

VAC²: Visual Analysis of Combined Causality in Event Sequences

Sujia Zhu, Guodao Sun, Yue Shen, Zihao Zhu, Wang Xia, Baofeng Chang,
Jingwei Tang, and Ronghua Liang

Abstract—Identifying causality behind complex systems plays a significant role in different domains, such as decision-making, policy implementations, and management recommendations. However, existing causality studies on temporal event sequences data mainly focus on individual causal discovery, which is incapable of exploiting combined causality. To fill the absence of combined causality discovery on temporal event sequence data, eliminating and recruiting principles are defined to balance the effectiveness and controllability of cause combinations. We also leverage the Granger causality algorithm based on the Reactive point processes to describe impelling or inhibiting behavior patterns among entities. In addition, we design an informative and aesthetic visual metaphor of “electrocircuit” to encode aggregated causality for ensuring that our causality visualization exhibits no node-overlap, no edge-intersection, and no link-ambiguity. Aggregation layout, diverse sorting strategies, and smooth interactions are also integrated into our directed, weighted, and parallel-based hypergraph for illustrating combined causality. Our developed combined causality visual analysis system, namely VAC², can help users effectively explore combined causes as well as individual causes. This interactive system supports multi-level causality exploration with diverse ordering strategies and a focus and context technique to help users obtain different levels of information abstraction. The usefulness and effectiveness of our work are further evaluated by conducting two case studies and a controlled user study on event sequence data.

Index Terms—Causal discovery, cause combination, directed hypergraph visualization, impelling and inhibiting behaviors, event sequence data

1 INTRODUCTION

CUSAL discovery mainly focuses on revealing deep and strong correlations among entities. Human activities form a large number of event sequences, where potential causal relations exist among these events. Discovering causality is a ubiquitous and crucial task in many applications such as medical treatment, marketing research, products recommendation, and policy implementation.

In recent years, researchers have been devoting efforts to analyze causal effects [1] and causality problems [2]. Most causality researchers focus on the individual causality discovery and pay little attention to combined causality discovery. An effect is commonly caused by confounding mutual impacts rather than a single impact. These multifaceted entities are deemed as a combined cause. A cause combination consisting of two or more impact entities is a cause of another effect entity, each entity individually might not be a cause. For example, three factors of ignition point, oxygen, and combustible matter can result in flame together. In fact, neither of these factors can generate flame individually. Thus, individual causality discovery, under an assumption of neglecting all the other necessary factors, may lead to misleading that every factor be an independent cause of flame.

Existing methods [3] for combined causality discovery are confined to multiple dimensional dataset and are not adaptable for inferring causal relationships among successive time-stamped

events in temporal event sequences. We would like to fill this gap of combined causality discovery for event sequences, but there are obstacles for modeling and analyzing combined causality. The first challenge is limited consideration when modeling combined causality for event sequences. Existing causal research [4] primarily focusing on the existence of causality, while ignoring deeper investigations into fine-grained causality, specifically whether effects are impelling or inhibiting. An impelling or inhibiting impact signifies that the occurrence of one event decreases or increases the probability of another event, respectively. In fact, elements in a cause combination may mutually affected by the others, resisting or promoting relationships. However, the lack of comprehensive causality of these complex intricate interactions in events make combined causality discovery a particularly challenging problem. The second challenge is visualizing directed hypergraph that obtained by our combined causality algorithm. For a directed combined causality, a specific entity may be included in other causes’ sets, and elements in these sets may exhibit intercausal relationships. As a result, visual clutter and ambiguity always occur, leading to excessive nodes overlapping and hyperedge crossings.

To address the first challenge, we proposed a fine-grained causality discovery algorithm for further construct combined causality model. Concretely, we recover the Granger causality among events across a set of event sequences based on Reactive point processes [5] [6] modeling, which reveals how the historical occurrence of an event increases or decreases the probability of current events. Then, filtering rules are defined for reducing the number of candidate cause combinations. The rules are strategically achieved by abandoning redundant combination entries and recruiting the useful combination entries. To deal with the second challenge, we design an informative and aesthetic visual

• Sujia Zhu, Guodao Sun, Yue Shen, Zihao Zhu, Wang Xia, Baofeng Chang, and Ronghua Liang are with the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China. Guodao Sun is the corresponding author.

E-mail: {zhusujia, shenyueeee98}@gmail.com, {guodao, xiawang, baofeng.chang, jinrhlhang}@zjut.edu.cn, iszhuzh@outlook.com.

metaphor of “electrocircuit” to encode and aggregate “AND” and “OR” causes for ensuring no node-overlap, no edge-intersection, and no link-ambiguity. We also integrate aggregation layout, diverse sorting strategies, a focus and context approach, and smooth interaction techniques into our parallel-based technique for illustrating combined causality. Based on above combined causal discovery and visual design, we design a causality analysis system for exploring causalities in event sequences. The system consists of three panels. In configuration panel, users can set parameters to filter significant causalities. In combined causality panel, a parallel-based hypergraph is designed for analyzing causality across multiple levels of information abstraction. In addition, characteristics of entities in statistic component help users further understand and confirm causalities. Influence propagations are tracked in propagation panel. VAC² is provided as an open source project in Github¹. Then we demonstrate the usefulness and effectiveness of our work by conducting two case studies and a controlled user study on our work and visual design.

To summarize, our key contributions include:

- We develop an fine-grained causal discovery algorithm, which reveals impelling and inhibiting influence among entities for the event sequence datasets.
- We propose a novel directed hypergraph visualization, which incorporates a causal discovery mechanism to detect combined causes among entities.
- We design VAC², a causality analysis system to help interpret causalities, and provide causal facts through case studies and user evaluation.

2 BACKGROUND AND RELATED WORK

Our work aims to visually analyze both individual and combined causes of events in the event sequences. Thus, the related works are concerned about causal discovery and visual analysis of causality.

2.1 Causal discovery

Causality reveals profound cause-effect relationships that traditional associations cannot accurately define. Fortunately, substantial fundamental research [1] and comprehensive reviews [2] [7] have been conducted for causality analysis. In this paper, we subsume the causal discovery algorithm into two types, namely, individual cause discovery and combined causes discovery.

For individual causality discovery [1], existing approaches can be divided into constraint-based, score-based, and hybrid algorithms. Constraint-based approaches, including PC [8], SGS [9], FCI [10], and their extended algorithms, mainly implemented conditional independence (CI) tests to delete the edges between independent variables in the hypothetical and fully connected graph. However, these approaches cannot quantify the causal strength between variables. Score-based approaches, including BIC score, GES [11] and F-GES [12], introduced structural equations or kernel functions to quantify causal strength between causes and effects. However, these approaches commonly involve many parameters, which have significant impacts on causal results during the optimization process. Hybrid approaches, including ARGES [13] and MMHC [14], combine constraint-based and score-based approaches by their consistency results to achieve causal discovery. However, when two results obtained by constraint-based and score-based approaches are in conflict, making a correct decision is challenging.

Many causality discovery approaches are implemented for time series data, such as information theory-based method [15], Dynamic Bayesian networks [16], CI test [17], and Granger causality test-based approaches [10]. There exist some surveys [7] [18] focus on causal discovery for time series data. The reviews introduce the major concept of causal discovery algorithm, implement experiments on representative methods, compare the performance metrics, and summarize the characteristics of different types of approaches. Among these approaches, Granger causality test is a classical and widely adopted method for temporal properties dataset in various domains, such as medicine [19], economic [20], environmental protection [21], and geography [22]. The intervention concept is introduced to determine the causal relations via judging whether a certain variable changed after another's intervention. Jin et al. [4] proposed a user-feedback mechanism to enhance performance of automatic causality analysis by utilizing Granger causality on the hawkes processes [23] for the event sequences. However, the above approaches cannot uncover fine-grained causality, such as impelling and inhibiting impact among variables or entities.

For combined causes discovery, existing causality studies mostly focus on individual cause discovery, and little attention is paid to combined causes discovery [3]. Combined causes can be identified by performing types of individual cause discovery, such as traditional CI test [17]. However, this types of approaches pose a challenge of low efficiency because of combinatorial explosion. Many researchers proposed approaches to improve the efficiency of causality discovery by preprocessing combined causes, such as a novel model [3], an efficient algorithm [17] [24] [25], and frequent itemset mining algorithms [26]. First, combined causes rules or principles [24] are defined via leveraging association and partial association based on the idea of causal relationships being persistent. Then, many researchers have proposed types of approaches to trade-off complexity of computation and accuracy. An additive noise model [3] is introduced to subsume all possible combined causes into three categories, give them formal and complete definitions, and detect combined causes. A multi-level strategy [25] is proposed to efficiently identify and validate potential combined causes by utilizing an efficient association mining approach in observational data. Ma et al. [17] also proposed a multi-level approach to detect the combined causes via HITON-PC algorithm, an efficient and commonly used local causal discovery method by CI test, when given a target variable. CCCRUD [26] and DCCRUD [27] are proposed by employing frequent itemset mining algorithms, to identify disjunctive combined causal rules and conjunctive combined causal rules from uncertain data. However, existing algorithms for combined causal discovery are primarily validated through qualitative assessments, significantly lacking quantitative evaluations due to the absence of ground-truth data.

Our work aims to analyze combined causes among entities in temporal event sequences. However, current approaches for combined causes discovery, designed for multiple dimensional datasets, which are not applicable for inferring causalities among successive time-stamped events in temporal event sequences. In addition, current causal discovery algorithm for event sequences, relies on the point processes, has been already quite complex and time-consuming. The data transformation approach may make subsequent computations more intricate and protracted. Thus, we extended causality model [4] by leveraging the Reactive point process [5] to fit in event sequences.

¹<https://zjutvis.github.io/VAC2/>

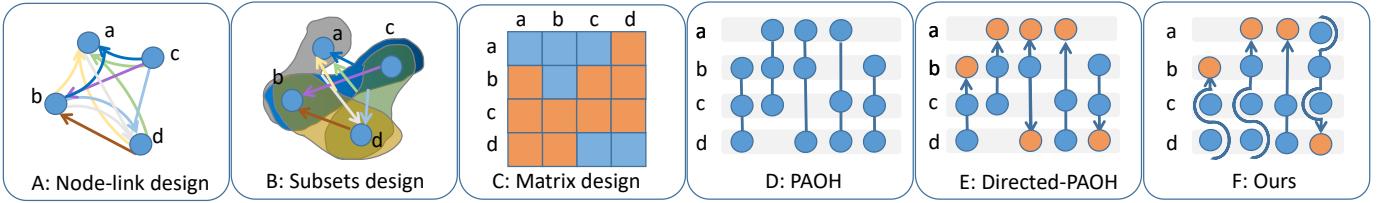


Fig. 1: Diverse techniques for a directed hypergraph visualization under the same data.

2.2 Visual Causality Analysis

Researchers investigate user perception and cognition in visual causality analysis [28] [29]. They contribute experiments [30] demonstrating an approach to offer design guidelines for visual encodings of causality [31].

Causal relations are commonly illustrated with a directed acyclic graph, where a node serves for a type of entities and a link with an arrow encodes the causal relation and direction between two kinds of entities, such as Causalflow [32] Compass [21], and DOMINO [33] for time series dataset, Questionnaire Explorer [34] for questionnaire data, and Outcome-explorer [35], Silva [36], Causality Explorer [37] and the Visual Causality Analyst [38] for multiple dimensional datasets. The above studies focus on visual analysis of simple graphs representing one-to-one causal relationships. In contrast, our investigation involves many-to-one causal relationships within a directed hypergraph, where multiple nodes collectively point to a terminal node. When attempting to depict our proposed many-to-one combined causality using traditional graphs, information is lost and ambiguities arise, one of which is that multiple combined causes might be mistakenly perceived as independent causes.

A hypergraph [39] [40] is a generalization of a traditional one-to-one-relationship graph where hyperedges denote entity relations and each edge set shares the common source entities or target entities. Many surveys [41] [42] and researchers [43] [44] on hypergraph visualization exist, where hypergraphs are usually visualized as node-link-based representations, subset-based hypergraph, matrix-based hypergraph, and parallel-based hypergraph. (1). Standard node-link-based representations are the basic and commonly used approaches [40]. Types of layout algorithms [41] [44] [45] are usually proposed to relax visual clutter and hyperedges intersections. (2). Subsets-based hypergraph drawing approaches [39] [44] are designed to address overlap issues of hyperedges. Hyperedges are illustrated with closed curves enveloping their vertices. (3). Matrix-based hypergraph drawing [46] methods can intuitively illustrate the evolution of relationships between sets and elements with specific matrix layout algorithm, such as Upset [47]. (4). Parallel-based hypergraph drawing approaches [48] depict vertices as equally spaced parallel horizontal lines, and edges as vertical lines. This type of hypergraph visualization can be able to address the visual clutter issues but leading to nodes repetition.

A directed hypergraph [49] is an inherent extension of a standard hypergraph, as well as a hypergraph is a generation of a generic graph. Types of applications [50] are implemented to help users effectively and intuitively obtain the correlations among entities in communication [51] [52], data mining [53], security [54], transportation [55] [56], and semantic networks [57]. A directed hypergraph can be illustrated by adding an additional visual channel into a conventional visualization of undirected hypergraph, which are often portrayed as standard or bipartite networks. For example, under the same data ($c \rightarrow b$; $d \rightarrow b$; $\{b, c\} \rightarrow a$; $\{b, d\} \rightarrow a$; $\{c, d\} \rightarrow a$; $\{a, b\} \rightarrow d$; $\{b, c\} \rightarrow d$), Fig. 1 illustrates diverse

visual designs and Table. 1 summarizes the comparison results of these visual design techniques.

- Fig. 1(A) shows a traditional directed hypergraph visualization. Hyperedges are bounded and terminate at one end node, while multiple edges originate from various source nodes target the end nodes, respectively [58] [59]. However, low scalability and too much edge crossing are key challenges for the edge-bounded approach.
- Fig. 1(B) displays a subsets-based design. Concave boundaries are adopted for drawing set regions. This visual design does not scale well for a larger number of sets intersect because the reused elements may lead to set shapes and overlaps quite complex.
- Fig. 1(C) illustrates a matrix-based approach, which is essentially an extension of the ordinary matrix diagram by adding a visual channel to encode the direction of correlations. However, the pixels representing hyperedges are reused, potentially resulting in ambiguity.
- Fig. 1(D) offers a parallel-based approach, *PAOH*, similar to our design. In contrast, *PAOH* is designed for dynamic and undirected hypergraph while our design aims at static and directed hypergraph. Other strategies (e.g., layout aggregations, sorting strategies, and interactive technique) in *PAOH* are also different and unavailable for our directed hypergraph.
- Fig. 1(E) depicts a feasible design for visualizing directed hypergraph by adding a visual channel to *PAOH*. However, starting and ending nodes can be confusing if incorporating the symbol of direction to *PAOH*. The elements reused in sets may lead to misleading information and visual ambiguity.
- Fig. 1(F) shows our proposed parallel-based design, where “electrocircuit” visual metaphor is embedded to encode and aggregate “AND” and “OR” causes. This visual design has no node-overlap, no edge-intersection, and no link-ambiguity on hypergraph. In addition, the aggregation strategies and the focus and context technique guarantee high scalability. But this design concurrently introduces low readability.

These visual designs are compared across several dimensions: the type of input data, key characteristics of technique, ambiguity, edge-intersection, scalability, and readability. The results of comparing diverse hypergraph visual designs are summarized in Table 1. Our parallel-based visualization is designed for a directed hypergraph. This visual design achieves a hypergraph with no node-overlap, no edge-intersection, and no link-ambiguity, showing a high scalability. The design suffers from low readability, which may result in a steep learning curve.

TABLE 1: Comparison results of directed hypergraph visualization.

Techniques	Directed	Aggregated	Ambiguity	Edge-intersection	Scalability	Readability
A:Node-link	■	□	□	■	Low	High
B:Subset design	■	□	■	■	Low	High
C:Matrix design	■	□	■	■	High	High
D:PAOH	□	■	■	□	High	Low
E:Directed-PAOH	■	■	■	□	High	Low
F:Ours	■	■	□	□	High	Low

■: Existence □: Non-Existence

3 TASK ANALYSIS AND PIPELINE

3.1 Task Analysis

This paper strives to visually analyze the individual and combined causality from the event sequence dataset. We summarize the significant analysis tasks in domain literatures as follows:

- T1 **Discovering the combined causality.** A cause combination consists of two or more elements and each of them cannot be an independent cause. These elements in cause combinations interact with each other and have causal relations with other entities. The causality of combination sets is demanding [4] and can be urgent in situations such as commerce [26].
- T2 **Detecting fine-grained causality.** Elements in a cause combination may interact with each other and play an impelling or inhibiting role in causality. Thus, impelling or inhibiting influence are desired to accurately express the features of causality between entities.
- T3 **Trading-off controllability and effectiveness on cause combinations.** Computing all possible combined causes is infeasible due to the exponential number of combined causes. Thus, diminishing redundant cause combination entries and retaining valid cause combination entries are indispensable [17].
- T4 **Providing an informative and intuitive visual metaphor.** Hyperedges crossing and nodes overlapping may create visual clutter and visual ambiguity [60], which can hinder users from exploring and investigating pattern recognition. Diminishing visual ambiguity and relaxing visual clutter are desired, such as providing an informative and pleasing visual metaphor for the combined causality.
- T5 **Highlighting and displaying causality patterns.** Commonly encountered tasks for graph analysis [61] is useful for our combined causality, a directed hypergraph. Significant causality patterns may be hidden in a complex causality. Exploring a complex causality is time-consuming. Thus, diverse causal patterns must be illustrated and highlighted to help users significantly improve exploration efficiency [32] [37].
- T6 **Exploring causal relations at a multilevel perspective.** Users may need to involve different levels of concepts, i.e., the overall trend of causality and focused causality, which is a complex scenario. Hence, a hierarchical exploration is a common approach for ensuring scalability in visual analysis [62] [63], which can help users organize their analysis and improve exploration efficiency.

These requirements help us obtain appropriate visual design principles and make informed decisions about our visual design.

3.2 Model Pipeline and User Workflow

We proposed a combined causal discovery algorithm for detecting impelling and inhibiting behaviors, two influencing factors of causality shown in Fig. 3(B). After extracting useful entities and filtering valid event sequences from the raw dataset, Reactive point processes is employed to model Granger causality for inferring the impelling or inhibiting behaviors in temporal event sequences. In this process, eliminating and recruiting principles on candidate cause combinations are defined for ensuring the effectiveness and controllability on cause combinations. Based on the above operations, we obtain a directed and weighted hypergraph, which uncovers the impelling and inhibiting behaviors among entities. To visually analyze causality, we design VAC², an interactive causality analysis system shown in Fig. 3(C) for supporting the user analysis task (T1-T6) with the system architecture and user workflow

featured in Fig. 3(A). The system shown in Fig. 6 includes three panels, namely, combined causality panel, propagation panel, and parameters configuration panel. A user can start with an overview of a novel, directed, weighted, and parallel-based hypergraph. Then causal patterns are presented in a parallel-based visualization with various ordering operations and aggregation strategies. Flexible interactions, such as filtering, brushing, and selecting focused causality, are provided to explore causalities in multiple levels of abstraction. Next, modifying incorrect causalities, analyzing causality patterns, and tracking causality propagations are also supported by various visual clues, ordering operations, and smooth interactions.

4 CAUSALITY MODEL

The key to our causality discovery algorithm lies in modeling the impelling and inhibiting influence contribution of historical events on current events.

When adopting existing impelling impact to both impelling and inhibiting impact through data transformation, inhibiting the occurrence of an event is equivalent to impelling the occurrence of event's complement. In this way, each individual point in event sequences need to be considered as the non-occurrence of the others. The event set E is represented as a sparse matrix, where the non-zero elements indicate the occurrence of events. In contrast, the complement matrix $\text{not } E$ converts the original zero elements to ones, leading to substantial overlaps and significantly increases the computational complexity. In addition, another technical challenge lies in transforming the impelling impact between vector/matrix obtained from existing models into inhibiting impact between events. For example, when identifying the causality among events: e_1, e_2, e_3, e_4, e_5 , an event point e_5 is converted into a vector: $[note_1, note_2, note_3, note_4, 0]$. This types of vector, composed of the non-occurrence states of events, is called a $\text{not } E$ vector. In this way, an event sequence: $[e_5, e_4, e_3, e_1, e_4, e_5, e_2]$ is transformed into a complement matrix shown in Fig. 2. Then, the output is causality between vector/matrix rather than the causality between events. However, A specific method is absent for transforming the impelling causality between “not E” vector/matrix obtained from existing models into inhibiting causality between events.

Sequence #1	u_5	u_4	u_3	u_1	u_4	u_5	u_2
Transformation		$\begin{bmatrix} \text{not } u_1 & \text{not } u_1 & \text{not } u_1 & 0 & \text{not } u_1 & \text{not } u_1 & \text{not } u_1 \\ \text{not } u_2 & 0 \\ \text{not } u_3 & \text{not } u_3 & 0 & \text{not } u_3 & \text{not } u_3 & \text{not } u_3 & \text{not } u_3 \\ \text{not } u_4 & 0 & \text{not } u_4 & \text{not } u_4 & 0 & \text{not } u_4 & \text{not } u_4 \\ 0 & \text{not } u_5 & \text{not } u_5 & \text{not } u_5 & \text{not } u_5 & 0 & \text{not } u_5 \end{bmatrix}$					

Fig. 2: Illustration of event sequence transformation when adopting existing causality discovery algorithm. An event sequence in the raw dataset is transformed into a matrix where the occurrence of certain event is considered as the non-occurrence of the others.

Based on the aforementioned considerations, it is not feasible to trivially adopting existing causality. Thus, we extended the existing causality model to identify fine-grained causality. Inspired by the causal discovery [4] that leverages Hawkes processes to describe how the historical occurrence of an event **increases** the probability of current events, we utilize Reactive point processes [5] [6] to model how the historical occurrence of an event **increases or decreases** the probability of current events. These two influencing factors are referred as impelling and inhibiting causal effects.

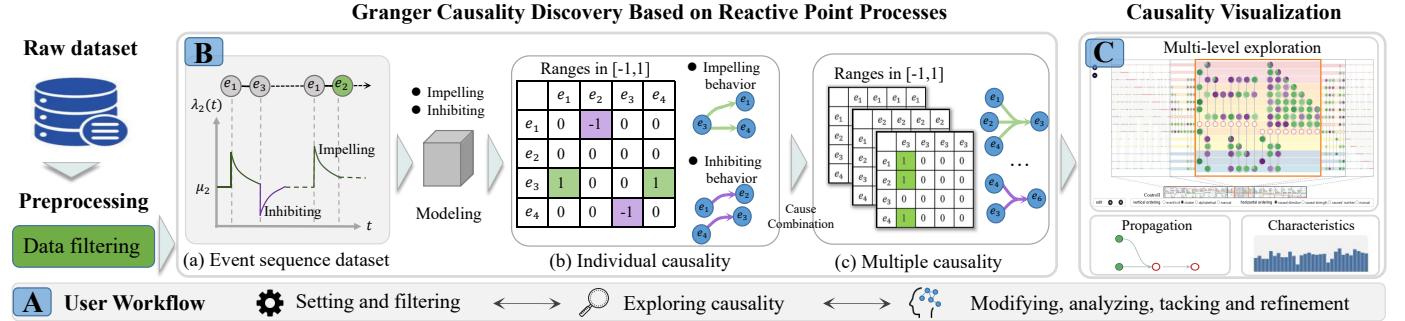


Fig. 3: The system architecture (A) and the pipeline of our combined causal discovery algorithm (B) which reveals impelling and inhibiting behaviors. The algorithm is based on Granger causality, which consists of three key steps: (a) training the Reactive point processes model to fit event sequence dataset, (b) discovering impelling or inhibiting influence on causality, which ranges in $[-1, 1]$ between two entities, and (c) inferring the cause combinations based on the individual causal discovery. The framework of our individual and combined causality analysis system (C) can be divided into the following steps: preprocessing and filtering raw dataset, modeling combined causal discovery and exploring directed hypergraph with diverse ordering operations and interaction techniques.

4.1 Fine-grained Causality on Reactive Point Processes

Point processes describes the distribution of random points (events) within a continuous time, which can naturally capture the temporal dependencies and dynamics between events. Thus, we leverage Reactive point processes to model Granger causality in event sequences. Then, the causality model is trained to fit event sequence dataset by optimizing objective function. We utilize parameters of the trained model to infer causal strength, as well as impelling or inhibiting behaviors of causality between events.

The causal discovery [4] uses hawkes processes to establish a model by relating impelling events. The occurrence probability of event \$e\$ is inferred from its conditional intensity function:

$$\lambda_e(t) = \mu_e + \sum_{e'=1}^E \int_0^t \phi_{ee'}(s) dN_{e'}(t-s) \quad (1)$$

where the first constant term \$\mu_e\$ denotes basic entity intensity and the second term represents the influence of history event \$e'\$ on entity \$e\$ at time \$t\$. \$\phi_{ee'}\$ is a kernel function to reflect influencing factors from event \$e\$ to \$e'\$. \$N_{e'}(t)\$ represents the number of event \$e'\$, which occurs at or before time \$t\$. The value of inferring impact factor is positive, a parameter obtained from the trained Hawkes processes model.

Hawkes processes only considers the impelling impact of events, while Reactive point processes model both impelling and inhibiting impact (T2) of chronological events on instantaneous intensity. The essential challenge of our model lies in modeling inhibiting causality. Considering that inhibiting causality is the opposite of impelling causality, our main idea is inhibiting the occurrence of an event \$E\$ is equivalent to impelling the occurrence of event's complement (*not* \$E\$). Fig.3(a) shows the event sequence modeling on Reactive point processes. The purple line represents inhibiting behavior of causality, decreasing the probability of an event occurring, and the green line represents impelling behavior of causality, increasing the probability of an event occurring. Then the occurrence probability of event \$u\$ can be expressed as follows:

$$\begin{aligned} \lambda_e(t) &= \mu_e + \sum_{e'=1}^E \int_0^t \phi_{ee'}(s) dN_{e'}(t-s) \\ &\quad - \sum_{e'=1}^E \int_0^t \psi_{ee'}(s) dN_{e'}(t-s) \end{aligned} \quad (2)$$

where \$\phi_{ee'}\$ and \$\psi_{ee'}\$ are kernel functions to reflect the impelling and inhibiting influencing factors from event \$e\$ to \$e'\$. The intensity

function \$\lambda_e(t)\$ combines an inhibiting term for accounting inhibiting effects and describing the causality intensity among entities from past entities. Given that \$\lambda_e(t)\$ describes the intensity of causal influence between entities, \$\lambda_e(t)\$ must be positive to ensure the non-negative intensity of the occurrence of events. Many researchers utilize a nonlinear function [6] [64] to guarantee non-negativity. Similarly, we leverage a nonlinear function \$\hat{\lambda}_e(t)\$ to guarantee the non-negativity of \$\lambda_e(t)\$.

$$\hat{\lambda}_e(t) = s \log \left(1 + e^{\frac{\lambda_e(t)}{s}} \right) \quad (3)$$

where \$s\$ is a small positive constant, which ensures that we can utilize the property of function \$g(x) = s \log(1 + \exp(x/s)) \approx \max\{0, x\}\$.

Then, we handle impact function as a linear combination of basis kernel functions, written as follows:

$$\begin{aligned} \phi_{ee'}(t) &= \sum_{m=1}^M a_{ee'}^m \kappa_m(t) \\ \psi_{ee'}(t) &= \sum_{m=1}^M b_{ee'}^m \kappa_m(t) \end{aligned} \quad (4)$$

where \$a_{ee'}^m\$ and \$b_{ee'}^m\$ are the coefficient of \$\kappa_m(t)\$, and \$a_{ee'}^m \in [0, 1]\$, \$b_{ee'}^m \in [-1, 0]\$. Using multiple basis functions improves the accuracy of model estimation but increases computing complexity. Thus, a commonly used basis function [23] is employed to balance accuracy and complexity of computation. To estimate and handle parameters in the model, we minimize objective function:

$$\min -L(\mu, a, b) + \alpha(\|A\|_F^2 + \|B\|_F^2) + \beta(\|A\|_1 + \|B\|_1) \quad (5)$$

where \$L(\mu, a, b)\$ is the Log-Likelihood function, which can be written as follows:

$$L(\mu, a, b) = \sum_{c=1}^C \left\{ \sum_{i=1}^{N_c} \log \hat{\lambda}_{e_i^c}(t_i^c) - \sum_{e=1}^E \int_0^{T_c} \hat{\lambda}_e(s) ds \right\} \quad (6)$$

where \$\mu\$ and \$a\$ are parameters that must be computed. \$\alpha\$ and \$\beta\$ are hyperparameters that control the regularization terms. The objective function is a convex function [6]. Thus, we choose the Gradient Descent method to minimize Equ. 5 and employ the *Monte Carlo* method to simplify the integral term for reducing the computational complexity as follows (see more details in supplementary materials):

$$\int_0^{T_c} \hat{\lambda}_e(s) ds = \frac{T_c}{N} \sum_{i=1}^N \hat{\lambda}_e(t^{(i)}) \quad t^{(i)} \sim U(0, T_c) \quad (7)$$

After obtaining parameters μ , a , and b , the causal strength of entities e on entity e' can be directly inferred from the impact coefficient a and b . Voting and trading-off strategies are utilized to determine causal strength by counting the number and the average of a and b in the coefficient. If e has an impelling impact on e' , we compute the average of positive coefficients in basis function from entity e to entity e' . For example, M_i represents coefficient set whose basis function coefficients are positive, negative, and zero. $M_{max} = \arg \max_{M_i} |M_i|$, which denotes the set with the largest number of elements among M_i . The causal strength can be inferred as follows:

$$C_{ee'} = \frac{1}{|M_{max}|} \sum_{m \in M_{max}} (a^m + b^m) \quad (8)$$

Therefore, we obtain a directed graph $G(V, DE)$, which unveils impelling and inhibiting behaviors on causality. Edges are weighted by causal strength θ_{uv} , where positive and negative value of θ_{uv} indicate impelling and inhibiting behaviors, respectively. Elements in nodes set V are the entities and edge set DE denote the causality between two entities. Then we leverage the above causal discovery algorithm to determine the causal strength of the following candidate cause combinations on effect entity.

4.2 Combined Causality Discovery

A cause combination consists of two or more entities, which individually might not be a cause. The elements in a cause combination interact with each other and jointly have a causal relation with other entities. We define “ $a \rightarrow b$ ” as a individual causality, where a and b refer to cause and effect entity. A combined causality “ $\{a, c\} \rightarrow b$ ” is then defined, where $\{a, c\}$ refers to a cause combination and b is an effect entity. The causality of these combinations is desired to be uncovered.

However, computing all possible combined causes is unnecessary and infeasible. First, a cause combination may exist invalid element or non-cause components, e.g., If X is a cause of Z and Y is not a cause of Z , while $\{X, Y\}$ is detected as a combined cause of Z . Then it is reasonable to assume that $\{X, Y\}$ can be a combined cause because of the existence of X has nothing to do with Y . Thus, we only consider the combined variables whose components are not causes. Second, the number of possible combined causes is 2^n given the limited entities, where n is the number of all entities. To reduce the number of cause combinations, the invalid cause combinations are decreased, and useful cause combinations are recruited based on a series of rules. Then we proposed principles of preprocessing the cause combinations to generate and decrease the number of candidate cause combinations to be verified. The preprocess contains two aspects.

(1) To guarantee effectiveness and controllability of candidate cause combinations (T1, T3), the following steps are defined. First, the causal relation between entities is intricate and ubiquitous. Thus, it is impossible and meaningless to combine all events as a combined cause. Defining the maximum element length of a combined cause is an efficient way and should be evaluated case by case to decrease the number of candidate cause combinations (e.g. 3). If we define maximum element length of a combined causes 3, all causes combinations that the element length is more

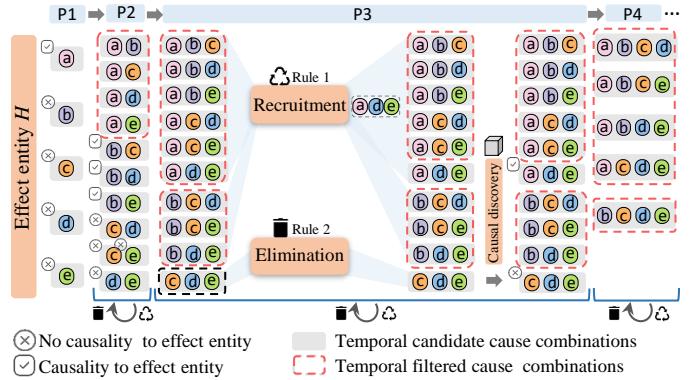


Fig. 4: Illustration of cause combinations process by eliminating the candidate cause combinations and recruiting the filtered combinations with the similarity and co-occurrence of entities.

than 3 would be meaningless. The maximum element length of a combined cause determines the number of iterations. Second, the iteration is implemented based on the result of previous iteration. Cause combinations are updated in each iteration. For example, cause combination of $\{a, c\}$ is regarded as a redundant combination if c has been detected as a cause entity. Fig. 4 illustrates cause combination process for a certain effect entity H . If a is a cause entity of H in process P1, we do not consider any cause combination entries containing a in next iteration. If $\{b, c\}$ is a cause combination of H in process P2, we do not consider any cause combination entries containing $\{b, c\}$. In this way, the number of candidate cause combination can be decreased.

(2) To further improve effectiveness and usability of candidate cause combinations, we eliminated the invalid cause combinations and recruited the filtered cause combinations. The similarities of entities are acquired by computing the vectors, which are obtained by training temporal event sequence data with Word2Vec [65]. If the similarity threshold value is set too high, there remains a substantial number of candidate combined causes to be verified, making verifying whether candidate combined causes are actual causes time-consuming. Conversely, lower threshold might lead to the oversight of valuable candidate combined causes. The co-occurrence number of entities are calculated from event sequence dataset. When all events in a candidate combined causes occur in the range of defined time interval, the occurrence of these events is seen as a co-occurrence. Then, we leverage similarity (threshold = 0.4) and chronological co-occurrence of cause entity for identifying the possibility of cause combination to determine to eliminate or recruit cause combinations. Take procedure P3 of Fig. 4 as an example, elimination or recruitment rules are defined.

Eliminating temporal candidate combinations. Candidate cause combinations are obtained shown in the black dotted box after procedure P2, where certain cause combinations may be redundant. Thus, these candidate cause combinations must be further filtered (T3) to decrease computation complexity and improve the effectiveness of cause combination. Considering that similar entities may together facilitate the occurrence of effect entity, we utilize Word2Vec to refine the selection of relevant causal combinations. Given that entities with a low co-occurrence are less likely to be a combination, we calculate the co-occurrence frequency of entities to filter out cause combinations involving elements with low co-occurrence.

Recruiting temporal filtered combinations. Redundant cause combinations are filtered shown in the red dotted box after

Algorithm 1 Cause Combination Algorithm

```

Input: Causal Graph  $G(V, E)$ , Maximum Length of Combination  $N$ 
Output: Hyperedges  $\Phi$ 
Initialize  $\Phi = \emptyset$  and candidate cause combinations  $\Psi = \emptyset$ 
Compute candidate cause combinations  $\Psi$  with principles
for  $i \leftarrow 2$  to  $N$  do
    for all  $i$ -combination  $\{v_1, v_2, \dots, v_i\}$  do
        if  $\{v_1, v_2, \dots, v_i\}$  in  $\Psi$  and satisfy recruitment rule then
            Filter  $\{v_1, v_2, \dots, v_i\}$  from  $\Psi$ 
        end if
        Combine  $\{v_1, v_2, \dots, v_i\}$  as  $v'$ 
         $V' \leftarrow \{v'\} \cup V \setminus \{v_1, v_2, \dots, v_i\}$ 
        Use  $V'$  to iterate new causal graph  $G'(V', E')$ 
        for cause  $u$  in  $V' \setminus \{v'\}$  do
            if  $v' \rightarrow u$  in  $E'$  then
                Add  $v' \rightarrow u$  to  $\Phi$ 
            end if
            if  $v'$  not in  $\Psi$  and satisfy elimination rule then
                Add  $v'$  to  $\Psi$ 
            end if
        end for
    end for
end for

```

procedure P2, where certain filtered cause combinations may be useful. Therefore, these filtered cause combinations must be further confirmed (T3). Considering that the similar entities may together facilitate the occurrence of the effect entity, we utilize Word2Vec and co-occurrence to detect cause combinations of similar causes. Given that entities with a high co-occurrence tend to be a cause combination, the filtered cause combinations are recruited when the entities in a cause combination have high similarity and high rate of co-occurrences.

4.3 Combined Causes Transformation

The co-occurrence of elements in a candidate cause combination is conceptualized as an occurrence of a virtual entity, which replaces the original events in the sequence to form a new input for causality discovery model. Thus, a virtual point is generated when all internal events of the candidate combined cause occur in the range of time interval threshold T .

In order to preserve the original information as much as possible, we only focus on internal elements and ignore external events of the candidate combination. For example, when $\{a, b\}$ is a candidate cause combination. Fig. 5 displays scenarios integrating combined causes into virtual points. Lines signify sequences, with red points and blue points respectively depicting internal and external elements of a candidate combination. As demonstrated in Fig. 5(E) and Fig. 5(F), blue points occur between red points (internal events are not adjacent). In instances external events occur between internal events, these external events are ignored and remain unchanged in the new sequence data. Thus, the solid line is converted into the dashed lines as follows.

Strategy not satisfied: As illustrated in Fig. 5(A), there only exist event a , and lack the occurrence of event b . Fig. 5(B) and Fig. 5(C) depict that the time interval of event a and b out of the range of the time interval threshold T . Thus, these scenarios do not satisfy the strategy for generating a virtual point.

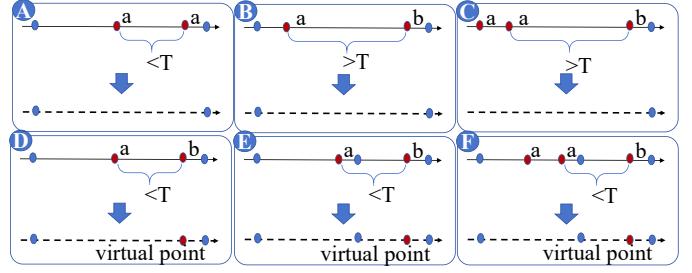


Fig. 5: Scenarios of integrating combined causes into virtual points.

Strategy satisfied: As demonstrated in Fig. 5(D) and Fig. 5(E), when time interval of event a and b is shorter than T , the occurrence of a and b is considered as a co-occurrence. Then a virtual point is generated to replace original events for forming a new input. In addition, if the first event occurs repeatedly (e.g. red point shown in Fig. 5(F)), we calculate time interval using the timestamp of its last occurrence (the second red point a in Fig. 5(F)).

In addition, a small value of threshold T indicates that events occurring in a small period are considered as a co-occurrence. In this way, events combinations occurring in a larger time interval may be exclusion, leading to ignoring of some useful data points. Conversely, a large value of threshold T represents that occurring of events in a large period is considered as a co-occurrence. In this way, most events combinations occurring in a long-time interval are retained, leading to the meaningless of events' co-occurrence. We set T 24 hours considering that Webpages visited within a day are potentially causally related.

The transformed data is fed into our causality discovery algorithm to ascertain whether this candidate combination is an actual cause. In this way, the existence and the (impelling or inhibiting) behavior of causality are determined with causal discovery algorithm. Algorithm 1 summarizes the process of candidate cause combinations by eliminating candidate cause combinations and recruiting filtered cause combinations. Then we obtain a directed hypergraph $DHG(V, DHE)$, where V is a set of vertices, representing the entities, and DHE is a set of directed hyperedges, denoting the combined causes. Each hyperedge $hypere \in DHE$ is a subset of V , serving as a collection of vertices that are connected by that particular hyperedge. This directed hypergraph is fed into the following visualization component.

5 CAUSALITY VISUALIZATION

With respect to the combined causality analysis tasks, this section introduces the details of each visual component, visual metaphor, and supported interactions. The user interface, illustrated in Fig. 6, includes three major components supporting interactive visual exploration: a combined causality panel (Fig. 6(A)) for showing the causality pattern and illustrating focused causality at different levels of information abstraction, a parameters configuration panel (Fig. 6(B)) for supporting users to efficiently obtain focused information, and a propagation panel (Fig. 6(C)) for tracking the influence propagation when given two entities.

5.1 Causality Aggregation and Ordering

Combined causality graph, indicated as a directed hypergraph, is used to illustrate combined causes. A cause combination consists of several impact factors, forming a set to point to an effect factor. Thus, our detected causality is a directed hypergraph, which contains sets information and directions information, as well as inevitably exists a large number of shared nodes and hyperedges

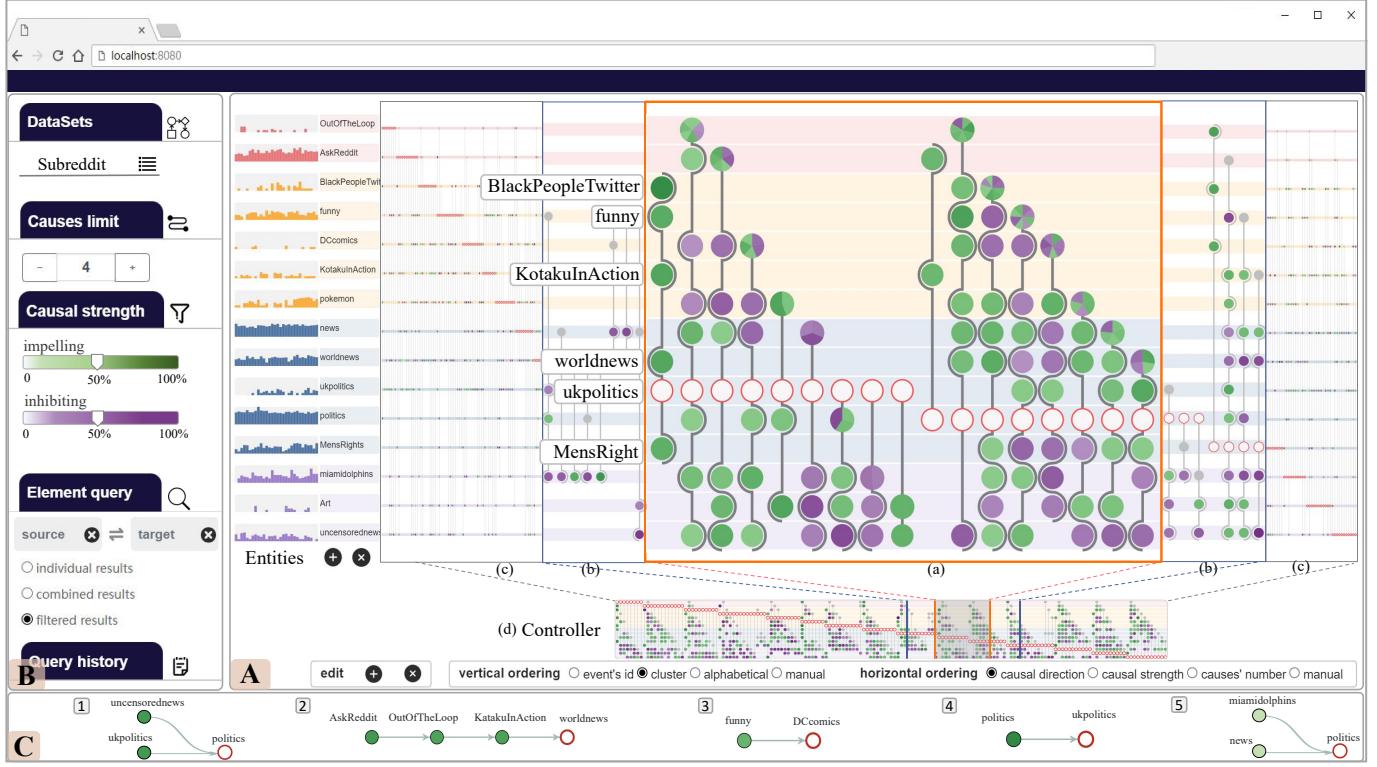


Fig. 6: *User interface*: (A) Combined causality panel for illustrating the individual and combined causes at different levels of causality abstraction with controller and diverse ordering. (B) Parameters configuration panel for providing users causality filters. (C) Propagation panel for finding and tracking causal propagation and storing the history of causal propagation.

intersections. One of our purposes is to illustrate both individual and combined causality. Due to the inherent challenges of edge intersection and node overlapping in graph visualization [60] troubling researchers for decades, no node-overlapping, no edge-intersection, and no link-ambiguity visualizations are necessary and desired to describe each pair of causality. Node-link graph drawing often leads to visual ambiguity that each entity in combinations can cause an effect independently. We designed a directed, weighted, and parallel-based hypergraph, which implements aggregation strategy, ordering operations, and multi-level exploration for causality analysis.

Parallel-based combined causality visualization (T4). focusing on no node-overlapping, no edge-intersection, and no link-ambiguity visualizations for combined causality, we utilize a parallel-based design to illustrate directed combined causality. This causality visualization aims to provide a comprehensive visual abstraction at various levels of causal summarization with a familiar node-link graph to present causalities. As illustrated in Fig. 8(A), the entities, aligned on the left side, are visually encoded with parallel horizontal tracks, where the tracks colors describe the communities of these entities, solid circles and hollow circles represent starting cause entities and ending effect entities, a line pointing to ending circle depicts and emphasizes the direction of causality, and a directed line connecting starting circle pointing to a red ending circle serves for a causality entry. The opacity of a cause circle represents causal strength, where darker color serves as higher causal strength. The color of circles depicts the characteristics of causalities for cause entities, where green and purple circles serve as impelling and inhibiting behaviors.

Visual metaphor (T4). Designing a formative and intuitive visual metaphor for multiple combined causes is an informative

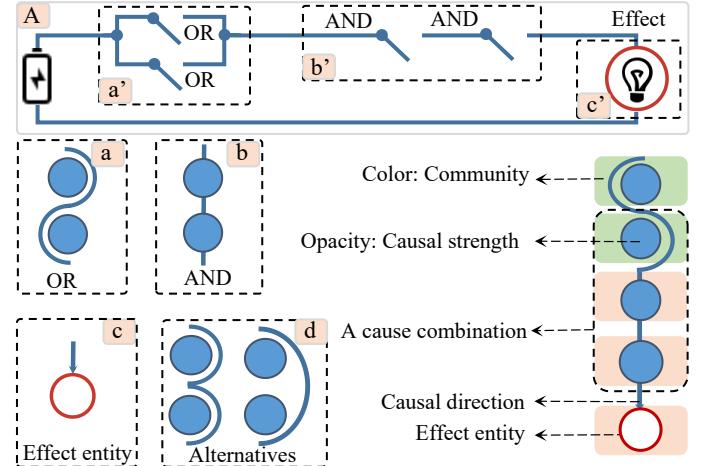


Fig. 7: Visual metaphors of “electrocircuit” (A) for illustrating the aggregated causality relationships. An arc surrounding a circle serves as an optional causal element (“OR”), a line crossing a circle encodes the combined causal element (“AND”), and a red circle depicts an effect entity.

and effective method for users to obtain the overall distribution of causal relations. As illustrated in Fig. 7(A), multiple causality relationships are aggregated as an informative visual metaphor of “electrocircuit”, which is a significant visual component in our system. Operations of merging optional causes (“OR”) and indispensable causes like functions of switches in a parallel circuit and a series circuit. A closed circuit can light a light bulb. Causes in Fig. 7(a) and (b) represent “AND” or “OR” switches in Fig. 7(a') and (b'). An effect entity in the lower left corner of Fig. 7(c)

represents a light in Fig. 7(c'). An arc surrounding a circle serves as an optional causal element (“OR”), a line crossing a circle encodes the combined causal element (“AND”), and a red circle encodes an effect entity. The right corner of Fig. 7 illustrates two pairs of causality, where a line crossing two circles is two indispensable cause elements (“AND”) and an arc surrounding two circles are optional causes element (OR). Two “AND” circles combining with a “OR” circle is a cause combination for an effect entity, such as a cause combination in the black dotted area. In addition, two visual metaphors (d) in the lower left corner of Fig. 7 are our alternative designs, which tend to lack pleasing and lead cognition bias on a cause combination. Thus, we utilize a pleasing and smoothing curve around optional causes, as well as a direct line crossing indispensable causes.

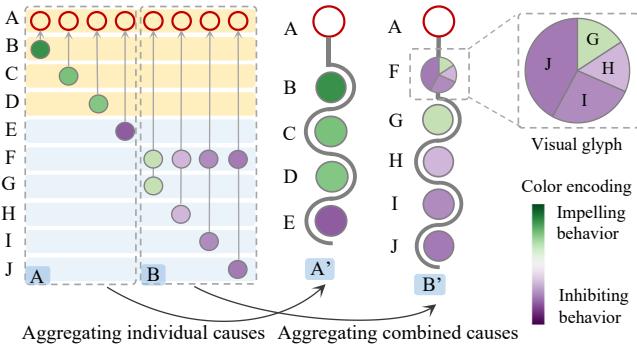


Fig. 8: Principles of causality aggregation to ensure informativeness and compact of similar cause combinations. Left original causality (a) and (b) are aggregated to the right metaphors (a') and (b').

Nodes binding and aggregating (T4, T5). Filtering and aggregating vertices to streamline graphs are common approaches to ensure readability and informativeness of hypergraphs, such as merging hyperedges that share the same entity. In our work, cause combinations are aggregated as visual metaphors of “electrocircuit” to optimize the horizontal space, eliminate demanding context switching and ensure scalability. As illustrated in Fig. 8, the matrix graph is the original causal hypergraph, where horizontal circles describe the same entity aligned on the left side and the solid or hollow circles encode cause or effect entities. Solid green or purple circles serve as the combined causes and red hollow circle represents the effect entity. A line connecting cause entity and effect entity means a pair of causal relationships. To aggregate causalities, causalities are divided into several parts according to counting the number of causes. For example, in Fig 5(a), the number of all causes for entity A is one. Each causality exists in the common effect entity A, resulting in their merging and aligning with AA (the effect entity). Causes are alternative causes (“OR”), which are aggregated by following the design of “electrocircuit” visual metaphor. In Fig 5(b), the element of all combined causes for entity A is two. Each causality has the common effect entity A, leading to their merging and then aligning with A (the effect entity). The causes share the common element F (“AND”), which also undergoes merging and alignment with entity F. Other causes are alternative causes (“OR”), which are aggregated according to the design of ‘electrocircuit’ visual metaphor.

Then the visual metaphor of “electrocircuit” in the middle of Fig. 8(a') and Fig. 8(b') are results of aggregating original causality(a)(b). We utilize an arc surrounding circles to indicate optional causes (“OR”) in combination, while a line crossing circles encodes indispensable causes (“AND”) in combination. Fig. 8(a)

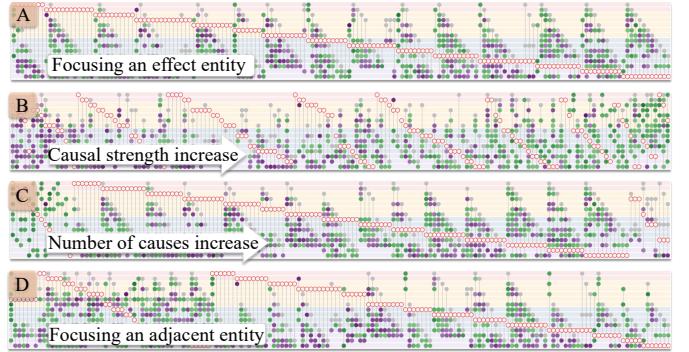


Fig. 9: Four types of hyperedges reordering for causalities. (A) Sorting by effect-orientation. (B) Sorting by causal strength. (C) Sorting by the number of causes. (D) Sorting by causal topology.

illustrates four causes, which can independently affect an effect entity. To save the horizontal space of edges, four pairs of causal relations are packed and collapsed in Fig. 8(a'), which saves the horizontal space of three edges. In the same way, Fig. 8(b) displays four pairs of causality. Entity F can be combined with the entity G, H, I, or J to affect entity A. These four pairs of causalities are aggregated in Fig. 8(b'). In addition, visual glyph is designed to merge and distinguish cause combinations. Pie chart in the upper right corner of Fig. 8(a) is utilized to illustrate the characteristic of aggregated causalities. The radian size of each sector represents causal strength, and the color of each sector represents impelling or inhibiting behaviors to compare various cause combinations.

Ordering and assignment (T5). Simplifying graphs with sorting nodes can improve visual aesthetics. Nodes and hyperedges refer to the entities and combined causes. Different orderings on nodes and hyperedges pose various visual clues, aiding diverse tasks, such as bringing similar causality together for uncovering communities and bringing top causality for highlighting the significant entities. The layout is modified for various requirements [61] with diverse row and column ordering strategies.

Row ordering. Nodes are sorted by ID, groups, alphabetical, and manual interactions to reveal and highlight causality patterns, such as central entities, communities, outliers, and scanning. In our work, following row ordering options are supported:

- **Base:** nodes are ordered by effect entities’ ID. This strategy aims to move causality belonging to a certain entity together. Users can focus on causality exploration on this entity.
- **Groups:** nodes are sorted by groups and similarities. We employ Louvain algorithm and Word2Vec for automatic community detection and similarity computation, which highlights potential communities and outliers. This ordering tends to reduce edge length when clustering similar entities.
- **Alphabetical:** nodes are sorted alphabetically. This ordering is beneficial for efficiently scanning, searching, and locating interested entities by name in long lists.
- **Manually:** nodes are sorted manually. Sometimes, automatic ordering may not satisfy certain needs, such as reducing the length of edges. This ordering is designed for manual reordering of horizontal entities.

Column ordering. Hyperedges are sorted by ID, causal strength, causal degree, and causal topology to highlight causality patterns, such as causality distribution, groups, and key nodes. In our work, users have following vertical ordering options:

- **Effect-orientation:** hyperedges ordered by effect entities are useful in certain applications shown in Fig. 9(A). This ordering moves vertices belonging to an effect entity together.

- *Causal strength*: hyperedges are ordered by causal strength shown in Fig. 9(B). This ordering filters causal strength and moves stronger causal edges together, which may reveal similar causal edges and enhance significant entities that have strong causal relations.
- *Degree of causality*: hyperedges are ordered by the number of combined causes shown in Fig. 9(C). This ordering tends to reveal the overall distribution of individual causalities as well as combined causalities.
- *Causal topology*: hyperedges ordered by the direction of a specific entity highlight key entity of interest shown in Fig. 9(D). This ordering reveals the topology information, such as finding adjacent nodes, counting related nodes, and identifying extremum on nodes.
- *Manually*: hyperedges are ordered manually. This ordering is designed for comparing the similar or interested causality manually.

Multi-level exploration (T6). Depicting causality at a different level of causal abstraction not only provides the distribution of causal information, but also the details of each pair of causality. Fig. 6 (A) shows that the focused area (a) provides and enhances a user’s concerned facts. Contextual area (c) is always necessary to enhance the expressiveness of a visualization. Transition area (b) provides a smoothing effect for aggregation/disaggregation from the context area to the focused one. Inspired by Multistream [63], controller panel (d) is designed, including brushing and linking techniques to help users handle and expand focus or context area. The information in a gray area in the controller is zoomed and expanded in the focused area. The causal relations in focused area are expanded by dragging the controller. When two sides of the controller are dragged, the gray area in the controller is broadened. When the focused area in the controller is dragged, the length of the contextual area is dynamically scrolled and updated. The focused area is broadened by dragging the two sides of the controller.

5.2 Auxiliary Visualization for Causality Exploration

Confirming and explaining impelling and inhibiting behaviors of causality are necessary, such as the navigation of amending the incorrect causalities and the influence propagation of the entities.

Causality propagation visualization (T5). Users always need to drill down into the causality and explore the influence propagation in large searching space. Causality propagation reveals how the stimulus propagates over and influences entities. The roots in a causal graph imply sources and the pathways uncover how a source entity propagates over and influences entities. Causality queries and propagation visualizations are supported in our system by clicking starting and ending entities. We employ Dijkstra algorithm [66] and Stratisfimal layout [67] for identifying and visualizing shortest influence pathways between specified entities.

Causality reconfirmation visualization(T2). Considering data noise, uncertainty, and complex interactions among entities, our data-driven causal discovery is not always faithful. The effect entities may be chronologically influenced by a pair of cause entities. We utilize traditional and easy-understanding multiple histograms to visualize the traffic volume information of each entity. Multiple histograms help users understand data characteristics, confirm causality behaviors, and navigate users modifying incorrect causality by comparing the chronology of entity occurrence.

5.3 User Interactions

Our work combines various basic and advanced interaction techniques, including multiple views linking and mouse hovering, clicking, dragging, and brushing.

Exploring causality at multiple levels (T6). Selecting and handling selected causality segments at various levels of abstraction. All causality can be expanded and merged interactively at a regular range number of causal relations by dragging the controller tool and brushing the overview of causalities.

Supporting diverse ordering options (T5). Entities and hyperedges are sorted by vertical and horizontal ordering (i.e., grouping similar entities, highlighting causality, and dragging focused entities manually) to efficiently recognize significant causality patterns.

Filtering causality and selecting subgraphs (T5). The causal strength and the causal combination depth filtering are supported. In addition, subgraph selection in the causal graph is supported to further explore interested causality.

Searching for causal propagations (T6). Users are supported to search for causal propagations by clicking the arbitrary visual encoding of entities (i.e., circles, layer bands, and texts) or entering two entities in input box.

Amending spurious causal relations. Considering that automatic causal discovery may not always be reliable, users’ domain experiences are supported to be incorporated into causal analysis to revise incorrect causality manually.

6 EVALUATION

This section evaluates the usefulness and effectiveness of our approach and visual design by two cases and a user study.

6.1 Performance of Causality Mechanism

The performance of combined causality detection algorithm is discussed in the following aspects.

Time complexity. The time efficiency of our algorithm is $O(INC)$ for parameter estimation, where I is the iteration number of gradient descents, N and C is the number of entities and sequences, respectively. Using Intel Core i5-6500 processor with 8GB memory, we conducted an experiment on four subreddit datasets, including 10 events, 15 events, 20 events, and 199 events, respectively. Table. 2 shows the time complex experiment on iterations of causality algorithm when threshold of causal strength is set 0.5. The result shows that time consuming of per iteration increases with the number of events and sequences increasing. The values recorded in the fourth column, approximately 45 seconds each, indicate a linear relationship between the time consuming of per iteration and the product of the number of events and the number of sequences.

TABLE 2: Time consuming of our causal discovery algorithm.

the number of events	the number of sequences	Time consuming of per iteration	Time consuming of per event	Number of causality impelling	Number of causality inhibiting
10	22	4.96	44.35	11	10
15	92	28.68	48.12	37	36
20	128	55.52	46.11	88	75
199	475	1890.30	48.00	492	337

Space complexity. For parameter estimation, the space complexity is $O(N^2M)$, where N is the number of entities and M is the number of basis function. In causality combination, the worst space complexity is $O(NC_N^K)$, where C_N^K is the combination number of entities, but in fact, the average space complexity is much less than the worst due to our proposed rule-based filtering in causality combination.

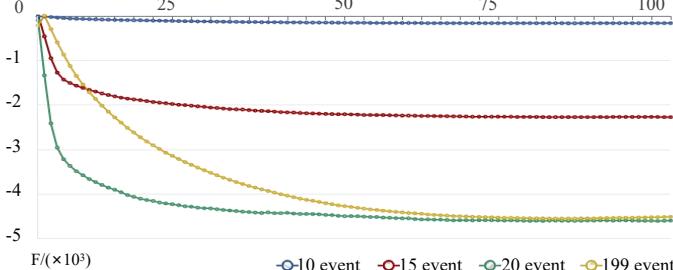


Fig. 10: The convergence of causality algorithm under four datasets.

Convergence of causality algorithm. Hyperparameters in our model may influence the stability of algorithm’s output, particularly when inputting data is small. Thus, to mitigate the uncertainty caused by hyperparameters, grid-based searching approach and gradual iterative optimization are utilized for gradually refining hyperparameters in causality discovery model. Subsequently, we determined the optimal hyperparameter values to be $s = 2$, $\alpha = 0.5$, and $\beta = 0.5$. To validate the convergence of our model under these parameter settings, we conducted an experiment on four datasets, including 10 events, 15 events, 20 events, and 199 events, respectively. As shown in Fig. 10, x axis presents the number of iterations, while y axis represents the value of objective function. Four distinct lines illustrate the performance of our causality algorithm in the 100 iterations under four datasets. The yellow line significantly increases in the initial iteration, followed by a subsequent decrease. In contrast, other lines’ rapid decline and followed by a gradual deceleration in the rate of decline. Ultimately, the values of lines converged toward a stable point. This observation indicates that the objective function tends to be convergent after iterations.

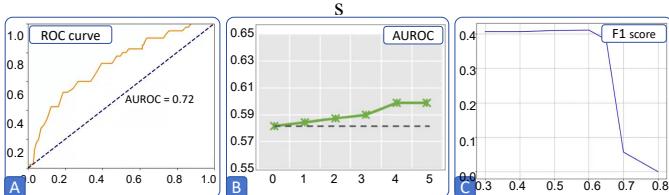


Fig. 11: The performance comparison. The ROC (A) curve and F1 score (C) of of our causality model. The AUROC values (B) of user-feedback causality model under iterations.

Accuracy of causality algorithm. We assess the accuracy by implementing the quantitative study on a well-known public news media dataset, MemeTracker dataset [68], which has the ground-truth causality for us to statistic Area Under ROC (AUROC) and F1 scores. Fig 11 summarize the performance metrics. Figure.11 (A) illustrates ROC curve of our causality algorithm and the value of AUROC is 0.72. Comparatively, Fig 11(B) illustrates the value of AUROC of user-feedback causality model [4] under several iterations. The horizontal axis represents the number of iterations, and the vertical axis represents the AUROC value of the interactive model [4] after iterations. The first value in Fig 11(B) represents the value of AUROC without user-feedback, and the following five values correspond to the AUROC values, ranges [0.57,0.61], obtained from five iterations to incrementally improve causality model. As shown in Fig 11(A), the value of AUROC for our method is 0.72, which is higher than the AUROC values of Jin’s method from five iterations (ranging from 0.57 to 0.61) on their user-feedback causality model. A higher AUROC value signifies enhanced accuracy of the causality results when benchmarked against the ground-truth causality. Thus, our method outperforms

to Jin’s method in terms of accuracy. In addition, as shown in Fig 11(C), we also statistic Precision and Recall to calculate F1 scores across different causal strength. Higher F1 score signifies a better model performance. The F1 score attains its peak value when the threshold is set to 0.6. At this setting, our approach exhibits acceptable overall efficacy at this threshold. Though our approach demonstrates potential for further refinement, it remains a valuable and effective method.

Comparative experiment of causality algorithm. We implemented a comparative experiment on our causality with a similar algorithm [4] across various causal strengths. The involved dataset is described in Section 6.2. Shown in Table 3, Each column shows the number of causality, the values in yellow and blue cells are the number of inhibiting causality and impelling causality. *Inhibiting causality*: There exists three pairs of inhibiting causality detected by our algorithm. Popularity is the common feature of the three causality, therefore similar topics have a mutual incentive causal effect. *Impelling causality*: in terms of the total number of impelling causality, our model obtained a higher count compared to that obtained by Jin [4]. We speculate that this phenomenon is caused by the wider coefficient term of the kernel function in our model. In addition, our model exhibits a higher concentration of causal strength within the interval (0.6,0.8], whereas the causal discovery [4] demonstrates a more dispersed distribution.

TABLE 3: The comparison of our algorithm with Jin et al. [4].

Causal Strength (0)	[0.4, 0.5)	[0.5, 0.6)	[0.6, 0.7)	[0.7, 0.8)	[0.8, 0.9)	[0.9, 0)						
Ours	0	1	0	6	0	36	0	11	3	0	0	0
Jin et al. [4]	-	2	-	2	-	2	-	8	-	11	-	5

Efficiency of cause combination. As mentioned in Section 4.2, we choose several strategies to decrease the search space. Therefore, the running time does not reach the worst time complexity in the actual experiment. We test our filter rules in a dataset with seventeen entities. As shown in Table 4. In the two-, three-, and four-entity combinations, the filtering rule reduces 11%, 30%, and 64% of combinations, respectively.

TABLE 4: The number of cause combinations based on rules.

Combination number	2	3	4	Average
Original	136	680	2380	1065
Filtered	120	471	849	480
Percentage	11%	30%	64%	55%

6.2 Case Study I: Causality Investigation for Subreddit

The first one is the subreddit data collected from October to November 2016. The dataset covers the top 15 subreddits occurring most frequently after filtering irrelevant entities by Word2Vec and co-occurrence. A subreddit represents a community with a specific area of interest. We filter and extract each user’s commenting trajectory including account name, subreddit and timestamp, on these 15 subreddits, including *News*, *Worldnews*, ..., and *funny*, as an event sequence.

First, we want to explore the causality from individual causality to combined causality. Thus, after ranking causality in ascending order by the length of cause combinations, the causality is illustrated in Fig. 12(A). The left square matrix diagram illustrates individual causal relationships, with the nodes predominantly green, indicating an impelling effect. The remaining portion illustrates combined causality, where the nodes are mainly purple, signifying an inhibiting effect. For individual causality, we found that most

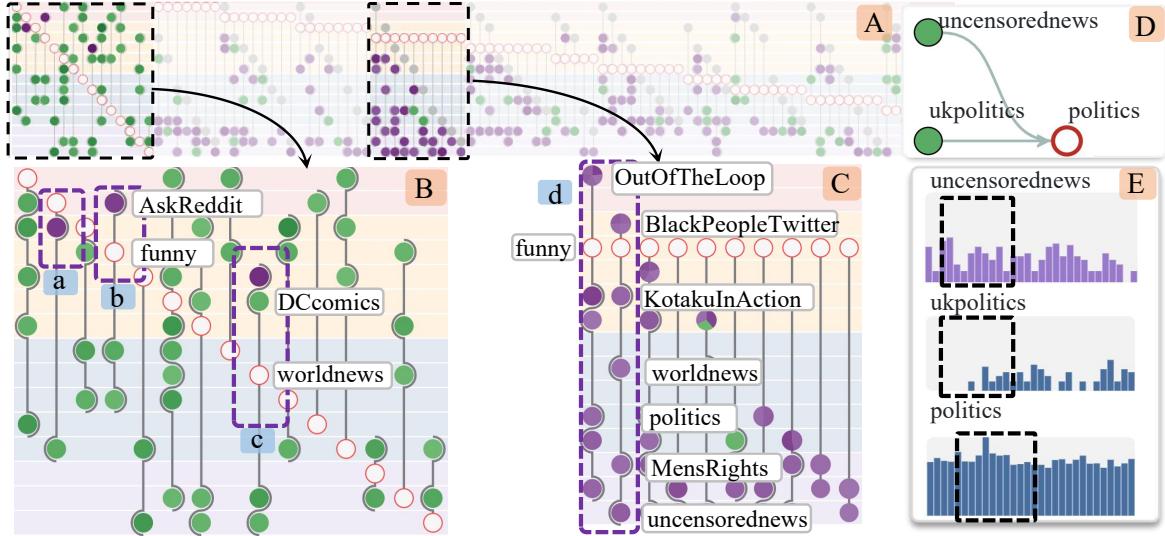


Fig. 12: The causality of subreddit dataset. (A) Overall distribution of causality among subreddits after ascending ordering by the number of cause combinations. Dedicated to (B) individual causality and (C) combined causality for analyzing interactions within the community. Searching for more causality propagation about related entities (D) and conducting preliminary verification of causality based on entities' traffic fluctuations in chronological order(E).

of causality pose an impelling effect, hypothesizing that this phenomenon may be due to the involved subreddits drawing large attention and popularity is their common feature, therefore similar topics tend to pose an impelling effect. For combined causality, we found that inhibiting causal effect dominate We speculate that the Internet's capacity and total visitation are constrained by the limit of humans' attention, leading to existence of inhibiting behaviors. Thus, the inhibiting behaviors easily occur in combined causality.

However, focusing on the individual causality illustrated in the purple box of Fig. 12(B), there exist three inhibiting causal relations. The subreddit *AskReddit* (a) is influenced by *BlackPeopleTwitter*, the subreddit *funny* (b) is influenced by *AskReddit*, and the subreddit *worldnews* (c) is influenced by *DCcomics*. The subreddit *funny* (b) is a humor depository and the subreddit *BlackPeopleTwitter* is a platform of screenshots of Black people being hilarious or insightful on social media. While *AskReddit* is the place to ask and answer thought-provoking questions. They belong to various types of community, thus, there exist inhibiting impacts from *BlackPeopleTwitter* to *AskReddit*, and from *AskReddit* to *funny*. We found *Suicide Squad* was released by *DCcomics* and attracted a lot of attention. This explosive topic has an inhibiting impact on long-term popular topic *worldnews*.

Then, focusing on the combined causes shown in the purple box of Fig. 12(C), the subreddit *funny* (d), a humor depository, is influenced by pairs of combinations. These combinations all contain necessary elements, contemporary issues-related topic *OutOfTheLoop* or humor-related topic *BlackPeopleTwitter*. The two subreddits, *funny* and *BlackPeopleTwitter* belong to humor topic, suggesting that similar subreddits tend to pose an inhibiting behavior in combined causality. We search for the original posts and comments in subreddit because we only have clickstream dataset and lack supplementary information. We speculate that while announcing the November 2016 U.S. presidential election results, users who like politically related subreddit *uncensorednews* and *worldnews*, simultaneously tend to explore the humor-centric *BlackPeopleTwitter* subreddit. Consequently, the traffic of a similar humorous subreddit *funny* tends to be reduced. Thus, *BlackPeopleTwitter* associates distinct political related subreddit

to be combined causes, resulting in an inhibiting causality with *funny*. These inhibiting causal relationships follow the principle of zero-sum game.

Further exploring politically related topic and searching for *ukpolitics* and *politics*, we obtained a combined causality shown in Fig. 12(D). The subreddit *ukpolitics* and *uncensorednews* forming a cause combination has an impelling impact on *politics*. To confirm the causal relationships, we check the attention distributions of subreddits in Fig. 12(E). The histograms of entity *ukpolitics* (e) and *uncensorednews* (f) in a combined cause have a similar peak when *politics* (g) also reaches a peak in corresponding chronological order. Thus, *ukpolitics* and *uncensorednews* have an impelling impact on *politics*. When advocating policy, based on the findings in our case and the agenda setting theory from social science, journalists can focus on the causal related topics to engage in the policy implementation process (e.g., designing targeted content in *BlackPeopleTwitter* to attract traffic from *funny*) actively and comprehensively, ensuring that policies align with the issues of public attention.

6.3 Case Study II: Causality Analysis of Website Panels

The second one comes from Internet Information Server logs for msnbc.com and news-related portions of msn.com for the entire day for September 1999. Each user's request for a page consists of an event sequence. There exist 16 entities, including *frontpage*, *news*, *health*, *living*, *weather*, *on-air*, ..., and *business*.

First, We aim to analyze the interactions between modules exhibiting strong causal relationships. Then, after listing causality in ascending order by causal strengths, the overall trend of causality shown in Fig. 13(A), most nodes are purple and a few retained nodes in the black dashed box are green. This means that most events have an inhibiting behavior and little events have an impelling behavior.

Then, focusing on the green causality and brushing them, we find these green nodes mainly come from yellow and blue communities shown in Fig. 13(B). The entities in the two clusters are daily topics, such as *Health*, *Living*, *Weather*, *Travel*, and *Sports*. When hovering over the causality, the details of causal relations are illustrated. For example, as illustrated in Fig. 13(C),

TABLE 5: Summary of user evaluation for various tasks. *Ent.*, “ $1 \rightarrow 1$ ”, “ $n \rightarrow 1$ ” and *Path* serve for counting entities, recognizing single edges, identifying hyperedges and finding propagation path, respectively. Parameters in *Sample Size* represent the number of entities, single edges and hyperedges in tasks. Length of bars depicts trimmed mean (in seconds) and errors bars serve for standard errors. Orange bars or blue bars represent the performance of parallel-based technique or node-link technique on completion time and accuracy. Faster or more accurate results are indicated in bold.

	Ours	Baseline							
Sample Size	Preference		Completion Time		Accuracy				
		Ent.	“ $1 \rightarrow 1$ ”	“ $n \rightarrow 1$ ”	Path	Ent.	“ $1 \rightarrow 1$ ”	“ $n \rightarrow 1$ ”	Path
3, 3, 1	4 6	5.0 4.2	13.2 9.4	10.9 8.2	9.8 7.2	100% 100%	100% 89.5%	100% 89.5%	100% 100%
4, 4, 4	14 2	4.5 4.6	8.0 34.5	13.0 23.5	9.7 14.0	100% 100%	100% 78.9%	78.9% 63.2%	100% 94.7%
6, 9, 4	17 1	6.1 5.6	15.4 36.1	20.5 43.3	17.8 19.6	100% 100%	100% 84.2%	94.7% 63.2%	100% 100%
7, 6, 7	17 0	7.0 6.0	17.0 29.9	16.5 31.8	17.1 11.4	100% 100%	94.7% 78.9%	89.5% 89.5%	94.7% 94.7%
8, 5, 6	12 4	6.6 6.4	14.6 35.3	14.2 29.5	18.7 13.2	100% 100%	100% 89.5%	100% 89.5%	84.2% 100%
10, 2, >10	16 1	8.9 9.0	18.6 30.9	22.3 39.5	22.4 23.1	95% 100%	100% 73.7%	100% 100%	63.2% 63.2%

two aggregated causalities exist. The first one is two pairs of combined causality where *Travel* can incorporate *on-air* or *Misc* respectively to influence *Sports*. The second one is three pairs of causality where *Living*, *Misc*, or *Business* can influence *Health* respectively, suggesting that similar kinds of topics tend to have an impelling behavior on each other.

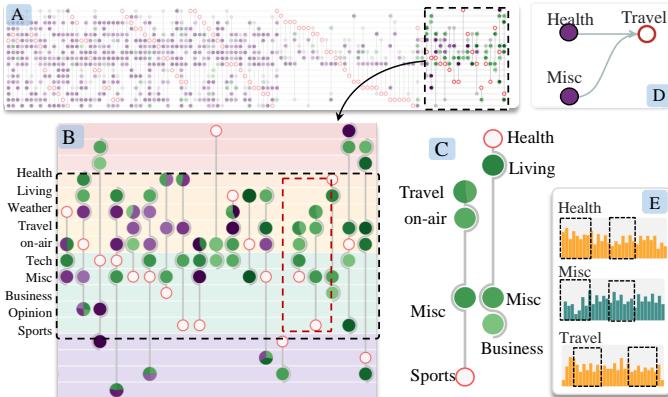


Fig. 13: The causality of Website Panels interactions. (A) The overall diagram depicting the causality between panels predominantly inhibiting effects, with a minor presence of impelling effects (B). Dedicated to impelling causality (C) for analyzing impelling interactions within panels. Searching related entities for obtaining more causality propagation(D). Implementing preliminary verification based on the temporal characteristics(E).

Then, tracking *Health* and *Travel*, as shown in Fig. 13(D), a combined causality exists that *Health* and *Misc* form a cause combination, and purple nodes present that the cause combination has inhibiting impact on *Travel*. To confirm the causality, we find that the histograms of the former two entities *Health* and *Misc* in the cause combination have two similar periodic peak distributions shown in Fig. 13(E). The histograms of the result entity *Travel* have two lead-lag valley values distributions in chronological order. Thus, we inferred that the combined causes of *Health* and *Misc* have an inhibiting impact on *Travel*. When advertising is integrated into a website, it may influence users' interactions with interconnected internet pages. These causality among internet pages help policymakers to promptly identify impacts and optimize resource allocation, enhancing policy flexibility and resource efficiency.

6.4 User Evaluation

6.4.1 Controlled experiment

We conducted a pilot user study to evaluate the readability of our parallel-based technique compared with node-link technique, where directed hyperedges of each combination causes are bounded to point to an effect entity as well as multiple edges of diverse individual cause directly point to an effect entity.

Participants. We recruited 19 participants (11 females, 8 males) to implement a controlled study. The participants varied in age from 20 to 28, including 2 undergraduate students, 10 master students and 7 doctoral students. Five participants, involving two control engineering majors and three computer vision majors, have no background in data visualization. Fourteen participants major in computer science and have more than one year's experience in data visualization. All participants are independent decision makers in the experiment.

Tasks. Relying on commonly encountered tasks while analyzing graph [61], we identified five types of readability tasks on browsing preference, counting entities, recognizing type of single edges or hyperedges, and finding the propagation path. The tasks are given below: *Q1*-which type of visual design do you like? *Q2*-how many entities? *Q3*-how many one-to-one relationships? *Q4*-how many many-to-one relationships? *Q5*-could you find the pathway from “entity a” to “entity b”? *Q1* is to determine which directed hypergraph is aesthetically pleasing in the two types of visualization technique (node-link technique or parallel-based technique). *Q2* serves for counting event elements. *Q3* and *Q4* are designed for determining single edges or hyperedges. *Q5* is to search for one linking pathway, an alternating sequence of nodes and links from starting to ending nodes. Both visualization techniques utilize identical inputs and answers. The order of tasks is randomized locally but is kept from low to high difficulties throughout the controlled study. To avoid tedious counting, we adjusted the answer “>10” rather than an exact value to improve participants satisfaction during survey experience.

Procedure. Our experiment is a within-subject design. Before the controlled user study, we additionally invited three participants, who are not included in the above 19 users, to test all tasks and give improving suggestions such as too complex problems, tedious issues, and repeating tasks. Tasks are finally determined after three iterations. Then 19 participants received instructions for the training on two types of visualizations. The training is divided into

three steps, including perception of visual metaphor, description of visual encoding and tasks testing. After 5-10 minutes training for each participant, each of the participants completed all questions independently. The experiment website presented questions and recorded time taken for each answer. Afterward, we collected and summarized comments after the experiment.

Hypotheses. Parallel-based technique does not reduce the visual clutter compared to node-link technique. Based on the tasks, we formulated hypotheses as follows:

- *Null hypothesis (H0):* Parallel-based technique performs worse in user preference(a), completion time(b) and accuracy(c) than node-link technique to accomplish corresponding tasks.
- *Alternative hypothesis (H1):* Parallel-based technique outperforms node-link technique in user preference(a), completion time(b), and accuracy(c) to accomplish corresponding tasks.

6.4.2 Result

We use multiple bar charts to analyze user preference, accuracy, and time performance on two techniques, as shown in Table 5. We also use significance test to analyze time [69] and accuracy [70] task-specific performance results.

User preference. User preference results are summarized in the second column of Table 5. Overall, participants prefer parallel-based technique compared to node-link technique. Neutral participants slightly shifted their preference towards our visual design. One commented that commonly used node-link technique “annoying and visually distracting for counting task.” When the number of entities is three, more participants support node-link design. One claimed that “While the non-overlapping design is pleasing, learning a new design is exhausting, then I vote for node-link graph.” However, most participants comment that they get visually lost while exploring complex node-link design.

Completion time. The completion-time results are summarized in the third column of Table 5, where the length of bars depict 10% trimmed mean (in seconds) and errors bars represent standard errors. Overall, using our parallel-based technique to accomplish counting tasks is significantly more efficient than using node-link technique with the number of entities increasing. In particular, for hyperedge counting tasks, our technique demonstrates a performance speed that exceeds the node-link technique by over 50%. Surprisingly, we found that time consuming of both techniques decreases when the number of hyperedges is above 10. Through investigating this surprising finding further, we observed that most participants give the answer at first glance rather than counting all relationship pairs. The node-link technique is faster for low difficulty sample size. No significant differences were found between two types of techniques aiming at *entities counting* task when the number of entities is less than three, while node-link technique slightly outperformed parallel-based technique in *path finding* task.

Accuracy. The accuracy results are summarized in the fourth column of Table 5. Overall, two types of visualization techniques tend to remain constant in accuracy across *entities counting* task and *path finding* task. Parallel-based technique performed above 10% more accurately than node-link technique *edges and hyperedges counting* task. Participants have weaker accuracy in node-link technique than in parallel-based technique. Meanwhile, parallel-based technique does not seem to considerably outperform node-link technique in *path finding* task.

Hypotheses summary. *Preferences:* Overall, 17 out of 19 participants commented that parallel-based technique was more aesthetically pleasing. Our hypotheses (H1) on Q1-participants’

preferences were also validated ($p = 1.75 * 10^{-5}$) with Barnard’s Test [70]. *Completion Time:* We employ Mann-Whitney Test [69], a non-parameter statistical test method used in two sample problems, to analyze time performance results. Mann-Whitney Test revealed a significant effect in time for Q3-edges counting tasks ($p = 1.64 * 10^{-11}$) and Q4-hyperedges counting tasks ($p = 6.85 * 10^{-6}$). We verified (H1) as, overall, participants are significantly faster with parallel-based technique than with node-link technique. We also found a significant effect in time for Q5-path finding task ($p = 0.94$), partially validating parallel-based technique performs worse than node-link technique in path finding task. No significant differences in time consuming were found for Q2-entities counting. *Accuracy:* We also utilize Barnard’s Test, an odds ratios test approach for contingency table, to analyze accuracy performance results. Our hypothesis (H1) on accuracy was also validated for Q3-edges counting tasks ($p = 3.94 * 10^{-6}$) and Q4-hyperedges counting task ($p=0.01$). Overall, while no significant differences in accuracy between two types of visualization techniques for Q5-path finding task and Q2-entities counting task.

6.4.3 User Feedback

After the above experiment, we described the visual encoding and user interactions in our system. User feedback is collected from ten participants, who have more than two years of data visualization experience, to gain insights into the effectiveness of our visual encoding and the usefulness and usability of our system. Each user spends about twenty minutes for system exploration. User feedback was summarized as the following three aspects:

Visual encoding. All participants commented that the parallel-based technique is informative, engaging, and well-designed, but not intuitive. One participant particularly claimed that “visual metaphor of “electrocircuit” for aggregated hyperedges is impressive”. Two participants commented that user interactions, especially the focus and context technique, are smooth and useful to explore causality from overview to details. However, three participants pointed out that they prefer traditional node-link graphs.

Scalability. All participants appreciated our combined causality analysis system. They all confirmed the usefulness and effectiveness of our visual design and interaction techniques. Seven of them commented that the aggregation strategy and the focus and context technique in our interactive system guarantee the scalability to visualize directed hypergraph. However, one participant raised scalability concerns about vertical axis for hundreds of entities.

Improvement. Users provided valuable suggestions on how to strengthen our visualization. Although the system received positive feedback during the interview from participants, they still raised concerns about visual perception of the topology in our visual design. In addition, the system lacks sufficient interpretability and explainability to support pattern recognition and decision-making. Two of them perceived that more explanations of impelling and inhibiting behaviors should be illustrated towards a better understanding of causalities.

7 DISCUSSION

The experiments on two case studies and a pilot user study confirm the usability and effectiveness of the approach for discovering impelling and inhibiting behaviors, as well as visual causality analysis system. However, the present work has some limitations.

Causal discovery algorithm. We attempt to integrate user feedback on causal diagnosis into the causal model to update our

causality model. However, our causal discovery algorithm is time-consuming. Although multiple parallel computations on multiple cores have been used, causal discovery cannot be computed in real time. Thus, human-in-the-loop cannot be reached due to the high complexity of the causal model. First, causality usually evolves over time, posing a challenge for a fixed kernel function to encompass all scenarios. Therefore, identifying a kernel function suitable for specific situations becomes a non-trivial task. Second, we have defined a threshold T to detect the co-occurrence of elements in a candidate combination. A smaller value of T enhances events' co-occurrence more meaningful, while a large value of T makes more useful data points retained. Thus, it is a non-trivial trade-off between the meaningfulness of events' co-occurrence and the retention of more extensive and valuable data points. Third, Monte Carlo method is employed to simplify the integration term, leading to reducing computational complexity but potentially lowering the accuracy of causality.

Visual design. The primary concerns of visual design limitation would be the number of entities (y-axis) and causalities (x-axis). Parallel-based visual design makes the vertical arrangement and horizontal repetition of entities, resulting in limited available space for entities. Then, using the partition hierarchy may partially mitigate this issue. Aggregating strategies and focus-and-context technique improves the scalability of visual design. Nevertheless, they do not scale infinitely. Second, to ensure accurate representation, maintain visual consistency, and address more diverse tasks, our approach is deficient in sufficient summarization. Although our proposed edge aggregation strategies and sorting approaches have already improve this issue. However, it still lacks intuitive high-level insights extraction. Third, result of controlled user study reveals that our parallel-based technique outperforms node-link technique in completion time and accuracy for counting task, especially when increasing the number of combined correlations. In addition, the visual encoding of "OR" or "AND" logic is not attractive, especially when the edges are light and thin. Compared with node-link technique, our parallel-based technique suffers from several challenges such as certain learning curve for visual design. Node-link technique outperforms our parallel-based directed hypergraph in counting tasks when the number of entities is less than 3. Thus, more intuitive visualization on combined causality is desired to provide easy-understanding information.

Pattern identification. Our detected causal hypergraph contains significant causal patterns. Although our system has provided diverse ordering strategies for assisting causal analysis, experts may need to drill down to special causality pattern mining. This work is the first attempt to use parallel-based, directed, and aggregated hypergraph to analyze impelling and inhibiting behaviors. Thus, causality pattern discovery approaches are still enlarged and designed to automatically identify significant impelling and inhibiting behavior patterns.

Future work has two directions, including combined causality discovery and directed hypergraph visualization. Our combined and fine-grained causality discovery need to be improved. An efficient combined causal discovery is desired to address the high complexity of impelling and inhibiting behavior modeling. In this way, human verification can be further embedded into the causality algorithm to incrementally improve and update causality model in real time. We also attempt to summarize the characteristics of causality patterns and apply image processing technology for extracting causal patterns automatically. Meanwhile, our proposed directed hypergraph visualization design also needs to be advanced. First,

vertical arrangement and horizontal repetition of entities sharply reduced the available space for entities in our parallel-based visual design. Partition hierarchy and aggregation strategies may partially mitigate this issue. Second, effective data summarization requires the integration of diverse analytical tasks, including community detection, topological features, network flow, etc. Thus, illustrating high-level information from graphical representations involves balancing between data abstraction and precise expression, which would be our future work. Third, our visual design does not provide an intuitive view of causality propagation. Overlaying more information (a visual encoding of causality pathway) to the parallel-based hypergraph may cause visual clutter but be a feasible method. Thus, trading-off visual clutter and informativeness would be our another future work. In addition, causalities always change dynamically, our parallel-based visualization is not well-extended to illustrate directed hypergraph because of redundant nodes. Merging similarities and distinguishing differences of dynamic directed hypergraph is a key challenge in our parallel-based visualization.

8 CONCLUSION

In this paper, we utilize Granger causality based on the Reactive point processes for unveiling the inhabiting and impelling behaviors. We also propose a novel directed hypergraph visualization, which incorporates a causal discovery mechanism to detect combined causes among entities. VAC² provides a parallel-based visual design for directed hypergraph, which embeds diverse sorting layouts, aggregating strategies, and focus-and-context techniques to help users efficiently analyze complex combined causality. Furthermore, we also conduct two case studies and a pilot user study to evaluate the usability and effectiveness of our work.

REFERENCES

- [1] J. D. Angrist, G. W. Imbens, and D. B. Rubin, "Identification of causal effects using instrumental variables," *Journal of the American Statistical Association*, vol. 91, no. 434, pp. 444–455, 1996.
- [2] R. Guo, L. Cheng, J. Li, P. R. Hahn, and H. Liu, "A survey of learning causality with data: Problems and methods," *ACM Computing Surveys*, vol. 53, no. 4, pp. 1–37, 2020.
- [3] H. Zhang, C. Yan, S. Zhou, J. Guan, and J. Zhang, "Combined cause inference: Definition, model and performance," *Information Sciences*, vol. 574, pp. 431–443, 2021.
- [4] Z. Jin, S. Guo, N. Chen, D. Weiskopf, D. Gotz, and N. Cao, "Visual causality analysis of event sequence data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1343–1352, 2020.
- [5] S. Ertekin, C. Rudin, and T. H. McCormick, "Reactive point processes: A new approach to predicting power failures in underground electrical systems," *The Annals of Applied Statistics*, vol. 9, no. 1, pp. 122–144, 2015.
- [6] X. Zhang, Y. Zhou, D. Wu, M. Hu, X. Zheng, M. Chen, and S. Guo, "Optimizing Video Caching at the Edge: A Hybrid Multi-Point Process Approach," *IEEE Transactions on Parallel and Distributed Systems*, 2022, to be appeared.
- [7] C. K. Assaad, E. Devijver, and E. Gaussier, "Survey and evaluation of causal discovery methods for time series," *Journal of Artificial Intelligence Research*, vol. 73, pp. 767–819, 2022.
- [8] D. Colombo, M. H. Maathuis *et al.*, "Order-independent constraint-based causal structure learning," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 3741–3782, 2014.
- [9] P. Spirtes, C. Glymour, and R. Scheines, "From probability to causality," *Philosophical Studies*, vol. 64, no. 1, pp. 1–36, 1991.
- [10] E. V. Strobl, S. Visweswaran, and P. L. Spirtes, "Fast causal inference with non-random missingness by test-wise deletion," *International journal of data science and analytics*, vol. 6, pp. 47–62, 2018.
- [11] D. M. Chickering, "Optimal structure identification with greedy search," *Journal of machine learning research*, vol. 3, no. 11, pp. 507–554, 2002.

- [12] J. Ramsey, M. Glymour, R. Sanchez-Romero, and C. Glymour, “A million variables and more: the Fast Greedy Equivalence Search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images,” *International Journal of Data Science and Analytics*, vol. 3, no. 2, pp. 121–129, 2017.
- [13] P. Nandy, A. Hauser, and M. H. Maathuis, “High-dimensional consistency in score-based and hybrid structure learning,” *The Annals of Statistics*, vol. 46, no. 6A, pp. 3151–3183, 2018.
- [14] I. Tsamardinos, L. E. Brown, and C. F. Aliferis, “The max-min hill-climbing Bayesian network structure learning algorithm,” *Machine learning*, vol. 65, no. 1, pp. 31–78, 2006.
- [15] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, “Causality detection based on information-theoretic approaches in time series analysis,” *Physics Reports*, vol. 441, no. 1, pp. 1–46, 2007.
- [16] K. P. Murphy, “Bayesian map learning in dynamic environments,” *Advances in neural information processing systems*, vol. 12, 1999.
- [17] S. Ma, J. Li, L. Liu, and T. D. Le, “Mining combined causes in large data sets,” *Knowledge-Based Systems*, vol. 92, pp. 104–111, 2016.
- [18] U. Hasan, E. Hossain, and M. O. Gani, “A survey on causal discovery methods for iid and time series data,” *arXiv preprint arXiv:2303.15027*, 2023.
- [19] A. K. Seth, A. B. Barrett, and L. Barnett, “Granger causality analysis in neuroscience and neuroimaging,” *Journal of Neuroscience*, vol. 35, no. 8, pp. 3293–3297, 2015.
- [20] J. Geweke, “Inference and causality in economic time series models,” *Handbook of econometrics*, vol. 2, pp. 1101–1144, 1984.
- [21] Z. Deng, D. Wang, X. Xie, J. Bao, Y. Zheng, M. Xu, W. Chen, and Y. Wu, “Compass: Towards Better Causal Analysis of Urban Time Series,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 1051–1061, 2022.
- [22] A. C. Lozano, H. Li, A. Niculescu-Mizil, Y. Liu, C. Perlich, J. Hosking, and N. Abe, “Spatial-temporal causal modeling for climate change attribution,” in *Proceedings of ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, pp. 587–596.
- [23] H. Xu, M. Farajtabar, and H. Zha, “Learning granger causality for hawkes processes,” in *Proceedings of the International Conference on Machine Learning*, vol. 48. PMLR, 2016, pp. 1717–1726.
- [24] Z. Jin, J. Li, L. Liu, T. D. Le, B. Sun, and R. Wang, “Discovery of Causal Rules Using Partial Association,” in *Proceedings of International Conference on Data Mining*, 2012, pp. 309–318.
- [25] J. Li, T. D. Le, L. Liu, J. Liu, Z. Jin, and B. Sun, “Mining causal association rules,” in *Proceedings of IEEE International Conference on Data Mining Workshops*, 2013, pp. 114–123.
- [26] M. Alharbi and S. Rajasekaran, “Conjunctive combined causal rules mining,” in *Proceedings of IEEE International Symposium on Signal Processing and Information Technology*, 2015, pp. 28–33.
- [27] M. Alharbi and S. Rajasekaran, “Disjunctive combined causal rules mining,” in *Proceedings of IEEE International Symposium on Signal Processing and Information Technology*, 2015, pp. 40–45.
- [28] C. Xiong, J. Shapiro, J. Hullman, and S. Franconeri, “Illusion of causality in visualized data,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 1, pp. 853–862, 2019.
- [29] C.-H. E. Yen, A. Parameswaran, and W.-T. Fu, “An Exploratory User Study of Visual Causality Analysis,” *Computer Graphics Forum*, 2019.
- [30] A. Kale, Y. Wu, and J. Hullman, “Causal Support: Modeling Causal Inferences with Visualizations,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 1150–1160, 2022.
- [31] W. S. Cleveland and R. McGill, “Graphical perception: Theory, experimentation, and application to the development of graphical methods,” *Journal of the American Statistical Association*, vol. 79, no. 387, pp. 531–554, 1984.
- [32] X. Xie, M. He, and Y. Wu, “CausalFlow: Visual Analytics of Causality in Event Sequences,” 2020.
- [33] J. Wang and K. Mueller, “Domino: Visual causal reasoning with time-dependent phenomena,” *IEEE Transactions on Visualization and Computer Graphics*, 2022.
- [34] R. Li, W. Cui, T. Song, X. Xie, R. Ding, Y. Wang, H. Zhang, H. Zhou, and Y. Wu, “Causality-based visual analysis of questionnaire responses,” *IEEE Transactions on Visualization and Computer Graphics*, 2023.
- [35] M. N. Hoque and K. Mueller, “Outcome-explorer: A causality guided interactive visual interface for interpretable algorithmic decision making,” *IEEE Transactions on Visualization and Computer Graphics*, 2021.
- [36] J. N. Yan, Z. Gu, H. Lin, and J. M. Rzeszotarski, “Silva: Interactively assessing machine learning fairness using causality,” in *Proceedings of CHI conference on human factors in computing systems*, 2020, pp. 1–13.
- [37] X. Xie, F. Du, and Y. Wu, “A Visual Analytics Approach for Exploratory Causal Analysis: Exploration, Validation, and Applications,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1448–1458, 2020.
- [38] J. Wang and K. Mueller, “The visual causality analyst: An interactive interface for causal reasoning,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 230–239, 2015.
- [39] B. Alsallakh, L. Micallef, W. Aigner, H. Hauser, S. Miksch, and P. Rodgers, “The State-of-the-Art of Set Visualization,” *Computer Graphics Forum*, vol. 35, no. 1, pp. 234–260, 2016.
- [40] M. T. Fischer, A. Frings, D. A. Keim, and D. Seebacher, “Towards a Survey on Static and Dynamic Hypergraph Visualizations,” in *Proceedings of IEEE Visualization Conference*, 2021, pp. 81–85.
- [41] X. Ouvrard, “Hypergraphs: an introduction and review,” 2020.
- [42] C. Vehlow, F. Beck, and D. Weiskopf, “The State of the Art in Visualizing Group Structures in Graphs,” in *Proceedings of Eurographics Conference on Visualization*, 2015, pp. 21–40.
- [43] M. Kaufmann, M. van Kreveld, and B. Speckmann, “Subdivision Drawings of Hypergraphs,” in *Proceedings of International Symposium on Graph Drawing*, I. G. Tollis and M. Patrignani, Eds., 2009, pp. 396–407.
- [44] B. Qu, E. Zhang, and Y. Zhang, “Automatic Polygon Layout for Primal-Dual Visualization of Hypergraphs,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 633–642, 2021.
- [45] S. Di Bartolomeo, A. Pister, P. Buono, C. Plaisant, C. Dunne, and J.-D. Fekete, “Six methods for transforming layered hypergraphs to apply layered graph layout algorithms,” *Computer Graphics Forum*, vol. 41, no. 3, 2022.
- [46] M. T. Fischer, D. Arya, D. Streeb, D. Seebacher, D. A. Keim, and M. Worring, “Visual analytics for temporal hypergraph model exploration,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 550–560, 2020.
- [47] A. Lex, N. Gehlenborg, H. Strobelt, R. Vuillemot, and H. Pfister, “Upset: visualization of intersecting sets,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1983–1992, 2014.
- [48] P. Valdivia, P. Buono, C. Plaisant, N. Dufournaud, and J.-D. Fekete, “Analyzing dynamic hypergraphs with parallel aggregated ordered hypergraph visualization,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 1–13, 2019.
- [49] G. Ausiello and L. Laura, “Directed hypergraphs: Introduction and fundamental algorithms—A survey,” *Theoretical Computer Science*, vol. 658, pp. 293–306, 2017.
- [50] G. Gallo, G. Longo, S. Pallottino, and S. Nguyen, “Directed hypergraphs and applications,” *Discrete applied mathematics*, vol. 42, no. 2–3, pp. 177–201, 1993.
- [51] J. Ren, H. Zhang, Z. Du, Y. Sun, H. Hu, and X. Zhu, “Weighted-Directed-Hypergraph-Based Spectrum Access for Energy Harvesting Cognitive Radio Sensor Network,” *IEEE Access*, vol. 8, pp. 68 570–68 579, 2020.
- [52] Y. Sun, Z. Du, Y. Xu, Y. Zhang, L. Jia, and A. Anpalagan, “Directed-hypergraph-based channel allocation for ultradense cloud D2D communications with asymmetric interference,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7712–7718, 2018.
- [53] S. Ranshous, C. A. Joslyn, S. Kreyling, K. Nowak, N. F. Samatova, C. L. West, and S. Winters, “Exchange pattern mining in the bitcoin transaction directed hypergraph,” in *Proceedings of International Conference on Financial Cryptography and Data Security*, 2017, pp. 248–263.
- [54] B. Wang, Y. Sun, T. Q. Duong, L. D. Nguyen, and N. Zhao, “Security enhanced content sharing in social IoT: A directed hypergraph-based learning scheme,” *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 4412–4425, 2020.
- [55] W. Liu, Y. Zheng, S. Chawla, J. Yuan, and X. Xing, “Discovering spatio-temporal causal interactions in traffic data streams,” in *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2011, pp. 1010–1018.
- [56] X. Luo, J. Peng, and J. Liang, “Directed hypergraph attention network for traffic forecasting,” *IET Intelligent Transport Systems*, vol. 16, no. 1, pp. 85–98, 2022.
- [57] J. Han, B. Cheng, and X. Wang, “Two-phase hypergraph based reasoning with dynamic relations for multi-hop kbqa,” in *Proceedings of International Conference on International Joint Conferences on Artificial Intelligence*, 2021, pp. 3615–3621.
- [58] L. H. Tran and L. H. Tran, “Directed hypergraph neural network,” *arXiv preprint arXiv:2008.03626*, 2020.
- [59] A. P. Volpertesta, “Hypernetworks in a directed hypergraph,” *European Journal of Operational Research*, vol. 188, no. 2, pp. 390–405, 2008.
- [60] T. Von Landesberger, A. Kuijper, T. Schreck, J. Kohlhammer, J. J. van Wijk, J.-D. Fekete, and D. W. Fellner, “Visual analysis of large graphs: state-of-the-art and future research challenges,” in *Computer graphics forum*, vol. 30, no. 6, 2011, pp. 1719–1749.

- [61] B. Lee, C. Plaisant, C. S. Parr, J.-D. Fekete, and N. Henry, "Task taxonomy for graph visualization," in *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*, 2006, pp. 1–5.
- [62] M. Brehmer and T. Munzner, "A multi-level typology of abstract visualization tasks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2376–2385, 2013.
- [63] E. Cuenca, A. Sallaberry, F. Y. Wang, and P. Poncelet, "Multistream: A multiresolution streamgraph approach to explore hierarchical time series," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 12, pp. 3160–3173, 2018.
- [64] H. Xu, Y. Zhen, and H. Zha, "Trailer generation via a point process-based visual attractiveness model," in *Proceedings of the International Conference on Artificial Intelligence*. AAAI Press, 2015, p. 2198–2204.
- [65] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.
- [66] M. Noto and H. Sato, "A method for the shortest path search by extended dijkstra algorithm," in *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, vol. 3, 2000, pp. 2316–2320.
- [67] S. Di Bartolomeo, M. Riedewald, W. Gatterbauer, and C. Dunne, "STRATISFIMAL LAYOUT: A modular optimization model for laying out layered node-link network visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 324–334, 2021.
- [68] S. Kumar, W. L. Hamilton, J. Leskovec, and D. Jurafsky, "Community interaction and conflict on the web," in *Proceedings of the 2018 World Wide Web Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2018, pp. 933–943.
- [69] Mann–Whitney Test. New York, NY: Springer New York, 2008, pp. 327–329.
- [70] G. Barnard, "Significance tests for 2×2 tables," *Biometrika*, vol. 34, no. 1/2, pp. 123–138, 1947.



Zihao Zhu is a master student at the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China. He received the B.Sc. in Information and Computational Science in College of Science, Zhejiang University of Technology. His main research interests are information visualization, visual analytics, and mathematical optimization.



Wang Xia is currently a PhD student at the college of computer science and technology, Zhejiang University of Technology, Hangzhou, China. Her research interests mainly include computer vision, information visualization and visual analytics.



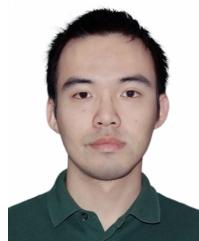
Baofeng Chang received B.Sc of automation at the College of Information Engineering, Zhejiang University of Technology. He is working toward Ph.D. degree in computer science and technology both from Zhejiang University of Technology. His main research interests are urban visualization, temporel sequence visualization, and graph visualization.



Sujia Zhu is currently a PhD student at the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China. Her major research interests include visual analysis, causality analysis, temporel sequence visualization, and high-dimensional data visualization.



Jingwei Tang is currently a Ph.D. student in the College of Computer Science and Technology, Zhejiang University of Technology, HangZhou, China. He received his B.E. degrees in communication engineering from Zhejiang University of Technology. His main research interests are data mining and information visualization.



Guodao Sun is an associate professor at the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China. He received the B.Sc. in computer science and technology, and Ph.D. degree in control science and engineering both from Zhejiang University of Technology. His main research interests are urban visualization and visual analytics of social media.



Ronghua Liang received the PhD degree in computer science from Zhejiang University in 2003. He was a research fellow at the University of Bedfordshire, United Kingdom, from April 2004 to July 2005 and as a visiting scholar at the University of California, Davis, from March 2010 to March 2011. He is currently a professor of computer science and dean of the College of Computer Science and Technology, Zhejiang University of Technology, China. He has published more than 50 papers in leading international journals and conferences including IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Visualization and Computer Graphics, IEEE Information Visualization, IJCAI, AAAI. His research interests include visual analytics, computer vision, and medical visualization. He is a member of the IEEE.



Yue Shen is a master student at the College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, China. She obtained the B.Sc. in software engineering. Her research interests include causality analysis and data visualization.