Causality-based Visual Analytics of Sentiment Contagion in Social Media Topics

Renzhong Li, Shuainan Ye, Yuchen Lin, Buwei Zhou, Zhining Kang,
Tai-Quan Peng, Wenhao Fu, Tan Tang, and Yingcai Wu

Abstract—Sentiment contagion occurs when attitudes toward one topic are influenced by attitudes toward others. Detecting and understanding this phenomenon is essential for analyzing topic evolution and informing social policies. Prior research has developed models to simulate the contagion process through hypothesis testing and has visualized user–topic correlations to aid comprehension. Nevertheless, the vast volume of topics and the complex interrelationships on social media present two key challenges: (1) efficient construction of large-scale sentiment contagion networks, and (2) in-depth explorations of these networks. To address these challenges, we introduce a causality-based framework that efficiently constructs and explains sentiment contagion. We further propose a map-like visualization technique that encodes time using a horizontal axis, enabling efficient visualization of causality-based sentiment flow while maintaining scalability through limitless spatial segmentation. Based on the visualization, we develop CausalMap, a system that supports analysts in tracing sentiment contagion pathways and assessing the influence of different demographic groups. Furthermore, we conduct comprehensive evaluations——including two use cases, a task-based user study, an expert interview, and an algorithm evaluation——to validate the usability and effectiveness of our approach.

Index Terms—Social Media, Causal Analysis, Visual Analytics

1 Introduction

Public sentiment toward social media topics critically informs social decision-making processes, from marketing strategy formulation [39] to public policy design [6]. The evolution of such sentiment is dominated by sentiment contagion [25], the propagation of sentiment between topics. For instance, supportive sentiment toward a protest can turn into passive sentiment toward other topics, particularly among frequent social media users [23]. Exploring the pathway of this contagion helps analysts find the origin of sentiment, understand the affected people and the evolution of topics. Such insights empower policymakers to formulate specific decisions to control negative sentiment development or promote the growth of positive sentiment around strategic topics.

Prior research primarily develops models [2, 16, 43] that construct user-level networks to depict the contagion process within a limited number of topics. However, the immense and multi-granular topics on social platforms pose two principal challenges to current approach. **First, constructing the contagion network that encompasses thousands of topics is inefficient.** Existing methods [6, 23] depend on hypotheses of network structure and propagation models [22, 30, 53], requiring enumeration of the hypotheses and statistical testing on each topic to determine the correct structure. Although accurate, high-volume topics cause an exponential increase in hypotheses, making it infeasible to find the network structure. **Second, analyzing sentiment contagion within complex networks is difficult.** Comprehending the sentiment contagion necessitates an understanding of features that lead to contagion, such as the transformation of sentiment between topics and the roles different demographic groups play [25]. Existing work has developed visualization methods to help analysts either explore the topic evolution [9, 11, 37] or identify characteristics of user behaviors [27]. Although promising, the spreading of sentiment requires a perspective on both topics and their impacts on different user types. This kind of analysis that simultaneously considers the interplay between sentiment, topic, and user as a triad is still under-explored.

In this work, we propose a visual analytics approach for causality-based sentiment contagion analysis across immense topics, users, and posts. To address the first challenge, we define a hierarchical Bayesian network for causality-based sentiment contagion modeling and present an efficient workflow. The workflow begins with transforming social media data into structured tabular representations of user sentiment per topic. Then, we apply the Fast Greedy Equivalence Search (F-GES) algorithm [41] to fit the contagion network from this tabular data. The temporal and hierarchical constraints specific to the sentiment contagion scenario are incorporated to enable efficient fitting.

For the second challenge, we propose CausalMap, a visual analytics system supporting three-stage exploration. The first stage enables hierarchical exploration of the contagion network through a scalable map-based visualization. Unlike traditional map visualizations [9, 10] that rely on animation to depict temporal progression, our approach encodes time on the x-axis. A comparative user study (Sec. 6.2) showed this design conveys more information with less interaction.

The second stage allows analysts to understand the details of a specific sentiment contagion pathway. Based on the causal model, we compute sentiment transformations and assess the importance of various user demographic groups. The system then presents an interpretation view to visualize the transformation and to quantify the impact of different user types on this pathway. In the third stage, our system enables what-if analysis of sentiment contagion. Analysts can intervene by either modifying sentiment distributions within a topic or altering the network structure, simulating downstream impacts. To evaluate our system, we conducted two case studies to demonstrate its usability, followed by an algorithm evaluation, an expert interview and a comparative user study that provided in-depth insights into the system and its visualization designs. The contributions are as follows:

- We propose a causality-based method for constructing hierarchical sentiment contagion networks in social media topics.
- We introduce a map-based visualization to efficiently visualize sentiment contagion networks and ensure scalability.
- We develop a visual analytics system to assist users in exploring, interpreting, and intervening the sentiment contagion, which is evaluated through a user study and an expert interview.

2 Related Work

This section reviews prior work on sentiment contagion, topic-based visualizations, and causal analysis that are related to our study.

- *R. Li, S. Ye, Y. Lin, B. Zhou, Z. Kang and Y. Wu are with the State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China. E-mail: {renzhongli, sn_ye, yuchenlin, zhoubuwei, kang264, ycwu}@zju.edu.cn.*
- *T. Peng is with the Department of Communication, Michigan State University, East Lansing, United States. E-mail: pengtaiq@msu.edu.*
- *W. Fu is an independent researcher. E-mail: f-wenhao@foxmail.com.*
- *T. Tang is with the Laboratory of Art and Archaeology Image, Zhejiang University, Hangzhou, China. E-mail: tangtan@zju.edu.cn.*
- *Tan Tang and Yingcai Wu are the co-corresponding authors.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

2.1 Sentiment Contagion and Topic-based Visualization

Traditional research on sentiment contagion explores how individuals' sentiments are influenced by others. Sentiment contagion manifests in two forms [25]: simple contagion, where individuals are influenced after single exposure [15], and complex contagion, requiring specific conditions to take effect. Research on online social media platforms mainly focuses on conditions for sentiment spread across topics and demographic group contributions via complex contagion [6]. For example, Yue et al. unveiled that positive sentiments regarding the European Parliament Elections spread more effortlessly in German compared to English [57]. While effective for limited topics, these approaches struggle with the vast number of topics on social media. In this work, we introduce a hierarchical Bayesian network [40] to model sentiment contagion between topics and employ a causality-based fitting algorithm to rapidly determine the network in a data-driven manner.

Topic-based visualizations [7, 8, 17, 18, 20, 26, 48] provide foundations for analyzing sentiment contagion as inter-topic relationships. These methods fall into two categories: relevance-based and directed-based. The relevance-based methods (e.g., HierarchicalTopics [18], TopicPanorama [37]) explore the undirected relevance of static topics. Directed-based approaches explore directed temporal relationships among topics. For example, ThemeRiver [26] introduced a river-based visualization design, with subsequent designs [13, 44] adopting this flow-like approach to track topic changes across different time segments. R-Map [10] employs a map-like visualization to depict the tree-like retweet structure under specific topics. These offer valuable inspiration for directed contagion networks. However, these works typically focus one or two elements within the triplet of topic, user, and sentiment, constraining comprehensive contagion analysis. To address this, we introduce a visualization system, CausalMap. The system begins by utilizing a time-coded map to visualize the sentiment contagion between topics. It then employs Sankey diagrams to illustrate the sentiment within topics and the sentiment transformation between topics. Finally, histograms are used to demonstrate the impact of different types of users on the sentiment transformation. These designs enable users to explore the interrelationships among topics, users, and sentiments.

2.2 Causal Discovery and Visualization

Sentiment contagion exhibits significant causal relationships across topics (refer to Sec. 4.1). Therefore, we introduce causality discovery and visualization methods which can be applied to assist in solving and exploring sentiments contagion network.

While randomized experiments [40] provide gold-standard causality verification, their cost limits application to few entities. The more widely accepted methods are divided into two main categories: constraints-based and score-based. Constraint-based algorithms initiate with fully-connected graphs and prune edges via conditional independence tests. This method, while computationally slower, provides accurate results and is commonly utilized for smaller datasets [45, 46]. In contrast, score-based algorithms incrementally build graphs by optimizing scores (BIC). These algorithms are faster and applied to large-scale datasets [34, 51]. However, the task of tailoring the model to the requirements of sentiment contagion scenarios is still a problem.

The visualization field has conducted in-depth research [46, 51, 58] on how to visualize causal networks. Initially, lots of empirical studies [29, 52, 55] were conducted to study the presentation of causal networks, with Bae et al. [3] highlighting node-link diagrams as an excellent choice. To visualize larger datasets, recent work improving scalability through specialized layouts [51] and subgraph division [34]. Yet, visualizing causal graphs with hundreds of topics in social media datasets remains an unresolved challenge. Concurrently, sentiment contagion analysis necessitates the integration of users, topics, and sentiments within the model and visualization framework to assist analysts in their analytical processes. In response to these requirements, our study initially presents a methodology for fitting a sentiment contagion network from users' sentiment distribution toward specific topics. Subsequently, we introduce a visualization approach based on map metaphors, designed to effectively display large-scale causal graphs. Finally, we complement the map view with auxiliary views that facilitate analysts in comprehending the details in the network.

3 Informing the Design

In this section, we introduce the problem of sentiment contagion and identify the design requirements that guide our design.

3.1 Data Description

The input for this problem consists of two types of data, namely, post data and user data. **Post data** stands for a post posted by a user. Each post is defined by its post ID, and related props, such as post time, text, etc. **User data** stands for a platform user. Each user is identified by his user ID, and related demographic info, such as position, age, etc.

3.2 Problem Formulation

We collaborated closely with four domain experts (EA, EB, EC, and ED) for one year to understand the domain problem, workflow, and distill design requirements. EA and EB are experienced analysts with over ten years of work in internet regulation. EC and ED are professors specializing in sentiment analysis and social media visual analytics, respectively. We collected relevant surveys [57] and held weekly meetings with our experts to acquire knowledge about the fundamental principles and typical patterns of sentiment contagion. We then focused on exploring experts' interests, analyzing workflows and practical tasks to specify stages for contagion analysis.

During the meetings, experts stated that the contagion analysis mainly consists of two core tasks. From a macro perspective, the objective of contagion analysis is to identify several key pathways of public sentiment development within vast datasets and to summarize the characteristics of these pathways in terms of sentiment transformation, user demographics, and topic evolution. From a micro perspective, contagion analysis focuses on specific topics of interest, such as those with negative public sentiment. It seeks to trace the origins of sentiment, understand the conditions that facilitate their spread, and address the sentiment at its source to manage the public discourse effectively.

Guided by these two core tasks, we propose a three-step visual analysis procedure. **Explore:** Initially, the primary task is to extract the propagation network from the social media dataset and identify several significant contagion pathways. **Interpret:** Subsequently, the focus shifts to understanding critical sentiment contagion pathways within the network, such as those with significant influence or relevance to important topics. It is essential to summarize these pathways based on patterns such as transformation in sentiment and variations across user demographic profiles. **Intervene:** Based on the comprehension of the pathway and the network, analysts can explore various intervention strategies for the contagion network. The outcomes of the strategies can help analysts control or manage certain phenomena.

3.3 Requirement Analysis

Based on these three stages and the feedback from the experts, we summarized six requirements to support sentiment contagion analysis.

Exploration Stage.

R1: Obtain an overview of the sentiment contagion network. At the beginning of the analysis, the system is tasked with deriving the sentiment contagion network from the dataset and presenting an overview. This overview should encompass two critical pieces of information: 1) Nodes: These refer to the topics and the associated sentiments of the users toward the topic. 2) Edges: These refer to the connections between topics, with a particular focus on how sentiment spreads from one topic to another. This overview should enable analysts to understand the dataset and find the main pathways for further analysis.

R2: Support multi-level contagion network analysis. Social media datasets are often vast, encompassing an uncountable number of posts. These posts, in turn, form a multitude of topics in a hierarchical structure. To meet the needs of experts, it is imperative to provide them with an analytical approach that is both progressive and scalable. This progression should not only facilitate a transition from the network to a specific pathway (R1) but also allow for a transition from a higher-level topic network to a lower-level topic network.

Interpretation Stage.

R3: Visualize the sentiment transformation within a pathway. During the exploration stage, analysts could select pathways at different levels for more detailed analysis. The first aspect that requires visualization within theses pathways is the transformation process of sentiment, which unfolds a narrative and shows important patterns.

R4: Visualize the influence of different types of users. Following the transformation of sentiment, the influence of different demographic groups on the pathway should also be presented. This pattern tells the narrative from the user aspect. For example, the attitudes of middle-aged individuals toward a presidential election may be more influenced by related policies, whereas younger individuals might be swayed by different factors. This information is instrumental in helping analysts gain a deeper understanding and make informed decisions accordingly.

Intervention Stage.

R5: Explore policy based on propagation structure. The first intervention policy involves disrupting the spread of sentiment from a specific topic. This policy is the most widely applied as it can swiftly and directly eliminate the influence of a particular topic. However, this policy may not always be effective; emotions from other nodes may bypass the blocked node, or the targeted node might not be the source of the sentiment that needs to be controlled. Analysts need to conduct what-if analysis to explore which nodes, if blocked, would be the most straightforward and efficient in curbing the spread.

R6: Explore policy based on sentiment distribution. While blocking certain topics is a common regulatory measure, it may provoke a public sentiment backlash and carry significant risks. The second frequently utilized strategy involves modifying the internal sentiment distribution of a topic through the release of information or other means [28]. This approach is effective but can be costly to implement and requires careful selection of the topics for what-if analysis. Conducting pre-experiments on existing data is a necessary and effective method to refine strategies. It allows for the assessment of potential impacts and the optimization of policy application before widespread deployment.

4 Causality-based Contagion Network

In this section, we first introduce the hierarchical Bayesian network model to fit the sentiment contagion network between topics. We then present a causality-based method to calculate the network using social media datasets and discuss how the model can meet the analysis requirements (R4-R6) for sentiment contagion networks.

4.1 Sentiment Contagion Network

The sentiment contagion network describes how the sentiments in the topics affect each other by two main components: nodes and edges. Specifically, a node denotes a topic T, aggregated from multiple posts. A node pair $\{T_A, T_B\}$ is connected by an edge if sentiment distribution in one node T_A affects the sentiment distribution in the other node T_B:

$$E(Distribution(T_B)|Distribution(T_A)) \neq E(Distribution(T_B))$$

where $Distribution$ refers to the distribution of user sentiment in the topic T and E denotes expectation. Moreover,

$$E(Sentiment(U,T_B)|Sentiment(U,T_A)) \neq E(Sentiment(U,T_B))$$

indicates the sentiment expressed by a user U in T_A influences his/her sentiment in node T_B. Beyond the network structure, the contagion network exhibits a hierarchical organization reflecting the multi-level nature of topics (R2), where nodes at higher levels contains a collection of lower-level topics and their connections.

To represent this structure, we propose a Bayesian network-based approach. The influence between topic sentiment distributions is captured through conditional probability distributions (CPDs) in the Bayesian network. Specifically, for a node T in any level, the sentiment of a user U toward a topic T_A is influenced by the sentiments expressed by the user U toward other topics T_i that are directed to T_A in the Bayesian network. For instance, in the simplest case where parental influences are independent, the sentiment can be expressed as:

$$E(Sentiment(U,T_A)) = \sum E(Sentiment(U,T_i)) * k_{transform}(T_A|T_i)$$

$k_{transform}$ represents the conditional probability, quantifying contagion likelihood between topics. To determine the connections across levels, we set up the inter-layer constraints. Specifically, we connect topic T_A to topic T_B only if the parent topic of T_A is connected to the parent topic of T_B, or T_A and T_B have the same parent topic. This modeling approach offers two advantages. Firstly, this model leverages a hierarchical network-based definition, shifting the focus of sentiment contagion analysis from individual users to topics. This transition enhances the scalability of the analysis. Secondly, by framing the sentiment contagion network as a Bayesian network, we transform the problem into one that can be efficiently solved using established methods.

4.2 Network Fitting

Based on the hierarchical Bayesian model, we provide a workflow to derive the sentiment contagion network. The workflow consists of two main stages: data preprocessing and causality-based fitting.

Data preprocessing. This process (Fig. 1A) consists of two steps: topic discovery and sentiment analysis within the identified topics.

First, we calculate the topic structure from posts. There are two common types of topic modeling methods: one based on statistical distribution, and the other [24] based on embedding. Embedding-based methods demonstrate excellent semantic understanding ability and perform well on long documents. However, these methods are computationally intensive and require the target dataset to be similar to pretraining dataset. Statistical methods can be further divided into matrix factorization [12] and LDA based approaches [4]. Matrix factorization methods often require intensive computations and are less interpretable. Given that posts are short texts with specific keywords like hashtags, unlike the long synonym-rich documents that embedding methods excel at analyzing, we opted for the LDA-based method. This choice aligns with our priorities of speed, stability, and interpretability.

We employ the elbow method [36] to determine the total number of topics. To ensure effective visualization in the subsequent steps, we impose a constraint on the number of subtopics in a topic. Since causal networks suitable for visual analysis typically contain around ten nodes [19, 34], we set the number of clusters to ten when the recommended number exceeds this limit. Users can also adjust this limit based on dataset characteristics.

In addition to directly limiting the number of LDA topics, visualization techniques such as bundling and algorithmic techniques like topic aggregation can also be used to address this issue. However, limiting LDA topic quantity is not about deleting excess clusters but is similar to visual bundling and topic aggregation. Adjusting the number of topics in LDA is equivalent to breaking down and merging the excess clustering content according to distribution features and previous clusters. Thus, limiting the number of LDA topics is a form of topic aggregation and a prerequisite for visual bundling. Therefore, we haven't added more complex solutions for aggregation or bundling. The drawback is also discussed in the discussion section.

By using the post clusters generated by LDA as input for multiple LDAs, we obtain a hierarchical topic tree. Specifically, in the first round, we run the algorithm on all posts. After obtaining several clusters, we use the posts in each cluster as input for the next round of LDA. This process is iterated for a specific number of rounds. In our system, we iterate three layers. As users explore deeper, lower-level clusters can be generated on demand. For each topic, we generate summaries via a Large Language Model (LLM) using keywords from LDA and the top 100 posts. As a fundamental natural language processing (NLP) task, common approaches [31] for summarization include traditional machine learning and deep learning approaches. These methods vary in stability, speed, and applicability. However, since the emergence of LLMs, the accuracy of these tasks on various datasets has been significantly improved compared to traditional methods and other deep learning approaches [1, 54]. For this reason, we employ LLMs to complete the NLP tasks within this workflow. The relevant prompts for these tasks are provided in the support material.

In the second step, we detect the sentiments of posts to obtain the sentiment distribution under each topic. We use the summary of topic in the first step to determine the sentiment and the degree of the sentiment of its posts toward the topics using LLMs. Similar to this survey [57], we use three values to mark the sentiment expressed by users: positive, neutral, and negative. Additionally, we employ a variable ranging from 0 to 1 to denote the degree of sentiment expressed in a post.

Causality-based fitting. There are a lot of causal discovery methods that can fit a single-layer Bayesian network from sentiment distribution

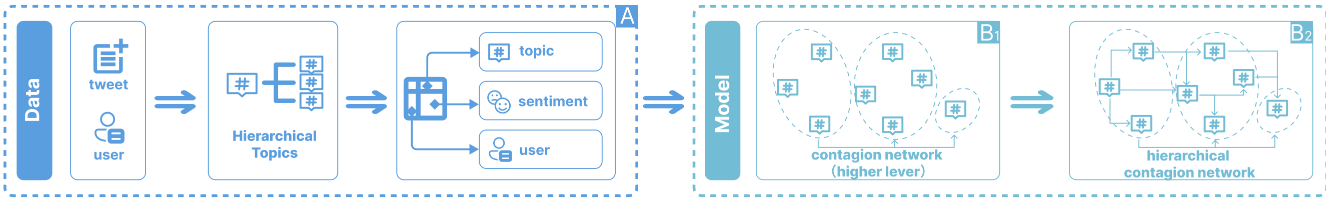

Fig. 1: The construction of the network involves three stages: (A) transforming social media data, which includes posts and users, into tabular data that represents the sentiment of users across various hierarchical levels of topics, (B_1) solving a single-layer Bayesian network from the tabular data of highest-level topic, and (B_2) leveraging the constraints from the upper-layer network to calculate the lower-layer networks using the tabular data.

under topics (Fig. 1B_1). As discussed in Sec. 2.2, causal discovery algorithms are mainly divided into two categories: constraint-based and score-based methods. Constraint-based methods use conditional independence tests to infer causal relationships. However, in social media datasets, the hierarchical and temporal nature gives a large amount of prior knowledge. Most constraint-based methods demand that the input prior knowledge cannot conflict with the independence tests. Considering robustness, we choose the score-based approach, which searches for the highest-scoring causal structure. The representative score-based method for observational data is GES. Compared to GES, F-GES enhances its speed through parallel processing and caching. Additionally, F-GES allows for missing data in samples. Many studies [34, 51] have adopted F-GES, which demonstrates its stability. Therefore, we select the F-GES method for causality-based fitting.

Within each layer of the hierarchical network, we employ the F-GES algorithm [41] to fit the Bayesian structure. To form the hierarchical contagion network, the Bayesian structure should maintain alignment (refer to Sec. 4.1) between different layers. For example, if topic T_A causes topic T_B, any subtopic of T_A should not be an effect of subtopic of T_B. Additionally, since social media datasets are inherently sequential, the modeling should adhere to temporal relationships. The past topics should not be the effect of future topics. Therefore, we propose two constraints for F-GES to satisfy these two conditions. Specifically, when adding a causal edge from topic T_A to topic T_B in F-GES, these topics should satisfy: 1) the start time of topic T_A is earlier than the start time of topic T_B, and 2) the parent node of topic T_A has a causal edge to the parent node of topic T_B. Alternatively, topic T_A and topic T_B share the same parent node. These constraints allow us to derive a Bayesian network between topics at each layer (Fig. 1B_2).

4.3 Network Analysis

This section introduces the Bayesian-based methods for finding contagion pathways and supporting demographic factor reasoning and what-if analysis (R4-R6).

The contagion pathway. After examining the global overview, analysts often need to delve into detailed contagion pathways, which is a linear sub-network as mentioned in R2. We propose a method to mine highly correlated contagion pathways. We select these pathways based on the Bayesian Information Criterion (BIC) score, which serves as a metric to assess the certainty of causality [51]. We use the decrease in the BIC score after removing an edge as a measure of certainty. We select the pathway by considering the average certainty over consecutive edges. Ultimately, we can calculate the composition and the scoring of each pathway, which are then employed for further visual analysis.

The impact of demographics. Within a specific contagion pathway, the influential demographic groups are essential to analyze (R4). To address this, we first establish criteria for defining the influence of certain demographic groups. Intuitively, a demographic group is considered influential if the removal of it causes the causal relationships within the pathway to diminish or disappear. Drawing from this concept, we utilize the impact of removing a demographic group on the BIC score as a measure of their importance within the Bayesian network. Based on this approach, we obtained the influence of different demographic groups and used it for subsequent visual analysis.

What-if analysis. Intervening into the contagion modeling process and conducting counterfactual evaluations is crucial for strategy development (R5, R6). We use model-based what-if analysis to perform the counterfactual evaluations. We employ the CPD for conducting what-if analysis and accessing what-if outcomes based on the Bayesian network structure. Corresponding to the requirements, we support two types of what-if analyses: structure-based (R5) and distribution-based (R6). In structure-based analysis, users can cut off the causal connections between a topic and the rest of the network. By using the CPD method, we observe how the removal of one topic's influence modifies the sentiment distribution within the affected topics. This modification corresponds to scenarios such as: "If the impact of the pandemic is disregarded, would people still support Biden's election?". In distribution-based analysis, users can force the sentiment distribution under a topic to be a specific value. We then calculate the effect and change of this node's distribution on its resultant nodes, ultimately reflecting all successor nodes. This modification corresponds to questions: "If people had a different attitude toward the epidemic, would they still support Biden's election?". These methods ensure analysts to complete what-if analyses in the intervention stage.

5 System Design

The sentiment contagion network output by our model is a complex hierarchical structure. To assist analysts in analyzing such network, we propose a map-based visualization system, CausalMap. In this system, we use orange red, green, and purple to represent positive, negative, and neutral sentiments, respectively (Fig. 2). Analysts can also customize these colors according to their analysis preferences. Our system comprises three views, namely, the overview, the interpretation view, and the post view. The analysis pipeline is aligned with the previously summarized three-step process: exploration-interpretation-intervention (refer to Sec.3.2). **In the exploration stage**, the overview empowers analysts to see the hierarchical network structure of sentiment contagion (R1, R2). High-weight contagion pathways (refer to Sec. 4.3) within the network are highlighted, enabling analysts to select pathways of interest for further interpretation. **In the interpretation stage**, analysts can obtain the details of a contagion pathway that they picked from the overview. The details include the sentiment transformation (R3) and the influence of different demographic groups (R4). **In the intervention stage**, analysts are empowered to directly test the impact of two types of strategies (R5 and R6). The influence of these strategic interventions will be visualized in both the overview and the interpretation view. During the whole workflow, analysts can see the posts corresponding to the current stage in the post view on demand.

5.1 Overview — Exploration Stage

To understand the overall situation (R1, R2), we propose a map-based visualization (Fig. 5A) to present the overview of contagion network, which shows the evolution of topics within a hierarchical network.

5.1.1 Contagion Network Layout

The sentiment contagion solved by the Bayesian model consists of both hierarchical structure and network structure. To support the overview

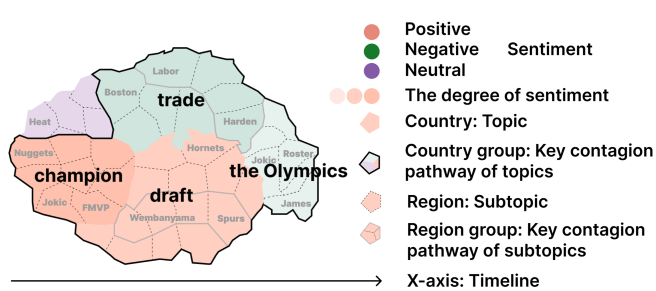

Fig. 2: Visual encodings of the overview in CausalMap, which is constructed based on the sentiment contagion network.

of both structures, we first conceptualize topics at different hierarchical levels as distinct layers within a map metaphor. For instance, topics at the top level are visualized as countries, with each country containing multiple regions, representing sub-topics. We arrange the lowest-level topics along the horizontal axis according to their begin time (Fig. 3A). Next, we employ a force-directed optimization technique to refine the vertical axis coordinates to position topics with direct influence relationships closer to each other (Fig. 3B). Finally, we determine the shape of lowest-level topics (Fig. 3C) and construct the map (Fig. 3D), ensuring that the visualization accurately reflects the hierarchical structure and network structure of the sentiment contagion network.

Layout procedure. To effectively present both the hierarchical and network structures with clarity, we employ a map-based visualization approach. Discussions on alternative visualizations are reserved under "Alternative Design". When determining the positions of various topics within the map, it is crucial to convey to analysts the propagation relationships of sentiment contagion. These relationships encompass two crucial pieces of information: the direction of contagion and the strength of the influence. Our layout algorithm is designed with the principle of prioritizing the visualization of these two aspects.

We first arrange the lowest-level topics along the x-axis of the map based on their timestamps. This arrangement takes into account the temporal constraints (refer to Sec. 4.2) preserved during the modeling process, ensuring that the contagion flows are depicted in a consistent, left-to-right direction. This enables analysts to quickly grasp the evolution of topics and aligns with their mental models, which is commonly found in causal visualization [34, 51] and flow-based social media visualization [26]. To determine the y-axis positions of each topic, we follow the principle that topics should be spaced apart from each other, while related topics should be brought closer together [9].

In our scenario, the influence relationships and the similarity between topics are factors to consider. Based on these two considerations, we propose a force-directed optimization enhanced with three types of forces. Specifically, all topics exert a repulsive force on each other. Topics connected by edges exert an attractive force on each other, with the force magnitude increasing with the certainty of the causality. Furthermore, topics that share the same ancestor topics are attracted to each other, with the strength diminishing as the level of the ancestor topic in the hierarchical structure increases. During the layout optimization process, nodes have degree of freedom only on the y-axis as the x-axis has been encoded for the timestamps. After the layout has stabilized, we obtain the positions of the lowest-level topics. We employ an algorithm proposed by G-Map [21] to partition the map based on the spatial position of the lowest-level topics and draw the boundaries for their parent topics at each hierarchical level.

Alternative design. The contagion network is a time-aware hierarchical network. To visualize the network, we have considered three common structures frequently used in the field of public sentiment analysis: Flow, Node-link, and Map. Flow visualization is a common method for presenting time-aware social media data [49], effectively demonstrating the direction of sentiment flow. Flow visualizations often incorporate whitespace to highlight the separation and merging of flows across multiple levels. In contrast, for sentiment contagion networks, the analysis focus is on key contagion pathways formed by a few high-weight, consecutive edges. Therefore, extensive whitespace for encoding separation and merging is not necessary. From another perspective, our design is similar to flows that remove the whitespace and add clear boundary. Node-link diagrams are another standard visualization technique [5]. Each topic has a clear boundary, and every edge is explicitly drawn. However, in the lower-level contagion network, there are too many nodes and edges to visualize all the information using a node-link diagram. Additionally, the nodes do not facilitate smooth transitions between different levels of exploration. Map visualization inherently makes full use of the available space and clearly presents each topic. By attributing a temporal dimension to the x-axis of the map, we effectively illustrate the direction of sentiment flow.

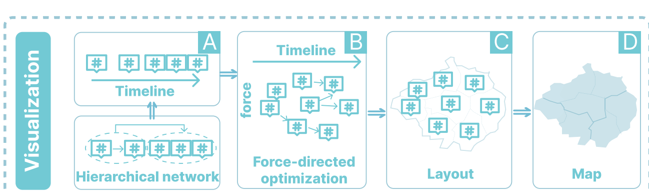

Fig. 3: The generation of overview involves four steps: (A) positioning the lowest-level topics along a temporal axis, (B) optimizing the y-axis placement of topics using force-directed graph techniques, (C) segmenting topics into map tiles, and (D) progressively merging the lowest-level map tiles to form a hierarchical map.

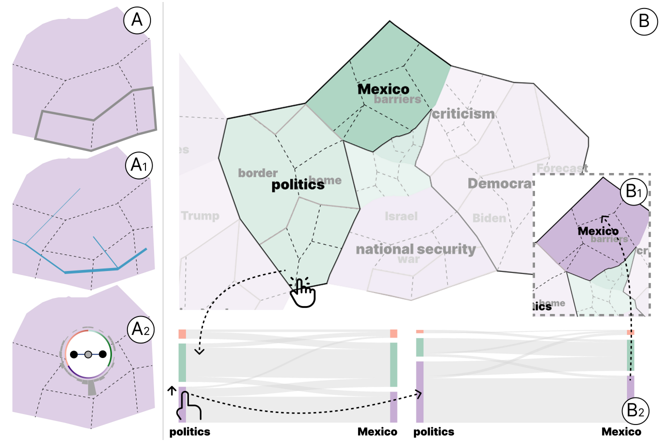

Fig. 4: (A) Design alternatives for contagion pathways. (B) What-if analysis via interaction with the Sankey diagram, results displayed in the Sankey diagram and overview.

5.1.2 Contagion Map

Based on the layout, we introduce a visualization to provide a visual summary of the hierarchical sentiment contagion network. As shown in Fig. 2, each country corresponds to a topic. We use geographical concepts as metaphors for different levels of topics, using the boundaries to distinguish between different topics. The color of each country represents the major sentiment of that area. The lightness of the color indicates the degree of the sentiment, where the darker color signifies higher degree. Furthermore, we use keywords from the LDA akin to city names to display the details of a topic.

Experts are more interested in the propagation pathways that encompass multiple edges (refer to Section 3) than single edges. The dense nature of the edges will also impair the user experience by cluttering the visualization. Therefore, we opt not to visualize individual edges directly. We highlight boundaries to denote pathways. Discussions on alternative visualizations of boundaries are reserved under "Alternative Design". Furthermore, we highlight certain pathways within subtopics to illustrate their structure, allowing analysts to grasp the propagation framework of the next hierarchical level directly from the current layer. We utilize a line chart (Fig. 5A_3) at the bottom of the map to display the quantity of posts with different sentiments across various time intervals. These visualizations facilitate analysts in comprehending the sentiment transformation across time, previewing the structure in next level, and locating key contagion pathway for further analysis.

Alternative design. When considering the visualization of contagion pathways, we explored three methods: boundary-highlighting, river-based, and glyph-based (Fig. 4A). In addition to the boundary-highlighting method we ultimately used, we initially considered rivers (Fig. 4A_1) due to their natural link structure and common use as a map metaphor [47]. River width could indicate connection strength, and branching could show subtopic hierarchy. However, a key limitation is that rivers make it hard to isolate specific contagion paths for analysis. Experts found that they needed to identify a significant section of the river before further analysis. This interaction process is less efficient and contradicts their mental models. The second design we considered was a node-link glyph representation to illustrate the propagation substructure of subtopics. For example, the glyph in Fig. 4A_2 indicates the pathway structure and the sentiment distribution. However, this approach was discarded due to its space-consuming nature and lack of intuitiveness. Upon recognizing these issues, we opted for a direct approach of highlighting the boundaries involved in a pathway. We display the pathways one by one based on their score (refer to Sec. 4.3), ensuring that new pathway does not overlap with those already placed. Since the system does not display all pathways, we allow analysts to select any two topics to explore the pathway between them. This method streamlines the analytical process, allowing analysts to swiftly select the important or interesting pathways for further analysis.

5.2 Interpretation View — Interpretation Stage

After selecting a contagion pathway, experts need to understand the evolution further from the aspect of sentiment transformation and user demographics (R3, R4). Therefore, we design the interpretation view (Fig. 5C) that consists of a Sankey diagram to visualize the sentiment transformation among topics and a set of histograms to reveal the influence of users with different demographic information.

Sentiment transformation. Considering the contagion pathway as a linear structure, it is essential to focus on the sentiment transformation between adjacent topics. We utilize a Sankey diagram (Fig. 5C_1) to present the transformation. In the Sankey diagram, each level denotes the sentiment toward a topic in the pathway. Each level is divided into three groups including positive, neutral, and negative sentiment. The flows between levels symbolize the transformation of sentiment. For instance, if a user has a positive sentiment toward both Topic A and Topic B, the positive sentiment group of Topic A will be connected by a flow to the positive sentiment group of Topic B. A wider link indicates a greater number of individuals experiencing the same transformation. This enables analysts to quickly grasp the core sentiment flow.

Demographic information. The demographic subview (Fig. 5C_2) consists of two parts. The right part shows the impact of different demographic groups. The impact is calculated based on the influence of a particular demographic group on the causal structure of the pathway (refer to Sec. 4.3). Since the impact weight can be positive or negative, the y-axis of the histogram can be both positive or negative number. The attribute values are arranged from left to right in descending order of impact weight. The left part shows the variance of corresponding impact distribution. Analysts can easily discern which groups of users cause the sentiment contagion. For instance, as depicted in Fig. 5C, it is observable that for individuals from North America, sentiment tends to spread more readily along this contagion pathway. By hovering over a specific area of the histogram, the sentiment flow of users within corresponding type is highlighted on the Sankey Diagram (Fig. 5A_4). In left part, the attributes are also ordered from top to bottom based on their variance, with those exhibiting the highest variance at the top. In such a design, analysts can quickly grasp which attribute of users can divide users into different parts efficiently according to their impact toward the pathway. And analysts can also understand how the specific value of a attribute influences the spread of the sentiment.

5.3 What-if Analysis — Intervention Stage

Analysts need to control the sentiment with strategies based on what-if analysis (R5, R6). We propose two corresponding interactions: modification of sentiment distribution (R6) and network structures (R5).

Sentiment Distribution. Our system allows analysts to modify the sentiment distribution within a specific topic and observe the subsequent effects on both the pathway and the entire network. This capability allows for targeted inquiries, such as assessing the potential impact if a topic were to be influenced by a coordinated inauthentic behavior campaign or if certain responses were to be suppressed. Analysts can freely adjust the sentiment distribution for a particular topic within the Sankey diagram of the spread pathway view by dragging the height of the corresponding rectangles (Fig. 4B). Once the modification is complete, it is processed by the backend, which calculates the impact of the intervention and sends the results back to the frontend.

We provide three visual cues to indicate the intervention's impact. Firstly, the background color of topics in the overview related to the intervened topic will change (Fig. 4B_1), implying the shift in major sentiment of this topic. Next, the height of sentiment blocks corresponding to relevant topics in the Sankey diagram indicates the intervention effect on relative contagion (Fig. 4B_2). Finally, the glyphs in the post view, which display the distribution of sentiments within topics (Fig. 5B_1), will also change. We use darker colors to represent an increase in the number of people and lighter colors to represent a decrease (Fig. 6B).

Network Structure. The second interaction involves eliminating the influence of a topic (R6). Analysts can click on an icon next to the topic (Fig. 5C_1) to remove its influence in the Sankey Diagram. The subsequent processing and visualization update follow the same procedure as the interaction of sentiment distribution modification.

5.4 Post View

The post view in our system is designed to display posts within different topics. We first use a rose chart to illustrate the specific distribution of degree levels across different sentiments within the topic. At the center of the chart is a color-coded circular axis, segmented into red, green, and purple sections representing positive, negative, and neutral, respectively. Each section of the axis transitions from light to dark, indicating the degree level of sentiment. Each post is marked with its sentiment on the left. Considering that each topic contains a large number of posts, we have sorted them based on the sum of likes, comments, and retweets, allowing analysts to see the posts that have a broader impact first.

With this design, analysts can not only perform sentiment contagion analysis at a high level but also examine the most granular level of posts from individual users. This capability strengthens their analytical insights and allows them to cite concrete examples in their final analysis. The post view thus serves as a versatile tool that bridges the gap between macro-level sentiment trends and micro-level post details.

6 Evaluation

In this section, we first evaluate the design of CausalMap from two use cases to demonstrate its usability. Second, we compare CausalMap and D-Map [9] to show that our visual design are more effective for analyzing sentiment contagion. Third, we conduct an algorithm evaluation to test the practical performance of our algorithm. Finally, we carry out an interview to discuss experts' views on CausalMap.

6.1 Use Case

The subsequent use cases are based on datasets collected from social media platforms over 2023 to 2024 with "NBA" and "basketball" as the keyword and over 2020 to 2024 with "president" as the keyword. Each dataset contains more than 200,000 posts. We follow Bob, a sentiment analyst, to explore the propagation of sentiment within the datasets.

6.1.1 Case 1: Sentiment Contagion in NBA Dataset

Bob first selected the National Basketball Association (NBA) dataset from 2023 to 2024. He wanted to know what issues are currently capturing the attention of basketball fans. In the overview (Fig. 5A), Bob first noticed that the most recent topic was related to *the Olympics*, which was located on the right side of the map. Through the post view summary (Fig. 5B), he discovered that the current discussions revolved around how the Paris Olympic basketball games would select players from NBA. Bob observed that the sentiment toward the Olympics was somewhat pessimistic. The highlighted contagion pathways related to *the Olympics* was: *champion* → *draft* → *trade* → *the Olympics*.

In the overview, he noticed that the sentiment of these four topics were predominantly positive for the first two and negative for the latter two. Additionally, the line chart (Fig. 5A_3) below indicated spikes in posts during July and October, which aligned with the timeline of this pathway. Bob clicked on the pathway for further exploration. The detailed sentiment transformation (Fig. 5C_1) was displayed in the Sankey diagram. It clearly illustrated the primary sentiment progression along this route: Positive → Positive → Negative → Negative. This indicated a shift from supporting the championship, to agreeing with the draft, to disapproving of trades, and finally feeling pessimistic about the Olympic Games. Moreover, the histogram in the interpretation view (Fig. 5C_2) categorized the influence of different demographics. There was a notable regional difference: North America closely mirrored the propagation pattern (Fig. 5C). Bob then examined posts under these four topics and found comments such as: "A strong center is really important", "Will Wembanyama be as strong as Jokic?", "Why aren't the Warriors trading for a center? They'll get demolished by the Nuggets!", and "JOKIC WILL REPRESENT SERBIA IN THE 2024

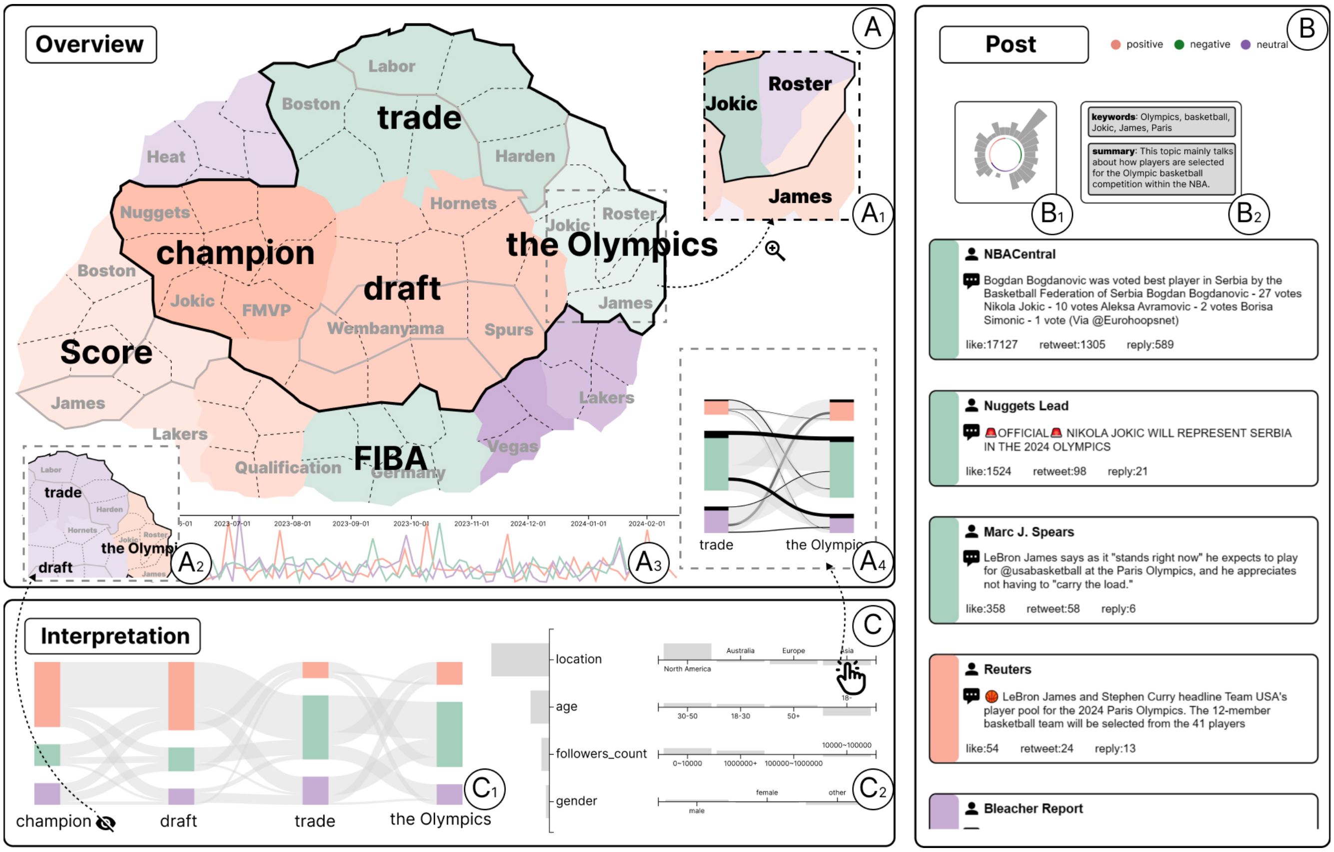

Fig. 5: (A) The overview displays the sentiment contagion network of the whole dataset. (B) The post view lists posts involved in different analysis stages. (C) The interpretation view provides a detailed sentiment evolution and impact of different demographic groups within a contagion pathway.

OLYMPICS". Bob hypothesized that this entire contagion might be driven by users' sentiment toward Jokic across different topics. He first zoomed in on *the Olympics* topic and discovered a subtopic focused on *Jokic* in the overview map, with a pessimistic sentiment (Fig. 5A_1). The summary of this subtopic indicated that it primarily discussed Jokic representing Serbia at the Olympics. Curious about how the sentiment would change if Jokic hadn't won the championship, Bob clicked on the icon next to the *champion* topic to exclude its influence on the subsequent topics. He noticed that the *trades* topic became neutral, and *the Olympics* topic turned positive in the overview map (Fig. 5A_2). This confirmed his hypothesis: Jokic's championship victory had a significant impact on NBA fans' perceptions.

6.1.2 Case 2: What-if Analysis in President Dataset

Following a process similar to Case 1, Bob explores the sentiment spreading pathway (Fig. 6C): *politics* → *Mexico* → *criticism* → *The Democratic*. Each of them is associated with a distinct public sentiment event: policies issued by the Biden administration, especially those concerning border policies; incidents at the U.S.-Mexico border; criticisms posts on relevant issues; and Biden's victory in the Democratic primaries. The major sentiments along this spreading chain were consistently negative and neutral throughout (Fig. 6A).

Insight 1: Bob aimed to explore how to transform neutral and negative sentiments related to the Democratic primaries into positive support for Biden. He initially attributed the neutral and negative sentiments to the public's skepticism toward policies related to the U.S.-Mexico border. Bob hypothesized that if the public more strongly supported Biden's policies, the sentiment along the chain would shift and finally lead to positive sentiment in *The Democratic*. However, after conducting a what-if analysis to increase positive sentiment in *politics*, he noticed only a marginal change. The negative sentiment in *Mexico* persisted and influenced attitudes toward the *The Democratic*. This led Bob to adjust the positive sentiment in *Mexico* alone, which had some effect but was not substantial. It indicated that there are additional sentiment pathways beyond attitudes toward U.S.-Mexico border events that spread to *The Democratic*. Upon adjusting both *Mexico* and *politics*, Bob noticed a positive shift in the public's attitude toward the Democratic Party (Fig. 6B). This revealed that *Mexico* and *politics* are two significant influencing factors. And *politics* had additional influence paths beyond *Mexico*, likely related to *national security* or other complex pathways, as observed in the overview. Therefore, Bob thought it was necessary to enhance public's comprehension of policies and address border issues simultaneously to get public support.

Insight 2: Bob noted that posts related to *Mexico* had been present on the map for an extended period. Aware that interventions on *Mexico* and *politics* could effectively influence public's support, Bob explored whether the timing of such interventions was crucial. By zooming in, he identified two temporally distinct subtopics: *barriers* in July and *immigration* in October. They corresponded to the Biden administration's opposition to floating barriers at the U.S.-Mexico border and the peak of immigration waves at the border in October, respectively. Bob found that the timing of intervention was important. Early intervention in the *barriers* topic could sway public sentiment (Fig. 6B_1), leading to more positive sentiments and not negatively impacting the primary election outcomes. Late interventions of *waves of immigration*, however, were less effective. Bob believed that the early negative sentiment had already spread to numerous other topics. Consequently, this strategy of late intervention did not succeed.

6.2 User Study

To evaluate the usability of the system (G1) and to understand whether the specific design choice of using the x-axis to represent time facilitated the design needs (G2), we designed a user study.

6.2.1 Baseline

Traditional sentiment contagion analysis [57] relies on statistical tests, which are time-consuming. Comparing visualization methods with such approaches is not meaningful. We analyzed existing map-based social media data visualization works [7–10]. We found that current visualizations all use animation to depict temporal changes. While animation can use more frames to show additional details, it also increases the user's cognitive load [50]. We aim to explore whether animation or the x-axis encoding is better suited for sentiment contagion analysis.

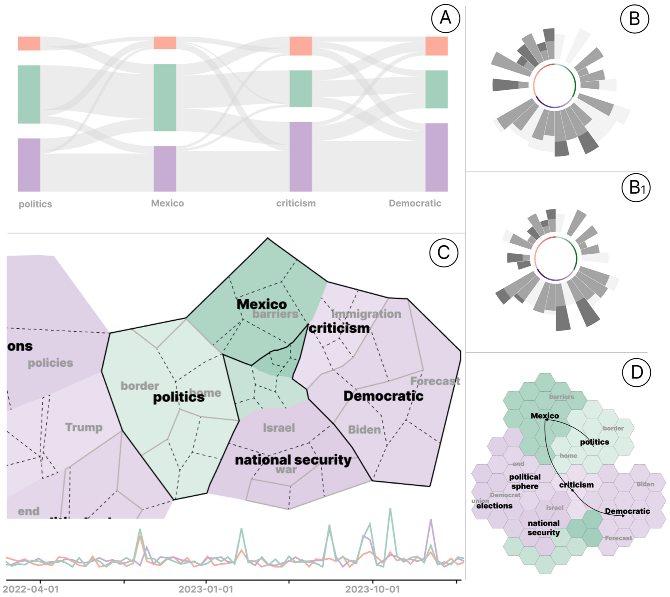

Fig. 6: (A) The sentiment evolution from *politics* topic to *Democratic* topic. (B) The glyphs of the *Democratic* topic following the what-if analysis. (C)The major sentiment contagion pathway between these two topics. (D) Main view of D-Map, the baseline used in the comparative user study.

However, these works focus on visualizing retweet relationships rather than sentiment contagion or causal links. Therefore, we cannot directly use these systems as baselines. In our study, we replaced the overview view of CausalMap to conduct the baseline system. For the baseline overview, we adopted the main view designs from D-Map (Fig. 6D). In D-Map, each topic has only one color, with subtopics as regular hexagons. D-Map doesn't use obvious x or y axes to map time. Instead, D-Map employs animation to illustrate the progression of sentiment spread. In each frame, transparency is utilized to highlight messages published simultaneously, allowing users to pause at any frame for in-depth exploration. Relationships between topics are indicated by arrows, which, in our baseline system, are repurposed to depict sentiment contagion spread. We kept other views to maintain the system's functionality and analytic capability for sentiment contagion. We also considered other map-based systems, such as R-Map and E-Map. However, R-Map requires distinguishing the importance of different topics, which is not part of our model, while E-Map uses color to encode time, preventing us from using color to express sentiment. The designs in D-Map can effectively visualize the propagation of sentiment over time within topics. Additionally, the input format is consistent with that of CausalMap, requiring only topics and the edges between them. Therefore, we chose D-Map as the baseline.

6.2.2 Participants and Study Setup

We invited 12 participants to take part in the study, all of whom have more than two years of experience in online public sentiment analysis. They all hold at least a bachelor's degree. The age range of participants is from 22 to 47 years old (mean = 29.0), with 3 females and 9 males. Their daily work involves analyzing public sentiment using Python and Excel, and none of them had prior experience with CausalMap or any map-based visual analytics systems. Overall, these participants possess basic knowledge in sentiment contagion analysis tasks. We can evaluate our hypotheses through their task completion times and feedback. The participants were invited to conduct the experiment in a laboratory setting. As this experiment was part of their collaborative work on research projects, no direct compensation was provided.

We used both CausalMap (S1) and the baseline system (S2) for the experiment, with the President dataset (D1) and the NBA dataset (D2) in Sec. 6.1.1. The 12 participants had no professional background in US presidential elections or the NBA. Four experimental conditions were constructed: [S1D1, S2D2], [S1D2, S2D1], [S2D1, S1D2], and [S2D2, S1D1]. The condition [S1D1, S2D2] represents completing the task first using system S1 on dataset D1, and then using system S2 on dataset D2. Participants were randomly and evenly distributed across these four conditions to complete the study.

6.2.3 Procedure and Task

Participants were first asked to watch two videos, each demonstrating the modules and interactions of S1 and S2 systems. After familiarizing themselves with the basic operation of both systems, participants were then given a training task using a small dataset of 1000 posts related to nuclear wastewater. Participants were free to explore the training system and watch the tutorial videos until they felt confident in their understanding of the system's interactions. Following the training, experiments were conducted according to the four group conditions.

The tasks were divided into two stages to address the visual analysis requirements. In the first stage, the tasks for both datasets were the same and consisted of three steps: (1) Select and summarize a contagion pathway from a narrative perspective. (2) Identify the main sentiment transformations and the demographic information of the primary spreading population within the pathway. (3) Analyze what strategies could be employed to transform the sentiment of the last topic in the pathway into positive/negative. In the second stage, we aimed to encourage more in-depth, exploratory analysis. For the presidential dataset, participants were asked to explore "Summarize the impact of the open border policy on Biden's election, and how voter attitudes toward Biden would change if this impact were ignored." For the NBA dataset, participants were asked to explore "The impact of Jokic's FMVP on Wembanyama's selection as the first overall pick, and how fan attitudes toward Wembanyama would change if this impact were ignored." We used the following methods to determine whether users had completed their tasks. For the descriptive tasks in the first step of the first stage and the second stage, participants were considered to have successfully completed the task once they summarized factually accurate answers to their satisfaction. For the second and third steps of the first stage, we compared their answers with the correct answers. Finally, participants were asked to fill out System Usability Scale (SUS) questionnaires and participate in a semi-structured interview to discuss the usability and efficiency differences between S1 and S2.

6.2.4 Results

We use the average SUS scores for the S1 system to evaluate the system's usability (G1) and compare task completion times on S1 and S2 systems to assess whether our design improved analysis efficiency and better met the design requirements (G2).

First, we analyze whether the datasets D1 and D2 had a noticeable impact on task completion time and SUS scores using a paired t-test (two-tailed). The results ($p = 0.279$ on time and $p = 0.185$ on SUS) indicate that the choice of dataset had no significant effect on the task completion time or SUS scores. For G1, the SUS score for S1 was 78.75, which indicates that the system is usable and can assist analysts in completing tasks. For G2, we hypothesized that the task completion time on S1 would be significantly shorter than on S2. After performing the t-test (one-tailed), the results ($p = 0.007$) demonstrate that the S1 design indeed offers superior efficiency.

Beyond quantitative experiments, we also conduct qualitative analyses. Through participant interviews, we explored the differences between S1 and S2, specifically the difference in time encoding using animation versus the x-axis. Users found both S1 and S2 to be user-friendly, with no significant difference ($p = 0.154$, two tailed t-test) in SUS scores between them. However, there was a notable difference in task completion time. Users noted that the exploration animations required more interaction, slowing down their analysis. We recorded the number of interactions, including clicking, scrolling, and dragging, in the overview required to complete the tasks. The number of interactions in S2 was significantly higher ($p < 0.001$) than in S1 using a paired t-test (two-tailed). Users found it challenging to view all sentiment information across time in S2, as they had to switch between different time segments. S2 used transparency to mark the current time versus past times, but this meant that even if users were aware of a block's exact time, the block's transparency changed when they viewed other times, making it harder to track. In contrast, S1 allowed users to see the time of all blocks directly, presenting more information in one view.

DATASET	METHOD	RECALL	PRECISION	F1
NBA	**CausalMap**	$\mathbf{0.63 \pm 0.04}$	$\mathbf{0.80 \pm 0.10}$	$\mathbf{0.70 \pm 0.06}$
	Baseline	0.88 ± 0.04	0.54 ± 0.05	0.67 ± 0.04
PRESIDENT	**CausalMap**	$\mathbf{0.65 \pm 0.08}$	$\mathbf{0.77 \pm 0.02}$	$\mathbf{0.70 \pm 0.04}$
	Baseline	0.84 ± 0.05	0.48 ± 0.02	0.61 ± 0.03

Table 1: Performance comparison between the algorithm part of CausalMap and the Baseline [56].

Additionally, the x-axis time representation in S1 allowed each contagion path to correspond to the sequence of topics from left to right. In S2, the sequence of events was indicated by arrows, with no clear spatial representation of the order of events. This created a cognitive burden for users. However, two users noted that the x-axis encoding design may counter users' intuition. This led them to feel that using our system required some learning, which is reflected in their scores of 4 and 5 on the "need to learn a lot" question. Overall, the use of the x-axis to represent time in a static layout conveyed more information, reducing the need for user interactions and easing the cognitive load. This design led to increased analysis efficiency.

6.3 Algorithm Evaluation

To evaluate our causality-based algorithm, we conducted a quantitative experiment. Our algorithm requires multiple unmasked posts from the same user. Therefore, most public datasets [32, 42] were unsuitable as they either were anonymized or only focused on comment/retweet structures. We used the two datasets from the case studies, including the NBA dataset and the President dataset. Three experts collaboratively labeled the information propagation paths within these datasets as ground truth, concentrating solely on the top-level topics generated by LDA. For comparison, we surveyed social media information propagation reconstruction algorithms [33]. To ensure fairness, we chose an explainable algorithm [56] that reconstruct the networks using both post and user features as the baseline. The results are shown in Tab. 1.

From the results, our method trades recall for precision, yielding a slightly higher F1. This difference likely arises because F-GES's calculation objective includes a structural-complexity penalty, making the results more precise. In contrast, the feature-based method doesn't dynamically adjust previously solved edges based on newly solved ones, leading to more edges in the output and thus higher recall. Overall, our algorithm performs well on real-world datasets. Moreover, the information propagation networks it solves can be better explained through the causal effects between the sentiment distributions of topics.

6.4 Expert Interview

To collect expert feedback, we conducted semi-structured interviews with three experts (E1-3). E1 and E2 are senior managers leading development teams for public sentiment monitoring and analysis systems, with responsibility for system implementation in regional agencies. E3 is a professor specializing in the field of sentiment contagion. None of them were involved in the system's design process.

E1 praised the Overview, stating that the map provided a crucial big picture for understanding the contagion, allowing for a quick grasp of the dataset. The design of zoom-in and post views also ensured that details were accessible. E2 concurred, noting that in their daily work topics are typically stored in textual tables, and our zoom-enabled map design could significantly enhance the efficiency. E2 also mentioned that the map's temporal layout allows for an intuitive understanding of topic evolution and the identification of propagation pathways of interest. Regarding the Interpretation view, E3 found these views well-suited for identifying spreading conditions in complex contagion scenarios within their field and for gaining a clearer understanding of different user characteristics. Both E2 and E3 highlighted the what-if analysis as particularly appealing. E3 remarked that conducting predictive research in their field is complex, and our what-if method could help identify directions for future intervention research. E2 noted that the system aligns well with their workflow, from detecting phenomena, explaining them, and intervening them. He believed the system could seamlessly integrate into their work. However, E1 also pointed out that while the what-if analyses are excellent, they are still predictions and do not equate to real-world interventions. He proposed future directions, such as controlling the training datasets of LLMs to create an intervened persona [28]. For example, simulating a world without COVID-19 through LLM without related training datasets to observe their perspectives on various events.

7 Discussion and Future Work

This section presents the implications, limitations, and future work.

7.1 Implications

This research is the first step toward the causality-assisted visual analysis of sentiment contagion network. Analysts can efficiently understand the large spreading network of sentiment in hierarchical topics, discover the impact of different demographic groups, and conduct what-if analyses to explore strategies.

Techniques. This study first introduces a method to model sentiment contagion and proposes a causality-based model to detect this structure from the social media dataset. Then this study introduces a layout algorithm to encode time axis in the map-like visualization to visualize the time-aware hierarchical sentiment contagion network.

Applicability. From the data processing perspective, our approach can be extended to any dataset with repeated user activities recorded in textual format. For instance, it can be applied to shopping review data to infer users' preferences for future purchases, or to medical records to predict potential diseases that may require prevention [35]. Once such data is transformed into a causal network, it can be directly analyzed using CausalMap. From the visualization perspective, CausalMap provides a method for displaying large-scale graph with directionality. This makes it suitable for visualizing large-scale causal graphs in causality analysis [51], or depicting large-scale time-series sensor data in urban data visualization [14, 38]. The previous map-based social media visualizations [9, 10] can also be easily adapted to these new designs.

7.2 Limitation and Future Work

This study has three limitations that present opportunities for future work. Firstly, our method may not yield accurate results when user posts on multiple topics are unavailable or when there is minimal user engagement on a particular topic. This is unavoidable when sampling and reasoning about sentiment distributions of individual users across topics One possible solution is to cluster similar types of users and examine the sentiment distributions of these clusters.

Secondly, our visualization method requires that a topic should not have too many subtopics simultaneously. Therefore, we have imposed a hard constraint on the number of subtopics under each topic. This reduces the system's scalability. However, considering the hierarchical design of subtopics within a topic, we have somewhat compensated for scalability by adding more levels and leveraging the map's infinite divisibility. Moreover, in the analysis of actual datasets, the optimal clustering quantity for most clusters does not exceed the constraint, so it rarely triggers the constraint. Consequently, it mainly serves as a boundary design to ensure system usability in rare cases.

Lastly, our system requires precomputing the causal network before visualization and interaction. The computation bottleneck of our LLM-based solution lies in the LLM request time. However, LLM is only used in the network generation phase, not in the analysis phase. Thus, with adequate preprocessing, our system can ensure smooth interaction. Developers can also improve the speed of calculating the sentiment contagion network by implementing parallelized query processing.

8 Conclusion

This paper introduces a visual analytics system designed to analyze sentiment contagion across social media topics. By employing a causality-based method, the system efficiently uncovers the time-aware hierarchical network of sentiment contagion, overcoming the scalability limitations of traditional methods. The system's map-like visualization provides a clear overview of sentiment flow and topic evolution over time. The interpretation view is also integrated with demographic information and sentiment transformation, thereby assisting analysts in gaining a comprehensive understanding. Validated through a user study and an expert interview, the system demonstrates its utility in social decision-making. Future research will focus on refining the model to bridge the gap from what-if analyses to actual interventions.

Acknowledgments

The work was supported by National Key R&D Program of China (2023YFB3107100), Zhejiang Provincial Natural Science Foundation of China under Grant No. LD25F020003, and NSFC (62421003, U22A2032, 62402437).

References

[1] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, et al. GPT-4 Technical Report. *arXiv preprint arXiv:2303.08774*, 2023. doi: 10.48550/arXiv.2303.08774 3

[2] J. Alvarez-Galvez. Network Models of Minority Opinion Spreading: Using Agent-Based Modeling to Study Possible Scenarios of Social Contagion. *Social Science Computer Review*, 34(5):567–581, 2016. doi: 10.1177/0894439315605607 1

[3] J. Bae, E. Ventocilla, M. Riveiro, T. Helldin, and G. Falkman. Evaluating Multi-Attributes on Cause and Effect Relationship Visualization. In *Proceedings of the International Conference on Information Visualization Theory and Applications*, pp. 64–74, 2017. doi: 10.5220/0006102300640074 2

[4] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3(Jan):993–1022, 2003. doi: 10.5555/944919.944937 3

[5] N. Cao, Y.-R. Lin, X. Sun, D. Lazer, S. Liu, and H. Qu. Whisper: Tracing the Spatiotemporal Process of Information Diffusion in Real Time. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2649–2658, 2012. doi: 10.1109/TVCG.2012.291 5

[6] D. Centola and M. Macy. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, 113(3):702–734, 2007. doi: 10.1086/521848 1, 2

[7] S. Chen, S. Chen, L. Lin, X. Yuan, J. Liang, and X. Zhang. E-Map: A Visual Analytics Approach for Exploring Significant Event Evolutions in Social Media. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 36–47, 2017. doi: 10.1109/VAST.2017.8585638 2, 7

[8] S. Chen, S. Chen, Z. Wang, J. Liang, Y. Wu, and X. Yuan. D-map+: Interactive Visual Analysis and Exploration of Ego-centric and Event-centric Information Diffusion Patterns in Social Media. *ACM Transactions on Intelligent Systems and Technology*, 10(1):1–26, 2018. doi: 10.1145/3183347 2, 7

[9] S. Chen, S. Chen, Z. Wang, J. Liang, X. Yuan, N. Cao, and Y. Wu. D-Map: Visual Analysis of Ego-Centric Information Diffusion Patterns in Social Media. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 41–50, 2016. doi: 10.1109/VAST.2016.7883510 1, 5, 6, 7, 9

[10] S. Chen, S. Li, S. Chen, and X. Yuan. R-Map: A Map Metaphor for Visualizing Information Reposting Process in Social Media. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1204–1214, 2019. doi: 10.1109/TVCG.2019.2934263 1, 2, 7, 9

[11] Z. Cheng, Z. Lin, Y. Yang, Z. Wei, and S. Chen. Interactive Simulation and Visual Analysis of Social Media Event Dynamics with LLM-Based Multi-Agent Modeling. *Visual Informatics*, p. 100260, 2025. doi: 10.1016/j.visinf.2025.100260 1

[12] J. Choo, C. Lee, C. K. Reddy, and H. Park. UTOPIAN: User-Driven Topic Modeling Based on Interactive Nonnegative Matrix Factorization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):1992–2001, 2013. doi: 10.1109/TVCG.2013.212 3

[13] W. Cui, S. Liu, Z. Wu, and H. Wei. How Hierarchical Topics Evolve in Large Text Corpora. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2281–2290, 2014. doi: 10.1109/TVCG.2014.2346433 2

[14] Z. Deng, D. Weng, 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*, 28(1):1051–1061, 2022. doi: 10.1109/TVCG.2021.3114875 9

[15] P. S. Dodds. A Simple Person's Approach to Understanding the Contagion Condition for Spreading Processes on Generalized Random Networks. *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks*, pp. 27–45, 2018. doi: 10.1007/978-3-319-77332-2_2 2

[16] P. S. Dodds. Slightly Generalized Contagion: Unifying Simple Models of Biological and Social Spreading. *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks*, pp. 67–80, 2018. doi: 10.1007/978-3-319-77332-2_4 1

[17] W. Dou, X. Wang, D. Skau, W. Ribarsky, and M. X. Zhou. LeadLine: Interactive Visual Analysis of Text Data Through Event Identification and Exploration. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 93–102, 2012. doi: 10.1109/VAST.2012.6400485 2

[18] W. Dou, L. Yu, X. Wang, Z. Ma, and W. Ribarsky. HierarchicalTopics: Visually Exploring Large Text Collections Using Topic Hierarchies. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2002–2011, 2013. doi: 10.1109/TVCG.2013.162 2

[19] M. Fan, J. Yu, D. Weiskopf, N. Cao, H.-Y. Wang, and L. Zhou. Visual Analysis of Multi-outcome Causal Graphs. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–11, 2024. doi: 10.1109/TVCG.2024.3456346 3

[20] J. Feng, K. Wu, and S. Chen. TopicBubbler: An Interactive Visual Analytics System for Cross-Level Fine-Grained Exploration of Social Media Data. *Visual Informatics*, 7(4):41–56, 2023. doi: 10.1016/j.visinf.2023.08.002 2

[21] E. R. Gansner, Y. Hu, and S. Kobourov. GMap: Visualizing Graphs and Clusters as Maps. In *Proceedings of the IEEE Pacific Visualization Symposium*, pp. 201–208, 2010. doi: 10.1109/PACIFICVIS.2010.5429590 5

[22] J. P. Gleeson and M. A. Porter. Message-Passing Methods for Complex Contagions. *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks*, pp. 81–95, 2018. doi: 10.1007/978-3-319-77332-2_5 1

[23] S. González-Bailón, J. Borge-Holthoefer, A. Rivero, and Y. Moreno. The Dynamics of Protest Recruitment through an Online Network. *Scientific reports*, 1(1):1–7, 2011. doi: 10.1038/srep00197 1

[24] M. Grootendorst. BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure. *arXiv preprint arXiv:2203.05794*, 2022. doi: 10.48550/arXiv.2203.05794 3

[25] D. Guilbeault, J. Becker, and D. Centola. *Complex Contagions: A Decade in Review*, pp. 3–25. Springer International Publishing, Cham, 2018. doi: 10.1007/978-3-319-77332-2_1 1, 2

[26] S. Havre, B. Hetzler, and L. Nowell. ThemeRiver: Visualizing Theme Changes over Time. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 115–123, 2000. doi: 10.1109/INFVIS.2000.885098 2, 5

[27] J. Heer and D. Boyd. Vizster: Visualizing Online Social Networks. In *Proceedings of the IEEE Symposium on Information Visualization*, pp. 32–39, 2005. doi: 10.1109/INFVIS.2005.1532126 1

[28] H. Hosseinmardi, A. Ghasemian, M. Rivera-Lanas, M. Horta Ribeiro, R. West, and D. J. Watts. Causally Estimating the Effect of YouTube's Recommender System Using Counterfactual Bots. *Proceedings of the National Academy of Sciences*, 121(8):e2313377121, 2024. doi: 10.1073/pnas.2313377121 3, 9

[29] A. Kale, Y. Wu, and J. Hullman. Causal Support: Modeling Causal Inferences with Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):1150–1160, 2021. doi: 10.1109/TVCG.2021.3114824 2

[30] W. O. Kermack and A. G. McKendrick. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, 115(772):700–721, 1927. doi: 10.1098/rspa.1927.0118 1

[31] D. Khurana, A. Koli, K. Khatter, and S. Singh. Natural Language Processing: State of the Art, Current Trends and Challenges. *Multimedia tools and applications*, 82(3):3713–3744, 2023. doi: 10.1007/s11042-022-13428-4 3

[32] J. Leskovec and A. Krevl. SNAP Datasets: Stanford Large Network Dataset Collection. `http://snap.stanford.edu/data`, June 2014. 9

[33] M. Li, X. Wang, K. Gao, and S. Zhang. A Survey on Information Diffusion in Online Social Networks: Models and Methods. *Information*, 8(4):118, 2017. doi: 10.3390/info8040118 9

[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*, 30(1):638–648, 2024. doi: 10.1109/TVCG.2023.3327376 2, 3, 4, 5

[35] C. Liu, Y. Chen, T. Liu, M. Gong, J. Cheng, B. Han, and K. Zhang. Discovery of the Hidden World with Large Language Models. In *Proceedings of the Advances in Neural Information Processing Systems*, vol. 37, pp. 102307–102365, 2024. doi: 10.5555/3737916.3741165 9

[36] F. Liu and Y. Deng. Determine the Number of Unknown Targets in Open World Based on Elbow Method. *IEEE Transactions on Fuzzy Systems*, 29(5):986–995, 2021. doi: 10.1109/TFUZZ.2020.2966182 3

[37] S. Liu, X. Wang, J. Chen, J. Zhu, and B. Guo. TopicPanorama: A Full

Picture of Relevant Topics. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 183–192, 2014. doi: 10.1109/VAST.2014.7042494 1, 2

[38] X. Luo, R. Jiang, B. Yang, H. Qin, and H. Hu. Air Quality Visualization Analysis Based on Multivariate Time Series Data Feature Extraction. *Journal of Visualization*, 27(4):567–584, 2024. doi: 10.1007/s12650-024-00981-3 9

[39] P. Parigi and R. Gong. From Grassroots to Digital Ties: A Case Study of a Political Consumerism Movement. *Journal of Consumer Culture*, 14(2):236–253, 2014. doi: 10.1177/1469540514526280 1

[40] J. Pearl et al. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2009. doi: 10.5555/1642718 2

[41] 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*, 3(2):121–129, 2017. doi: 10.1007/s41060-016-0032-z 1, 4

[42] A. Rogers, A. Romanov, A. Rumshisky, S. Volkova, M. Gronas, and A. Gribov. RuSentiment: An Enriched Sentiment Analysis Dataset for Social Media in Russian. In *Proceedings of the International Conference on Computational Linguistics*, pp. 755–763, 2018. 9

[43] J. N. Rosenquist, J. Murabito, J. H. Fowler, and N. A. Christakis. The Spread of Alcohol Consumption Behavior in a Large Social Network. *Annals of Internal Medicine*, 152(7):426–433, 2010. doi: 10.7326/0003-4819-152-7-201004060-00007 1

[44] G. Sun, Y. Wu, S. Liu, T.-Q. Peng, J. J. Zhu, and R. Liang. EvoRiver: Visual Analysis of Topic Coopetition on Social Media. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1753–1762, 2014. doi: 10.1109/TVCG.2014.2346919 2

[45] J. Wang and K. Mueller. The Visual Causality Analyst: An Interactive Interface for Causal Reasoning. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):230–239, 2016. doi: 10.1109/TVCG.2015.2467931 2

[46] J. Wang and K. Mueller. Visual Causality Analysis Made Practical. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 151–161, 2017. doi: 10.1109/VAST.2017.8585647 2

[47] Q. Wang, K. Xu, and R. S. Laramee. Demers Cartogram with Rivers. *Visual Informatics*, 8(3):57–70, 2024. doi: 10.1016/j.visinf.2024.09.003 5

[48] X. Wang, W. Dou, Z. Ma, J. Villalobos, Y. Chen, T. Kraft, and W. Ribarsky. I-SI: Scalable Architecture for Analyzing Latent Topical-Level Information from Social Media Data. *Computer Graphics Forum*, 31(3pt4):1275–1284, 2012. doi: 10.1111/j.1467-8659.2012.03120.x 2

[49] X. Wang, S. Liu, Y. Chen, T.-Q. Peng, J. Su, J. Yang, and B. Guo. How Ideas Flow across Multiple Social Groups. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*, pp. 51–60, 2016. doi: 10.1109/VAST.2016.7883511 5

[50] P. Wouters, F. Paas, and J. J. van Merriënboer. How to Optimize Learning from Animated Models: A Review of Guidelines Based on Cognitive Load. *Review of Educational Research*, 78(3):645–675, 2008. doi: 10.3102/0034654308320320 7

[51] 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*, 27(2):1448–1458, 2021. doi: 10.1109/TVCG.2020.3028957 2, 4, 5, 9

[52] C. Xiong, J. Shapiro, J. Hullman, and S. Franconeri. Illusion of Causality in Visualized Data. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):853–862, 2020. doi: 10.1109/TVCG.2019.2934399 2

[53] O. Yağan and V. Gligor. Analysis of Complex Contagions in Random Multiplex Networks. *Physical Review E*, 86(3):036103, 2012. doi: 10.1103/PhysRevE.86.036103 1

[54] A. Yang, A. Li, B. Yang, B. Zhang, B. Hui, B. Zheng, B. Yu, C. Gao, C. Huang, C. Lv, et al. Qwen3 Technical Report. *arXiv preprint arXiv:2505.09388*, 2025. doi: 10.48550/arXiv.2505.09388 3

[55] C.-H. E. Yen, A. Parameswaran, and W.-T. Fu. An Exploratory User Study of Visual Causality Analysis. *Computer Graphics Forum*, 38(3):173–184, 2019. doi: 10.1111/cgf.13680 2

[56] J. Yin, H. Jia, B. Zhou, T. Tang, L. Ying, S. Ye, T.-Q. Peng, and Y. Wu. Blowing Seeds Across Gardens: Visualizing Implicit Propagation of Cross-Platform Social Media Posts. *IEEE Transactions on Visualization and Computer Graphics*, 31(1):185–195, 2025. doi: 10.1109/TVCG.2024.3456181 9

[57] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin. A Survey of Sentiment Analysis in Social Media. *Knowledge and Information Systems*, 60:617–663, 2019. doi: 10.1007/s10115-018-1236-4 2, 3, 7

[58] X. Zhang, X. Yang, H. Hu, and H. Qin. Visual Causal Analysis of Multivariate Time Series. *Journal of Visualization*, 28(2):323–339, 2025. doi: 10.1007/s12650-024-01041-6 2