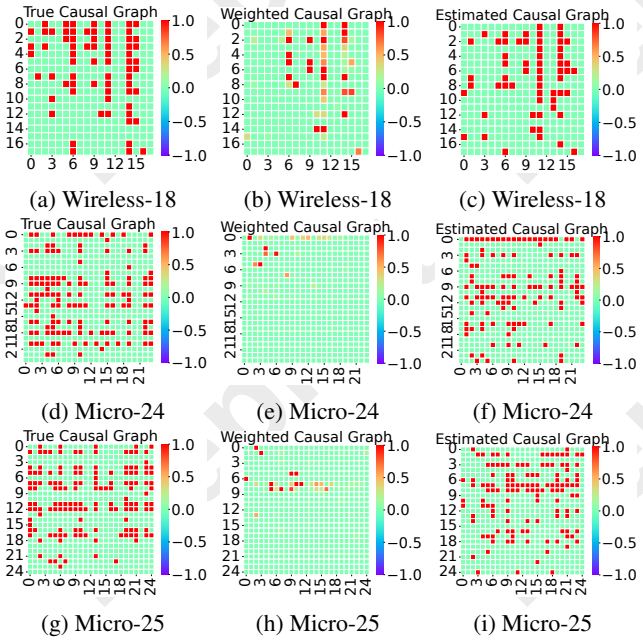


Figure 4: Performance of TCCD and TCCD$^{-e_{\text{remove}}}$ on three datasets

leads with an F1 of 0.4516 and Precision of 0.4809, though SHD is slightly higher than S^{2}GCSL (132 vs. 124). Traditional methods (PC, GES) perform poorly (F1 < 0.37), and ADM4 show high Recall but low Precision, indicating oversensitivity to weak causal signals. In general, TCCD effectively balances inference accuracy and structural correctness, establishing itself as a robust solution for large-scale causal discovery in telecommunication networks.



(a) Wireless-18 (b) Wireless-18 (c) Wireless-18

(d) Micro-24 (e) Micro-24 (f) Micro-24

(g) Micro-25 (h) Micro-25 (i) Micro-25

Figure 5: Experimental results of three cases.

Results on Synthetic Datasets. We further evaluate the performance of TCCD and five baseline models across six dimensions on the synthetic datasets (Fig.3). Fig.3 shows the F1 scores of different methods under various parameter settings (α, μ, δ), with additional evaluation metrics provided in Appendix E. Overall, our TCCD performs well in most cases, achieving relatively high F1 scores across multiple parameter configurations. In particular, it demonstrates strong stability and generalization ability when the sample size is large or the number of event types increases. Although THP [Cai *et al.*, 2022] and S^{2}GCSL [Li *et al.*, 2024] achieve slightly higher F1 scores under certain specific parameter settings, their overall performance is less stable compared to TCCD. Notably, our method maintains superior generalization ability and causal structure learning performance, especially in high-dimensional and complex event sequences.

Ablation Study. To validate the effectiveness of the proposed TCCD framework, we conduct ablation experiments on three real-world datasets. Table 2 presents the comparison experimental results between TCCD and its variant TCCD$^{-\text{MCTS}}$, where TCCD$^{-\text{MCTS}}$ denotes a simplified version of TCCD that only uses influence-driven local optimization to obtain the initial causal structure s^0 without the MCTS-based structure refinement. Additionally, Fig.4 shows the comparison results for another variant TCCD$^{-e_{\text{remove}}}$, which removes the mechanism of eliminating invalid edges from TCCD to assess the impact of this mechanism. As shown in Table 2, one the Wireless-18 dataset, TCCD outperforms TCCD$^{-\text{MCTS}}$ in Precision, Recall, and F1, with Recall increasing by 4.35% and F1 by 2.67%. On the more challenging Micro-24 and Micro-25 datasets , while overall performance decreases, TCCD consistently surpasses TCCD$^{-\text{MCTS}}$ in Recall and F1. In Micro-24, Recall improves by 2.92% and F1 by 2.41%, despite a slight drop in Precision, indicating better causal recovery. In Micro-25, Precision slightly increases, with Recall and F1 improving by 3.29% and 1.78%, respectively. Moreover, the identical SHD values for TCCD and TCCD$^{-\text{MCTS}}$ suggest that MCTS enhances structural weights rather than altering causal edge count. Even TCCD$^{-\text{MCTS}}$ outperforms the baseline S^{2}GCSL (Table 1), validating the effectiveness of our approach. The superior performance of TCCD can be attributed to two key factors: First, unlike S^{2}GCSL, which primarily rely on statistical correlations, our approach integrates network topology and temporal decay to capture spatiotemporal causality; Second, MCTS enables a more thorough exploration of causal structures, avoiding local optima and improving inference accuracy. Additionally, as shown in Fig.4, removing invalid edges in TCCD$^{-e_{\text{remove}}}$ significantly boosts F1 on Wireless-18, highlighting its role in mitigating erroneous causal links. On Micro-24 and Micro-25, TCCD$^{-e_{\text{remove}}}$ remains competitive, further confirming the contributions of each module in TCCD.

Case Study. Going deeper into the analysis, we evaluate TCCD on three real-world causal networks with $\mathcal{N} = \{18, 24, 25\}$ using True Causal Graph, Weighted Causal Graph (S^w), and Estimated Causal Graph. Fig.5(a)(b)(c) show how the method identifies local structures in the 18-event network by capturing causal relationships $1 \to 3$, $16 \to 14$, and $17 \to 6$ with high fidelity through locality-aware optimization based on influence propagation. For the Micro-24, Fig.5(d)(e)(f) demonstrate effective identification of dense regions (e.g., 8-13) while balancing detailed structural recovery with global sparsity via MCTS. Fig.5(g)(h)(i) reveal the robustness of the method in handling 25 event types. In particular, Fig.5(i) more accurately captures the key causal relationships around nodes 4-12 while reducing false positives in the sparse regions around nodes 20-24. Ultimately, the TCCD framework integrates MCTS with continuous optimization to build accurate causal structures with precision across diverse network scales.

6 Final Remarks

In this paper, we propose TCCD, a framework that combines MCTS with continuous optimization for efficient causal structure learning in large-scale telecommunication networks. By utilizing MCTS's exploration-exploitation strategy and gradient-based optimization, TCCD balances global search and local refinement. It extends the S^{2}GCSL likelihood by integrating discrete DAG search with continuous learning, enabling dynamic feedback to enhance generalization. Future work includes scaling MCTS to larger search spaces via logical action representations [Zhuo *et al.*, 2011; Jin *et al.*, 2022], and adapting the framework to diverse event patterns and temporal dynamics.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 62076263).

References

[Achab *et al.*, 2018] Massil Achab, Emmanuel Bacry, Stéphane Gaïffas, Iacopo Mastromatteo, and Jean-François Muzy. Uncovering causality from multivariate hawkes integrated cumulants. *Journal of Machine Learning Research*, 18(192):1–28, 2018.

[Bhattacharjya *et al.*, 2022] Debarun Bhattacharjya, Karthikeyan Shanmugam, Tian Gao, and Dharmashankar Subramanian. Process independence testing in proximal graphical event models. In *Conference on Causal Learning and Reasoning*, pages 144–161. PMLR, 2022.

[Cai *et al.*, 2022] Ruichu Cai, Siyu Wu, Jie Qiao, Zhifeng Hao, Keli Zhang, and Xi Zhang. Thps: Topological hawkes processes for learning causal structure on event sequences. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[Cheng *et al.*, 2023] Yuxiao Cheng, Runzhao Yang, Tingxiong Xiao, Zongren Li, Jinli Suo, Kunlun He, and Qionghai Dai. Cuts: Neural causal discovery from irregular time-series data. *arXiv preprint arXiv:2302.07458*, 2023.

[Chickering, 1996] David Maxwell Chickering. Learning bayesian networks is np-complete. *Learning from data: Artificial intelligence and statistics V*, pages 121–130, 1996.

[Chickering, 2002] David Maxwell Chickering. Optimal structure identification with greedy search. *Journal of machine learning research*, 3(Nov):507–554, 2002.

[Idé *et al.*, 2021] Tsuyoshi Idé, Georgios Kollias, Dzung Phan, and Naoki Abe. Cardinality-regularized hawkes-granger model. *Advances in Neural Information Processing Systems*, 34:2682–2694, 2021.

[Itô, 1971] Kiyosi Itô. Poisson point processes attached to markov processes. In *Proc. 6th Berk. Symp. Math. Stat. Prob*, volume 3, pages 225–240, 1971.

[Jin *et al.*, 2022] Kebing Jin, Hankz Hankui Zhuo, Zhanhao Xiao, Hai Wan, and Subbarao Kambhampati. Gradient-based mixed planning with symbolic and numeric action parameters. *Artif. Intell.*, 313:103789, 2022.

[Kocsis and Szepesvári, 2006] Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In *European conference on machine learning*, pages 282–293. Springer, 2006.

[Lachapelle *et al.*, 2019] Sébastien Lachapelle, Philippe Brouillard, Tristan Deleu, and Simon Lacoste-Julien. Gradient-based neural dag learning. *arXiv preprint arXiv:1906.02226*, 2019.

[Li *et al.*, 2022] Wenqian Li, Yinchuan Li, Shengyu Zhu, Yunfeng Shao, Jianye Hao, and Yan Pang. Gflowcausal: Generative flow networks for causal discovery. *arXiv preprint arXiv:2210.08185*, 2022.

[Li *et al.*, 2024] Mingjia Li, Shuo Liu, Hong Qian, and Aimin Zhou. A simple yet scalable granger causal structural learning approach for topological event sequences. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024.

[Liniger, 2009] Thomas Liniger. *Multivariate hawkes processes*. PhD thesis, ETH Zurich, 2009.

[Liu *et al.*, 2024] Yuequn Liu, Ruichu Cai, Wei Chen, Jie Qiao, Yuguang Yan, Zijian Li, Keli Zhang, and Zhifeng Hao. Tnpar: Topological neural poisson auto-regressive model for learning granger causal structure from event sequences. In *AAAI*, pages 20491–20499, 2024.

[Qiao *et al.*, 2023] Jie Qiao, Ruichu Cai, Siyu Wu, Yu Xiang, Keli Zhang, and Zhifeng Hao. Structural hawkes processes for learning causal structure from discrete-time event sequences. *arXiv preprint arXiv:2305.05986*, 2023.

[Runge *et al.*, 2019] Jakob Runge, Peer Nowack, Marlene Kretschmer, Seth Flaxman, and Dino Sejdinovic. Detecting and quantifying causal associations in large nonlinear time series datasets. *Science advances*, 5(11):eaau4996, 2019.

[Shojaie and Fox, 2022] Ali Shojaie and Emily B Fox. Granger causality: A review and recent advances. *Annual Review of Statistics and Its Application*, 9(1):289–319, 2022.

[Spirtes and Glymour, 1991] Peter Spirtes and Clark Glymour. An algorithm for fast recovery of sparse causal graphs. *Social science computer review*, 9(1):62–72, 1991.

[Spirtes *et al.*, 2000] Pater Spirtes, Clark Glymour, Richard Scheines, Stuart Kauffman, Valerio Aimale, and Frank Wimberly. Constructing bayesian network models of gene expression networks from microarray data. 2000.

[Spirtes *et al.*, 2001] Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, prediction, and search*. 2001.

[Tsamardinos *et al.*, 2006] Ioannis Tsamardinos, Laura E Brown, and Constantin F Aliferis. The max-min hill-climbing bayesian network structure learning algorithm. *Machine learning*, 65:31–78, 2006.

[Wang *et al.*, 2021] Xiaoqiang Wang, Yali Du, Shengyu Zhu, Liangjun Ke, Zhitang Chen, Jianye Hao, and Jun Wang. Ordering-based causal discovery with reinforcement learning. *arXiv preprint arXiv:2105.06631*, 2021.

[Xu *et al.*, 2016] Hongteng Xu, Mehrdad Farajtabar, and Hongyuan Zha. Learning granger causality for hawkes processes. In *International conference on machine learning*, pages 1717–1726. PMLR, 2016.

[Zhang *et al.*, 2020] Wei Zhang, Thomas Panum, Somesh Jha, Prasad Chalasani, and David Page. Cause: Learning granger causality from event sequences using attribution methods. In *International Conference on Machine Learning*, pages 11235–11245. PMLR, 2020.

[Zhang *et al.*, 2021] Keli Zhang, Shengyu Zhu, Marcus Kalander, Ignavier Ng, Junjian Ye, Zhitang Chen, and Lujia

Pan. gcastle: A python toolbox for causal discovery. *arXiv preprint arXiv:2111.15155*, 2021.

[Zheng *et al.*, 2018] Xun Zheng, Bryon Aragam, Pradeep K Ravikumar, and Eric P Xing. Dags with no tears: Continuous optimization for structure learning. In *NeurIPS*, 2018.

[Zhou *et al.*, 2013] Ke Zhou, Hongyuan Zha, and Le Song. Learning social infectivity in sparse low-rank networks using multi-dimensional hawkes processes. In *Artificial intelligence and statistics*, pages 641–649, 2013.

[Zhu *et al.*, 2019] Shengyu Zhu, Ignavier Ng, and Zhitang Chen. Causal discovery with reinforcement learning. *arXiv preprint arXiv:1906.04477*, 2019.

[Zhuo *et al.*, 2011] Hankz Hankui Zhuo, Qiang Yang, Rong Pan, and Lei Li. Cross-domain action-model acquisition for planning via web search. In *ICAPS*, 2011.